

Left Branch Extraction in Lower Sorbian

ANDREAS PANKAU
Free University Berlin

ABSTRACT

Corver (1990) and Bošković (2005) note that the availability of Left Branch Extraction – the extraction of material to the left of the head out of DPs, APs, and PPs – in a language correlates with the presence and absence of definite articles. Languages with definite articles do not allow Left Branch Extraction (like English), languages without definite articles do allow Left Branch Extraction (like Polish). They further suggest that the two phenomena are causally linked, claiming Left Branch Extraction in languages with definite articles is blocked by the DP projection. In this paper, I show that the presence of definite articles does not imply the absence of Left Branch Extraction. The evidence comes from Lower Sorbian, an endangered Slavic language spoken in South East Germany. I show that Lower Sorbian has both definite articles and Left Branch Extraction.

KEYWORDS Left Branch Extraction · DP/NP generalizations · articles · definiteness · Lower Sorbian

1 INTRODUCTION

Many languages are subject to the Left Branch Condition (Ross 1967/1986), which bars extraction of elements that appear on the left side within some constituent, for example determiners in DPs, degree adverbs in APs, or adjectives in DPs. English is such a language. It is impossible to extract a wh-determiner out of a DP, a degree adverb out of an AP, or an attributive adjective out of a DP.

- (1) a. *Which_i did you see [t_i car]?
 b. *Terribly_i he is [t_i tired].
 c. *Interesting_i he read [t_i books].

As Ross himself noted, the Left Branch Condition is not universal. Many languages allow extractions that violate this constraint; this type of extraction is called LEFT BRANCH EXTRACTION (LBE). Slavic languages are among the languages with LBE. In Polish for example it is fine to extract wh-determiners, degree adverbs, and attributive adjectives, as shown in (2).

- (2) a. Które_i widziałeś [t_i auto]?
 which saw.2.SG auto
 ‘Which car did you see?’ (Rappaport, 2000, 165)
 b. Okropnie_i on był [t_i zmęczony].
 terribly he was tired
 ‘He was terribly tired.’ (Talić, 2017, 419)
 c. Ciekawe_i on czytał [t_i książki].
 interesting he read books
 ‘He read interesting books.’ (Cegłowski, 2017, 346)

Standardly, the availability of LBE in a language is assumed to depend on the presence of a definite article (Corver, 1990; Bošković, 2005). In particular, the implication in (3) is assumed to hold cross-linguistically.

- (3) If a language L has a definite article, then L does not have LBE

This implication allows three types of languages: languages that have definite articles and lack LBE

(e.g. English); languages that lack definite articles and have LBE (e.g. Polish); and languages that lack both definite articles and LBE (e.g. Estonian). What is excluded under (3) is a language that has a definite article and allows LBE. The aim of this paper is to show that Lower Sorbian is exactly the type of language predicted not to exist: Lower Sorbian has both a definite article and allows LBE. Lower Sorbian therefore not only presents a clear counterexample to the implication in (3), it also provides evidence against any analysis that links the presence of a definite article to the unavailability of LBE.

The paper is organized as follows. After a brief introduction to Lower Sorbian in section 2, I will show in section 3 that Lower Sorbian allows LBE in all its flavors. In section 4, I show that Lower Sorbian possesses a definite article.

2 LOWER SORBIAN

Lower Sorbian (Iso. *dolnoser(b)ski*) is a West Slavic minority language indigenous to the south east of the German province Brandenburg. Nowadays, Lower Sorbian is spoken only in a few villages to the north and north east of Cottbus (Iso. *Chóšebuz*), namely in Dissen (Iso. *Dešno*), Striesow (Iso. *Strjažow*), Drachhausen (Iso. *Hochoza*), Schmogrow (Iso. *Smogorjow*), Fehrow (Iso. *Prjawoz*), Drehnow (Iso. *Drjenow*), Turnow (Iso. *Turnow*), Tauer (Iso. *Turjej*), Peitz (Iso. *Picnjo*), Jänschwalde (Iso. *Janšojce*), Bärenbrück (Iso. *Barbuk*), and Heinersbrück (Iso. *Móst*). Lower Sorbian is highly endangered: the language stopped to be transmitted intergenerationally after World War II. Only a few native speakers are left (approximately 200), all of whom are bilingual (Lower Sorbian, German) and mostly older than 80. The use of Lower Sorbian is restricted to private conversations and native speakers are usually reluctant to use it with strangers.

The data reported in this paper come from three sources. The first source are data collected in two meetings with a native speaker of Lower Sorbian in Berlin and Cottbus in November and December 2018. The second source are data from the publically available DoBeS corpus (*Dokumentation Bedrohter Sprachen*, documentation of endangered languages), which can be accessed either through the webpage of the Max-Planck institute in Nijmegen (bit.ly/2OCiK8g, after registration) or through the webpage of the Sorbian Institute in Cottbus (bit.ly/2OC5qAA, without registration, but with limited search options). The third source are written records of colloquial Lower Sorbian.¹ Data from the native speaker receive no special indication. Data from the DoBeS corpus are always indicated by ‘DoBeS’, followed by (i) the place of the recording, (ii) the date of the recording, and (iii) the name of the file the sentence is taken from. Data from written sources are indicated by reference to the publication.

3 LEFT BRANCH EXTRACTION IN LOWER SORBIAN

All West Slavic languages allow LBE (Corver, 1990; Blum, 2013). Lower Sorbian is no exception to this generalization. There are three subtypes of LBE: INTERROGATIVE, ADVERBIAL, and ADJECTIVAL LBE (illustrated in (2-a)-(2-c), respectively).² Lower Sorbian possesses all three types.

The first subtype, interrogative LBE, is illustrated in (4).

- (4) a. Wšykne su chwalili, **kaku**_i my mamy [t_i dobru wódu].
 all are praised what we have good water
 ‘All were praising what good water we have.’

(DoBeS, Fehrow, MEW-2015-07-07, MEW-200)

¹The restriction on colloquial Lower Sorbian is important. Although there exist many publications in Lower Sorbian, these are mainly published in a standardized version of Lower Sorbian. This standardized variety, however, differs from the colloquial variety on all levels of grammar, not only lexically. The differences can be drastic. For example, whereas written Lower Sorbian possesses relative pronouns (based on interrogative pronouns), relative pronouns are completely absent from the colloquial language, which employs a relative particle (*ak* or *ako*) and resumptive pronouns instead.

²The distinction between these three types is based on the following observations. First, interrogative LBE is also fine in German and Dutch wh-exclamatives (Corver, 1990). Second, adverbial LBE is also found in languages that allow no other type of LBE, like Icelandic (Talić, 2017). Third, adjectival LBE is a marked option even for languages that allow all other types of LBE (Fanselow & Féry, 2013).

- b. Wěš, **wjele**_i ja som zachadne lěto [t_i malinow] měl?
 know.2.SG how.many I am previous year raspberries had
 ‘Do you know how many raspberries I had last year?’
 (DoBeS, Striesow, FKT-2015-04-15, FKT-001)

In the domain of interrogative LBE, one observes three effects reported for other Slavic languages. First, extraction of a D-Linked wh-determiner is degraded (Blum, 2013) in contrast to extraction of a non-D-Linked wh-determiner.

- (5) **Kaki**_i / ?**kótry**_i jo rozpadnuł [t_i dom]?
 what which is collapsed house
 ‘What/which house collapsed?’

Second, Lower Sorbian allows EXTRAORDINARY LBE, by which Bošković (2005, sec. 6) refers to the effect that, when a wh-determiner is extracted out of a DP that is the complement of a preposition, the preposition is pied piped. Similarly to other Slavic languages, Lower Sorbian allows extraordinary LBE, too, as shown in (6).

- (6) a. [**Na kajke**]_i wón jo skocył [t_i kšywo]?
 on what he is jumped roof
 ‘On what roof did he jump?’
 b. [**W kakich**]_i ty kul_i [t_i cřejach] chóžiš?
 in what you PRT shoes walk.2.SG
 ‘In what shoes are you walking?’ (Schwela, 1911, 26)

Third, Lower Sorbian disallows DEEP LBE, by which Bošković (2005, 8) refers to LBE that targets the left branch of a constituent that is a complement to a noun. Deep LBE is generally ungrammatical in Slavic languages, and it hence comes as no surprise that deep LBE is out in Lower Sorbian as well, as shown by example (7).

- (7) ***Kótreje**_i wón jo wizeł [pšijašela [t_i mamy]]?
 which he is seen friend mother
 Intended: ‘The friend of which mother did he see?’

The second subtype, adverbial LBE, is illustrated in (8).

- (8) **Tak**_i jo była ta wójna how [t_i šlimna].
 so is been the war here horrible
 ‘The war was so horrible here.’ (DoBeS, Heinersbrück, MEW-2012-04-24, MEW-045)

Adverbial LBE is not restricted to movement to SpecCP, as in (8), but can also result from scrambling, as the following sentence illustrates.

- (9) Nět jo ten rotnik **tak**_i był [t_i zły], až wón jo jogo pšašal.
 now is the gatekeeper so been bad that he is him asked
 ‘Now the gatekeeper was so angry that he asked him.’ (von Schulenburg, 1930, 152)

Adverbial LBE is much more common than interrogative LBE. There are approximately 20 instances of adverbial LBE in the DoBeS corpus, but only 4 instances of interrogative LBE. My informant, however, judged both types of LBE as equally good and normal. The contrast is then most likely a side effect of the elicitation method used for DoBeS, namely elicitation of narrations by a single informant. In such a context, questions are unlikely to be used by the informant.

As for the third subtype, adjectival LBE, given in (10), my informant accepted it, but considered it a marked option.³

³It is basically impossible to evaluate this judgment with the DoBeS because the DoBeS corpus is not tagged. That is, the words lack all sort of grammatical information, including their word class. One can therefore not search for adjectives in general, but only for specific adjectives.

- (10) ?**Rědny**_i wón ma [t_i dom].
 beautiful he has house
 ‘He has a beautiful house.’

Similar to other Slavic languages, Lower Sorbian allows DOUBLE AP LBE. By this, Bošković (2005, 12) refers to the extraction of an adjective out of a DP with more than one adjective. This is possible only when the extracted adjective is in some way featurally distinct from the adjective that is not extracted. Witness the contrast in (11).

- (11) a. **Wjele**_i su byl_i tam [t_i młode žowća].
 many are been there young girls
 ‘Many young girls have been there.’ (DoBeS, Peitz, MEW-2012-02-04, MEW-041)
 b. ***Drogotne**_i wón ma [t_i rědne pyšnotki].
 precious he has beautiful jewelry
 Intended: ‘He has precious beautiful jewelry.’

In both examples, there is a DP with two adjectives, of which the first is extracted. But only (11-a) is grammatical. According to Bošković’s (2005) analysis, the relevant factor distinguishing (11-a) from (11-b) is that in (11-a) the adjective *wjele* is featurally distinct from the adjective *młode*, since *wjele* bears some quantifier-like feature.⁴ In (11-b), however, the two adjectives *drogotne* and *rědne* are plain attributive adjectives, and hence not featurally distinct from each other. Crucially, when *drogotne* in (11-b) is focused, the sentence becomes much better because then, *drogotne* bears a feature that *rědne* does not bear, namely [+focus], as shown in (12).

- (12) ?**DROGOTNE** wón ma [t_i rědne pyšnotki].
 precious he has beautiful jewelry
 ‘He has PRECIOUS beautiful jewelry.’

There is one property that sets Lower Sorbian apart from all other Slavic languages: it allows adjectival LBE of more than one attributive adjective. This option is usually out in other Slavic languages, for example in Serbo-Croatian (13-a) and Bulgarian (13-b).

- (13) a. ***[Visoke lijepe]**_i on gleda [t_i djevojke].
 tall beautiful he watches girls
 Intended: ‘He is watching tall beautiful girls.’ (Bošković, 2005, 12)
 b. ***[Malki-te žälti]**_i prodava [t_i kotki].
 small-DEF yellow sells cats
 Intended: ‘He sells the small yellow cats.’ (Stojković, 2019, ex.13a)

In Lower Sorbian, however, extracting two adjectives is as good as extracting only one.

- (14) ?**[Drogotne rědne]**_i wón ma [t_i pyšnotki].
 precious beautiful he has jewelry
 ‘He has precious beautiful jewelry.’

Importantly, this applies only to adjectives: extracting a determiner together with an attributive adjective is completely ungrammatical.

- (15) ***[Jaden / ten rědny]**_i wón ma [t_i dom].
 one the beautiful he has house
 Intended: ‘He has a/the beautiful house.’

⁴The status of *wjele* as an adjective in (11-a) is confirmed by the case marking of *młode žowća*, which is nominative plural. If *wjele* were a quantifier, the genitive plural *młodych žowćow* would be expected. *wjele* can also be used as an adverb meaning ‘often’ in Lower Sorbian, so there is the possibility that *wjele* in (11-a) is not part of the DP at all. This, however, is unlikely the case in (11-a): the speaker talks about the time after WWII and the situation in a village that was no longer German. What surprises the speaker is the high amount of young German girls still living there and not that young German girls would often show up.

In this latter respect, Lower Sorbian differs from Serbo-Croatian, where LBE of a demonstrative and an attributive adjective, as shown in (16-a), is fine, but patterns with Bulgarian, where this type of LBE is impossible, as shown in (16-b).

- (16) a. [Onu staru]_i prodaje [t_i kuću].
 this old sells house
 ‘He is selling that old house.’ (Bošković, 2015, 418)
- b. *[Tezi malki]_i prodava [t_i kotki].
 these small sells cats
 Intended: ‘He sells these small cats.’ (Šimík, 2014, ex 13b)

To conclude this section, Lower Sorbian behaves like a typical Slavic language: it allows all sorts of LBE and it shows all the properties typical of LBE observed in other Slavic languages (with the exception of allowing LBE of more than one attributive adjective).

4 A DEFINITE ARTICLE IN LOWER SORBIAN

I have shown so far that Lower Sorbian possesses LBE. Since Lower Sorbian is a West Slavic language and since all West Slavic languages have LBE, this seems rather trivial. I argue in this section that this is not trivial at all because Lower Sorbian also possesses a set of determiners that are definite articles. The relevant determiners are *ten* (MASC.SG), *ta* (FEM.SG), *to* (NEUT.SG), *tej* (DUAL), and *te* (PLURAL); in the remainder of the paper, I will refer to them simply by *ten*. That Lower Sorbian possesses definite articles, however, is unexpected given the implication in (3). In order to show that *ten* is indeed a definite article, I first argue that its morphosyntactic behavior is not that of an adjective, but that of a separate D°-category. Second, I argue that *ten* is not a demonstrative but a true definite article, because *ten* occurs in contexts typical for definite articles but untypical for demonstratives.

4.1 MORPHOSYNTACTIC PROPERTIES

Corver (1990) and Bošković (2005, 2009) present a number of arguments that definiteness markers in Slavic (demonstratives and possessives according to them) are adjectives but not definite articles, and hence not located in D°. These arguments concern:

- the position of the definiteness marker;
- the inflectional properties of the definiteness marker vis-à-vis adjectives;
- the possibility to stack definiteness markers;
- the occurrence of the definiteness marker in position typical for adjectives;
- the modification properties of adjectives.

I will discuss each argument in turn and conclude that the argument either shows that *ten* in Lower Sorbian is a definite article or that the argument is not sound.

First, *ten* must precede an adjective and must not follow it; moreover, *ten* must not appear after the noun:

- (17) ten řědny dom
 the beautiful house
 ‘the beautiful house’
- a. *řědny ten dom
 beautiful the house
- b. *řědny dom ten
 beautiful house the

equally compatible with an analysis of *ta* sitting in D° and *moja* being an AP. Other languages with definite articles show that this option is attested. In Italian, a language with definite articles, the co-occurrence of an article and a possessive pronoun, as in (21-a), is unproblematic. Even in German, also a language with a definite article, a demonstrative pronoun, which is definite as well, can appear on top of a possessive pronoun, shown in (21-b).

- (21) a. la mia casa
the my house
'my house'
b. dieses mein Haus
this my house
'this house of me'

In a nutshell, the ungrammaticality of (18-b) is due to an idiosyncrasy of English, and the grammaticality of (18-a) is unrelated to the lack of a definite article in Serbo-Croatian. The contrast in (18) shows nothing about the status of a definiteness marker as an A° or a D°, nor does it show that possessive pronouns are D°-elements (cf. also Plank, 1992).

The fourth argument concerns the contrast between English and Serbo-Croatian in (22-a)–(22-b).

- (22) a. Ova knjiga je moja.
this book is my
'This book is mine.'
b. *This book is my.
c. Ten njej był mój.
the not.is been my
'This one was not mine.'

According to Corver (1990, 332) and Bošković (2005, 5), the predicate position after a copula is a "typical adjectival position". As (22-c) shows, Lower Sorbian patterns with Serbo-Croatian. Also this argument is doubtful. On the one hand, both Corver (1990) and Bošković (2005) only show that possessive pronouns occur in predicate positions. This at best shows that possessive pronouns are adjectives; it does not show that articles are adjectives as well. This only follows once possessive pronouns and definiteness markers are equated. But as seen in the discussion surrounding the third argument, this equation does not go through. Note also that a demonstrative in the post-copula position is ungrammatical under a predicative interpretation in Polish, which is unexpected given their apparent status as adjectives.

- (23) *Mój kraj jest ten/tym.
my country is this
'My country is this'

On the other hand, the assumption that the predicate position after a copula is a typical adjective position is wrong. In German, for example, all types of elements can occur there, adjectives, nouns, but also prepositions, and even elements that are restricted to this very position (similar to Engl. 'broke').

- (24) Der Mann ist groß / ein Idiot / zu / pleite.
the man is tall an idiot to broke
'The man is tall/an idiot/drunk/broke.'

To sum up, the fourth argument in itself shows nothing about determiners and rests on the very same false premise as the third argument.

The fifth argument concerns the impossibility of modifying a prenominal possessive that is based on a noun, so-called possessive adjectives. The argument runs as follows. In Serbo-Croatian, a possessive adjective can neither be modified by an adjective, as shown in (25-a), nor by a possessive pronoun, as shown in (25-b).

- (25) a. bogati sused-ov konj
 rich neighbor-ADJ horse
 * 'the horse of the rich neighbor'
 ✓ 'the rich horse of the neighbor'
- b. *Moj brat-ov prijatelj spava.
 my brother-ADJ friend sleeps
 Intended: 'The friend of my brother sleeps.' (Bošković, 2005, 7)

According to Bošković (2005) that possessive pronouns pattern with adjectives concerning their impossibility to modify possessive adjectives is a trivial consequence from the fact that they are both adjectives. In other words, adjectives cannot modify possessive adjectives; and since possessive pronouns are adjectives as well, they are covered by that statement. Lower Sorbian again patterns with Serbo-Croatian.

- (26) To su {*naš / *našogo} nan-owe knjigly.
 that are our.NOM our.GEN father-ADJ books
 'Those are the books of (*our) the father.'

But this fifth argument is problematic as well. First, this argument again only shows that possessive pronouns pattern with adjectives, but not that definiteness markers are adjectives. This only follows once the two are equated. Second, Bošković (2009) fails to properly address Pereltsvaig's (2007) objection that the impossibility to modify a possessive adjective could simply result from a more general constraint that restricts possessive adjective formation to heads only (that is, to simplex nouns), which is what one finds in Russian. Importantly, the same restriction is at work in Lower Sorbian: possessive adjectives can be formed on simplex nouns only (Schwela 1906, 96; Richter 1980, 83; Corbett 1987, fn. 17).⁵ The impossibility of (26) is therefore irrelevant.

To conclude this section, *ten* behaves like a definite article with respect to two of the five diagnostics presented at the beginning of this section; the other three were argued to be no diagnostics for definite articles.

4.2 PRAGMATIC PROPERTIES

In order to show that *ten* in Lower Sorbian is a definite article, one also needs to show that *ten* behaves pragmatically like a definite article. This is what I will do in this section. In particular, I argue that *ten* is a true definite article and not merely a demonstrative determiner.

I adopt standard assumptions about definiteness (Lyons, 1999) according to which the definite article marks the referent of a DP as being familiar and unique. By familiarity, one refers to the requirement that the referent of a definite DP is in some sense given to both interlocutors.

- (27) [Context: out of the blue, no prior mentioning of dogs; A says to B:]
 # The dog attacked me. / ✓ A dog attacked me.

In (27), the use of the definite article is ungrammatical because in an out of the blue context, all DPs are new and not given. By uniqueness, one refers to the requirement that the referent of a definite DP must be the only element within the discourse that satisfies the description of the definite DP.

- (28) [Context: two identical buckets are next to each other; A says to B:]
 # Give me the bucket. / ✓ Give me a bucket.

What goes wrong in (28) is that using *the bucket* is only adequate if exactly one bucket is contextually established; but in this context there are two buckets, and the uniqueness requirement is violated.

⁵Data from the *Sorbischer Sprachatlas vol. 15* (Fasske, 1995) challenge this claim. In this work, a number of examples are given that show that also Lower Sorbian allows possessive adjective formation based on phrases; that is, examples as in (26) are well-formed for many Lower Sorbian speakers. Importantly, there are no examples where *ten* appeared internal to the phrase on which the possessive adjective is formed. This receives an easy treatment once *ten* is treated as a determiner: the relevant constraint would bar D°-elements to occur inside a possessive adjective, but not adjectives.

A typical use of a definite article is that of marking an ANAPHORIC DEFINITE. An anaphoric definite is a DP whose definiteness is established through the discourse: it has already been mentioned before in the discourse and is now referred back to. In Lower Sorbian, anaphoric definites are typically marked by *ten*, as in (29).

- (29) A pón jo kšěł mješ cywo jagły z młokom warjone. A pón jo jana žeńska
and then is wanted have all millet with milk cooked and then is a woman
jomu raz pšinasła te jagły wjelgin warjece.
him once brought the millet very hot
'And he wanted to have all the time cooked millet. And then a woman brought him the
millet very hot.' (Šliziński, 1964, 58)

But examples such as in (29) are inconclusive when it comes to the status of *ten*. On the one hand, also demonstratives can mark a DP as being an anaphoric definite (Hawkins, 1978). The DP in (21-b) for example contains precisely such a demonstrative. On the other hand, even languages without definite articles prefer marking anaphoric definites, namely with demonstratives (Šimík, 2014). In other words, (29) is compatible with an interpretation of *ten* as a definite article but it is also compatible with an interpretation of *ten* as a demonstrative.

Fortunately, the usage conditions of demonstratives and definite articles overlap only partly. Although demonstratives can mark anaphoric definites, there are other types of definite DPs whose definiteness does not depend on the discourse (Hawkins, 1978; Schwarz, 2009). Crucially, with such definites, definite articles are fine but demonstratives are excluded.

The first such type are SITUATIONAL DEFINITES. A situational definite is a DP whose definiteness is supplied by the situation, as in (30).

- (30) [Context: A and B sit in a café but are not served; A says to B:]
✓When will the waiter come? / # When will this waiter come?

The waiter need not have been talked about before to use the definite article here. What matters is that the situation supplies a unique referent: going to a café implies that one is served by exactly one person. Situational definites cannot be marked by demonstratives, neither in languages that possess both articles and demonstratives like English, (30), nor in languages that possess only demonstratives like Czech, as shown by (31). In Czech, situational definites are not marked at all, as in (31-a), (Šimík, 2014, 2).

- (31) Jsem učitelem na škole a bavím se se svým kolegou.
am teacher at school and talk REFL with POSS colleague
'I am a teacher at a school and talk to a colleague of mine.'
a. Ředitelka mě pozvala na kafe.
director me invited for coffee
'The director invited me for coffee.'
b. #Ta ředitelka mě pozvala na kafe.
this director me invited for coffee

Importantly, in Lower Sorbian, *ten* can be used to mark situational definites.

- (32) a. [Context: A and B sit in a café but are not served; A says to B:]
Žo jan wóstanjo ten kelnař?
where only remains the waiter
'Where is the waiter?'
b. Pón jo se ten móst ćepanuł a pón je była ći tom niksnu
then is REFL the bridge collapsed and then is been at the merman
'Then the bridge collapsed and then she was with the merman.' (Šliziński, 1964, 61)

(32-a) is identical to (30). (32-b) also instantiates a situational definite. According to Sorbian

mythology, every lake is inhabited by one, and only one merman. If a bridge collapses, one falls into a lake, and this situation then implies the situational definite ‘the merman.’

The second type are BRIDGING DEFINITES. By bridging definite (or associative definite), Hawkins (1978) refers to DPs that are definite because their referent is inferable from another DP explicitly or implicitly mentioned in the discourse. In (33), *the author* in the second sentence is definite because of the DP *a book* from the first sentence; the inference is that every book has an author.

(33) John bought a book yesterday. The author was French.

Using a demonstrative in such a context is out, as shown in (34-a). Again, this holds both for languages with articles and for languages without articles. In Czech, which belongs to the latter languages, bridging definites receive no marking either, as shown in (34-b-i) (Šimík, 2014, 4).

- (34) a. John bought a book yesterday. # This author was French.
 b. Honza si včera koupil knihu.
 Honza REFL yesterday bought book
 ‘Honza bought a book yesterday.’
 (i) Autorem je Francouz.
 author is French
 ‘The author is French.’
 (ii) #Tím autorem je Francouz.
 this author is French

In Lower Sorbian, *ten* marks bridging definites.

- (35) a. Pěš jěžo z kólasom. Naraz se pšelamjo to wóžidło.
 Pěš drives with bike suddenly REFL breaks the handlebar
 ‘Pěš rides the bike. All of the sudden, the handlebar breaks.’
 b. Kjarčmje su sejželi por buri. ... We tym pšišel ten stary góspodar.
 pub.LOC are sat few peasants in that came the old innkeeper
 ‘A few peasants sat in a pub ... Then the old innkeeper came down.’
 (Šliziński, 1964, 59)

(35-a) is basically identical to (33): every bike has a handlebar. In (35-b), *ten stary góspodar* is definite because it can be inferred from the DP *kjarčmje*: every pub has an innkeeper.

The third type are WEAK DEFINITES (Carlson et al., 2006). Weak definites refer to DPs marked with a definite article that do not seem to refer to a unique referent.

- (36) [Context: Peter sees his father reading the Washington Post, the New York Times, and The Atlantic; he says to his brother:]
 ✓Let’s come back later, he’s reading the newspaper.
 # Let’s come back later, he’s reading this newspaper.

Despite the context with several newspapers, using *the newspaper* is licit. There are various ways to deal with this problem (cf. Schwarz 2009; Aguilar-Guevara & Zwarts 2010), all of which boil down to the idea that uniqueness is satisfied but that the referent is non-trivial. Also weak definites do not allow the use of a demonstrative. Crucially, weak definites in Lower Sorbian are typically marked by *ten*:

- (37) Kito jo šel tek rad tyženjeju jaden ab dwa raza do teje kjarcmje.
 Kito is went also gladly week.DAT one or two times to the pub
 ‘Kito also enjoyed going to the pub once or twice per week.’
 (Elikowska-Winklerowa, 2013, 21)

The town talked about in (37) has three pubs, so the speaker could not have meant a specific pub, which excludes the interpretation of a situational definite. In the discourse preceding (37), no specific pub is mentioned at all, so that (37) is not an instance of an anaphoric definite. Moreover,

the preceding discourse only introduces Kito but not any kind of information about what his favorite pub out of the three pubs could be. This information cannot be not part of the common ground (Kito was neither mentioned earlier nor is he any local celebrity), so that again *teje kjarcmy* is not an anaphoric definite. Instead, (37) means that Kito enjoys going to a certain kind of place, namely to a pub, and not that Kito enjoys going to a specific pub.

The fourth and final type are INHERENT DEFINITES. By this, one refers to all DPs that are definite because of their semantics. Two examples are given in (38).

- (38) a. The sun is shining.
b. The smartest person is John Smith.

(38-a) is an inherent definite because the world we live in is such that there is only one sun. (38-b) is an inherent definite because the meaning of superlatives refers to a unique referent of which the relevant property holds. Crucially, demonstratives are bad with inherent definites, for reasons I return to immediately, both in languages with articles and in languages without articles, like in Czech, where they again receive no marking at all (Šimík, 2014, 2).

- (39) a. #This sun is shining.
b. #This tallest person in town is John Smith.
c. #Na tom měsíci by mohl být primitivní život.
on this moon SBJV could be primitive life
c' Na měsíci by mohl být primitivní život.
on moon SBJV could be primitive life
'The moon could harbor primitive life.'

Inherent definites can also be marked by *ten* in Lower Sorbian:

- (40) a. Te nejzlepší zrnka su pón narosli, gaž ...
the best grains are then grown when
'The best grains grew then, when...' (DoBeS, Sielow, MEW-2012-10-23, MEW-064)
b. Tak jo to słyńcko rědnje swěšiło.
so is the sun beautifully shined
'The sun was shining so beautifully.' (DoBeS, Burg, MEW-2011-12-18, MEW-024)

So far, I have focused on contexts where definite articles are fine but where demonstratives are bad. I showed that *ten* is fine in such contexts, indicating strongly that *ten* is a definite article and not a demonstrative. This conclusion is strengthened by a second set of context where definite article and demonstratives differ. In this second set, demonstratives are fine whereas definite articles must not be used. Also there, *ten* patterns with definite articles.

The first context concerns UNIQUENESS. As shown above in (28), definite DPs are unique. Now crucially, DPs marked with a demonstrative are not unique: replacing *the* by *this* in (28) yields a grammatical sentence.

- (28') [two identical buckets next to each other]
✓Give me this bucket.

Conversely, replacing *the* by *this* is infelicitous in a context where an inherently unique DP is given, as in (40), because *this* implies non-uniqueness. Exactly the same can be observed in Lower Sorbian. Lower Sorbian has developed a new series of demonstratives, formed by prefixing *toś* or *tam* to *ten* (*tošten* is a proximal demonstrative, *tamten* a distal one). Replacing *ten* for *tošten* yields a grammatical sentence in a context such as (28), as shown in (41-a), and also in (41-b), where the question implies more than one referent.

- (41) a. [two identical buckets next to each other]
Daj mi ten zborék. / ✓Daj mi tošten zborék.
give me the bucket give mi this bucket

- b. Kótru žeńsku sy ty wízel? # Tu žeńsku. / ✓Toótu žeńsku.
 which woman are you seen the woman this woman.
 ‘Which woman did you see? This woman.’

If *ten* were a demonstrative, it should not imply uniqueness, similar to a demonstrative. But it does imply uniqueness, indicating forcefully that *ten* is a definite article and not a demonstrative.

The second context are what Lakoff (1974) called AFFECTIVE READINGS. By affective reading, one refers to DPs marked by a demonstrative where the demonstrative expresses a typically negative attitude towards the referent of the DP. Importantly, such DPs need not be previously mentioned in the discourse, for example in (42).

(42) This Trump!

(42) conveys that the speaker expresses his negative attitude toward Trump. Also, (42) can be used to start a conversation about Trump. What is equally important is that the definite article does not have an affective reading. Whatever (43) means, it certainly does not have the affective component one finds in (42).

(43) # The Trump!

Importantly, in Lower Sorbian only *toóten* allows affective readings. *ten* does not have such a reading but is simply as awkward as (43).

(44) [Context: complaining about Donald Trump ...]

- ✓Toósten Trump! / #Ten Trump!
 this Trump the Trump
 ‘This Trump!’

Again, that *ten* does not have an affective reading provides the final piece of evidence that *ten* is a true definite article in Lower Sorbian and not a demonstrative. To conclude, I argued in this section that *ten* is a definite article and not a demonstrative.

5 CONCLUSION

In conclusion, I have shown in this paper that, first, Lower Sorbian – an endangered West Slavic language spoken in the south-east of the German province Brandenburg – allows LBE. Second, I argued that the distribution of the definiteness markers *ten*, *ta*, *to*, *tej*, and *te* and their morphosyntactic properties indicate strongly that these definiteness markers are true definite articles, and not demonstratives. Lower Sorbian therefore possesses two properties, namely LBE and a definite article, which according to many theories of LBE a language cannot have in combination. Consequently, the existence of a language like Lower Sorbian argues against any theory that establishes a connection between the possibility of LBE and the presence or absence of a definite article.

ACKNOWLEDGMENTS

I wish to thank the audience of FASL 28 in Stony Brook for helpful comments (in particular Tanya Bondarenko, Barbara Citko, Daniela Čulinović, Colin Davis, Lanko Marušič, Asya Pereltsvaig, and Jana Willer-Gold). I also thank Radek Šimík for the many discussions about definiteness and related aspects. Fabian Kaulfürst and Marcin Szczepański deserve credits for their help with the DoBeS corpus. Many thanks go to my informant Madlena Norbergowa for her time and patience. I also wish to thank the anonymous reviewers whose comments significantly improved this paper. This research was partly funded by a grant of the Serbski institut in Bautzen and Cottbus.

CONTACT

Andreas Pankau
Andreas.Pankau@fu-berlin.de

REFERENCES

- Aguilar-Guevara, Ana & Joost Zwarts. 2010. Weak definites and reference to kinds. In *Proceedings of SALT 20*, vol. 20, 179–196.
- Blum, Simon. 2013. *Left Branch Extraction und NP-Struktur im Obersorbischen* [Left Branch Extraction and the structure of NP in Upper Sorbian]. BA thesis, University of Potsdam.
- Bošković, Željko. 2005. On the locality of Left Branch Extraction and the structure of NP. *Studia Linguistica* 59(1). 1–45.
- Bošković, Željko. 2009. More on the no-DP analysis of article-less languages. *Studia Linguistica* 63(2). 187–203.
- Bošković, Željko. 2015. On multiple left-branch dislocation: Multiple extraction and/or scattered deletion. In Małgorzata Szajbel-Keck, Roslyn Burns & Darya Kavitskaya (eds.), *Proceedings of Formal Approaches to Slavic Linguistics: The First Berkeley Meeting*, 20–35. Ann Arbor, MI: Michigan Slavic Publications.
- Carlson, Greg, Rachel Sussman, Natalie Klein & Michael Tanenhaus. 2006. Weak definite noun phrases. In Davis Christopher, Amy Rose Deal & Youri Zabbal (eds.), *Proceedings of the 36th Annual Meeting of the North East Linguistic Society* 1, 179–196. Amherst, MA: GLSA Publications.
- Cegłowski, Piotr. 2017. Extraction facts and the internal structure of nominal constructions in Polish–(A report on) an empirical study. *Poznan Studies in Contemporary Linguistics* 53(3). 345–372.
- Corbett, Greville G. 1987. The morphology/syntax interface: evidence from possessive adjectives in Slavonic. *Language* 63(2). 299–345.
- Corver, Norbert. 1990. *The Syntax of Left Branch Extraction*: University of Tilburg dissertation.
- Elikowska-Winklerowa, Marja. 2013. *Stary Żylowař Erwin Mjeršeńc wulicujo* [Old Erwin Mjeršeńc from Żyłow tells stories from his life]. Cottbus: Schule für Niedersorbische Sprache und Kultur.
- Fanselow, Gisbert & Caroline Féry. 2013. A comparative perspective on intervention effects on left branch extractions in Slavic. In Sulym Wolodymyr, Mychajlo Smolij & Chrystyna Djakiw (eds.), *Non Progredi Est Regredi. Festschrift für Alla Paslwaska*, 266–295. Lwiw: Pais.
- Fasske, Helmut. 1995. *Sorbischer Sprachatlas: Syntax. 15* [Linguistic Atlas of Sorbian, vol. 15, Syntax]. Bautzen: Domowina Verlag.
- Hawkins, John A. 1978. *Definiteness and Indefiniteness: A Study in Reference and Grammaticality Prediction*. London: Croom Helm.
- Lakoff, Robin. 1974. Remarks on This and That. In Anthony Bruck, Robert A. Fox & Michael W. La Galy (eds.), *Papers from the 10th regional meeting, Chicago Linguistic Society*, 345–256. Chicago, IL: Chicago Linguistic Society.
- Lyons, Christopher. 1999. *Definiteness*. Cambridge: Cambridge University Press.
- Pereltsvaig, Asya. 2007. The universality of DP: A view from Russian. *Studia Linguistica* 61(1). 59–94.

- Plank, Frans. 1992. Possessives and the distinction between determiners and modifiers (with special reference to German). *Journal of Linguistics* 28(2). 453–468.
- Rappaport, Gilbert C. 2000. Extraction from nominal phrases in Polish and the theory of determiners. *Journal of Slavic linguistics* 159–198.
- Richter, Heinz. 1980. *Die Possessivadjektive im Sorbischen unter Berücksichtigung der benachbarten slawischen Sprachen [The possessive adjectives in Sorbian considering the neighboring Slavic languages]*: University of Leipzig dissertation.
- Ross, John Robert. 1986. *Infinite Syntax*. New Jersey: Ablex publishing corporation Norwood.
- von Schulenburg, Willibald. 1930. *Wendische Volkssagen und Gebräuche aus dem Spreewald [Lower Sorbian legends and customs from the Spree forest]*. Cottbus: Schule für Niedersorbische Sprache und Kultur.
- Schwarz, Florian. 2009. *Two Types Of Definites In Natural Language*: University of Massachusetts, Amherst dissertation.
- Schwela, Gotthold. 1906. *Lehrbuch der Niederwendischen Sprache. Erster Teil: Grammatik [Textbook of the Lower Sorbian language. First part: grammar]*. Cottbus: Author's Edition.
- Schwela, Gotthold. 1911. *Lehrbuch der Niederwendischen Sprache. Zweiter Teil: Übungsbuch [Textbook of the Lower Sorbian language. Second part: workbook]*. Cottbus: Author's Edition.
- Śliziński, Jerzy. 1964. *Sorbische Volkserzählungen [Sorbian folktales]*. Berlin: Akademie Verlag.
- Stojković, Jelena. 2019. The fleeting DP in Bulgarian and Macedonian. The view from Left Branch Extraction. Talk given at the Kolloquium Slawistische Linguistik (HU Berlin), November 18th 2019.
- Talić, Aida. 2017. Adverbial Left-Branch Extraction and the Structure of AP in Slavic. In *Proceedings of Formal Approaches to Slavic Linguistics. The NYU meeting 2015*, 399–418. Ann Arbor, MI: Michigan Slavic Publications.
- Šimík, Radek. 2014. Definiteness and articleless languages. Course notes, Letní škola lingvistiky, August 17-22 2014, Dačice.