



**HAL**  
open science

# Syntactically conditioned code-switching? The syntax of numerals in Beni-Snous Berber

Lameen Souag, Fatma Kherbache

► **To cite this version:**

Lameen Souag, Fatma Kherbache. Syntactically conditioned code-switching? The syntax of numerals in Beni-Snous Berber: The syntax of numerals in Beni-Snous Berber. *International Journal of Bilingualism*, 2016, *Construction and Building Materials*, 20 (2), pp.97 - 115. 10.1177/1367006914536002 . halshs-01376208

**HAL Id: halshs-01376208**

**<https://shs.hal.science/halshs-01376208>**

Submitted on 4 Oct 2016

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution - NonCommercial - NoDerivatives 4.0 International License

# Syntactically conditioned code-switching? The syntax of numerals in Beni-Snous Berber

**Lameen Souag**

LACITO, CNRS, France

**Fatma Kherbache**

University of Abou Bakr Belkaid, Tlemcen, Algeria

[Postprint version of an article published in: *International Journal of Bilingualism* 2016, Vol. 20(2) 97–115. DOI: 10.1177/1367006914536002. For correct pagination, please refer to the printed version; the page number can, however, be approximated well by adding 96 to the page numbers herein.]

## Abstract

*Aims and Objectives/Purpose/Research Questions:* Grammatical rules in one language that induce the speaker to switch to another language (Matras' "bilingual suppletion") are reported for two languages: Beni Snous Berber (Destaing 1907), Jerusalem Domari (Matras 2012). Few details are available, yet the two cases show greater similarities than expected if any grammatical rule could specify switching. This paper seeks to describe the phenomenon more precisely and to provide a principled explanation for the similarities.

*Design/Methodology/Approach:* Destaing claims that numerals greater than ten obligatorily select for Arabic nouns, but provides few examples and no frequency indications. The description was made more precise by examining all numeral+noun phrases in published texts, supplemented by elicitation. Contact effects were distinguished from retentions through comparison with nearby Berber and Arabic varieties.

*Data and Analysis:* Destaing's (1907) text corpus yielded 213 numeral+noun phrases; another 169 were elicited in 2013 from seven surviving speakers. In both datasets, counterexamples to Destaing's original claim were found, but switching was restricted to numerals greater than ten, and Fisher's exact test confirmed this factor's significance.

*Findings/Conclusions:* Beni-Snous Berber – like Domari – shows a statistically significant tendency to use Arabic nouns with numerals for which Arabic and Berber selectional requirements conflict. Modern speakers additionally show optional syntactic calquing in such cases, accompanied by fewer switches. These facts are predicted by the hypothesis that "bilingual suppletion" is induced by words shared across the two languages with different selectional requirements.

*Originality:* "Bilingual suppletion" remains little-researched, and its relative frequency in corpus data has not previously been examined. No restrictive explanation for it has previously been proposed.

*Significance/Implications:* The results indicate that "rules" inducing language switching (always optional) can emerge naturally from conflicts between grammars within bilingual populations, rather than needing to be directly encoded in the grammar.

## Keywords

borrowing, code-switching, bilingual suppletion, numerals, Berber, Arabic

## 1. Introduction

In his recent grammar of Domari, Matras (2012, p. 34) draws attention to an unusual phenomenon he labels “bilingual suppletion”: certain constructions within otherwise monolingual Domari discourse – specifically, comparatives and numeral-noun phrases for numerals 4 through 10 – appear to, respectively, require (Matras, 2012, pp. 206–207) or strongly encourage (Matras, 2012, pp. 191–201) switching into Arabic. Thus, not only in texts but even in elicitation via Hebrew, speakers systematically produce:

- |   |  |
|---|--|
| (1) <i>taran zara,</i><br>three boy,<br>“three boys”, | ’ <i>ARBAʕ</i> <i>WLĀD</i><br>four            boys <sub>Ar</sub><br>“four boys” (Matras, 2012, p. 197) |
|---|--|

The comparatives can easily be analysed as large-scale suppletive borrowing, as Matras notes, but such an analysis is not feasible for the numeral-noun phrases, in that it would require one to postulate that every single noun has a borrowed counterpart used only in a very restricted syntactic environment. Matras (2012, p. 381) takes this as evidence that “the grammatical rules of Domari permit and at times instruct speakers to avail themselves of those parts of the repertoire that are used in interaction in Arabic even in Domari conversation”. This implies that, if all speakers of a language are bilingual, its grammatical rules can make reference to material in another language, and that our understanding of what counts as a possible grammatical rule must be broad enough to encompass such cases.

The phenomenon Matras points out is surprising enough to deserve further cross-linguistic investigation. His interpretation of it, however, is more problematic, in that it suggests a wider range of possibilities than seem to be observed. If the grammatical rules of a language can make reference to material in another language, then nothing rules out the possibility of obligatory multi-word switches unrelated to lexical borrowing. Yet, although Domari has retained numerals up to three and sometimes higher from Indo-Aryan, the tendency (not obligation) to switch all following nouns applies only with numerals borrowed from Arabic, and among them only with those which select for a different number in Arabic (singular) than in Domari (plural). As will be seen below, the same generalisations hold for Beni-Snous Berber. This suggests that the phenomenon is somehow connected to lexical borrowing and selectional requirements.

One way to motivate such a connection is to consider how the bilingual speaker partitions his or her linguistic repertoire. Suppose that minority language A is under the influence of dominant language B. All speakers of A are likely to use elements originating in B, formally indistinguishable (where the two languages’ morphology and phonology permit) from their counterparts in B. Monolingual speakers of A have little choice but to treat such elements as part of A, and bilingual speakers dealing with them are likely to notice that such elements are used and understood by monolinguals whereas other elements of B are not, encouraging them to treat them as part of both A and B (cf. Matras, 2010, p. 67). However, if there are few or no monolingual speakers of A, then in principle bilingual speakers have another option: to treat such elements as short code-switches into B.

The difference between borrowings and single-word code-switches – already much debated (Myers-Scotton, 1992; Poplack, Sankoff, & Miller, 1988; Poplack, 2012) – at first sight appears purely academic for cases where phonology and morphology provide no basis for the distinction. In fact, however, Myers-Scotton’s Embedded Language Island Hypothesis suggests that it should make a difference under certain circumstances.

The revised version of this hypothesis (Myers-Scotton, 1997, p. 250) states that “When there is insufficient congruence between the lemma underlying an E[mbedded] L[anguage] content morpheme and its M[atrix] L[anguage] counterpart at one or more of the three levels of lexical structure, the only way to access the EL element is in an EL island.” If a word identical in A and B has irreconcilable selectional requirements in the two languages, and is being treated by the bilingual speaker as part of B, then by this hypothesis its use within A should be expected to trigger an EL Island. This predicts not only that apparent cases of multi-word “bilingual suppletion” should involve words shared with the dominant language having different selectional requirements, but also that if, through calquing caused by increased influence from the dominant language, the selectional requirements become identical, the usage of “bilingual suppletion” should – somewhat counter-intuitively – decrease, rather than increase. As will be seen in Sections 3 and 4, both predictions are supported by diachronic data for the only other case of “bilingual suppletion” known to the authors, first described more than a century earlier – Beni-Snous Berber.

Describing the Berber language of the Beni-Snous in western Algeria, Destaing (1907, pp. 212–213) reports that “From “ten” to “nineteen” (inclusive), the number is followed by the Arabic singular noun... After the numerals “twenty, thirty, forty (etc.)”, one uses the singular Arabic noun.” He supports this claim with several examples, including one apparent minimal pair:

(2)	<i>STTA</i>	<i>n</i>	<i>tsennan</i> ,	<i>EḤDAḤC-ER</i>	<i>MRA</i>
	six	GEN	women .CST,	eleven-LK	woman <sub>Ar</sub>
	“six women”,			“eleven women”	

Unfortunately for linguistics, the Beni-Snous have since collectively shifted to Arabic. However, in terms of published materials, their language remains, in some respects, better documented than Domari. Destaing (1907) provides a quite detailed grammar, along with a collection of texts, to which he later (1914) added an extensive dictionary. Moreover, we need not rely exclusively on Destaing’s grammar and texts: a few elderly speakers of the language remain. Together, these sources enable us to check Destaing’s description. The data thus obtained make it clear that, as in Domari, the preference for switching to Arabic phrases with certain numerals holds only as a statistical tendency, and allows us to see how it has changed with the selectional requirements of borrowed numerals.

In this paper, the orthography used for Berber and Arabic reflects the Kabyle Latin standard: *e* = /ə/, *y* = /j/, *c* = /ʃ/, *j* = /ʒ/, *ǧ* = /dʒ/, *č* = /tʃ/, *ε* = /ʀ/, *h* = /ħ/, underdot = pharyngealisation, underline = spirantisation. Forms indistinguishable from Arabic are capitalised. See Appendix for abbreviations used in glosses.

## 2. Background

### 2.1 The language and its speakers

Beni-Snous Berber, spoken in the mountains west of Tlemcen near the Algerian-Moroccan border, belongs to the Zenati subgroup of Northern Berber. While it shows slightly higher levels of Arabic influence than most Northern Berber varieties (Kossmann, 2013, p. 110), it neither presents difficulties for genetic classification nor shows signs of massive structural imposition; in particular, its verbal inflection and core case-marking adpositions are entirely Berber and its syntax is quite normal for Berber.

There is no evidence to suggest massive code-switching in Beni-Snous Berber either. Excluding the noun+numeral phrases under discussion, multi-word insertional code-switches are strikingly rare in Destaing’s

texts – his entire corpus yields no more than a few religious formulae (some arguably alternational) such as *Allah yexlef* “may God compensate you” (p. 360), which would certainly have formed part of the competence of monolingual speakers, if any existed; two or three vocatives such as *ya Selṭan* “O Sultan” (p. 360); and a few compounds that could equally be analysed as loanwords, eg *la-beṣṣeḥ* “certainly” (p. 348). Alternational code-switches, absent even from stories whose Tarifiyt counterparts require Arabic formulas at certain points (Maarten Kossmann, pc) are even rarer – no more than ten are found, all in quoted speech and all more or less formulaic (eg *la ilaha illa Allah* “There is no god but God”, p. 264). Informal conversation might have featured more switching, but the speakers’ competence evidently included a basically monolingual mode.

The small villages that make up the Beni Snous community are spread across the mountains around the Tifousser (Khemis) and Tafna (Oued Lekbir) valleys. The Beni Snous (Berber *At Snus*) have been present around Tlemcen for many centuries; they were already mentioned by the 14<sup>th</sup>-century historian Ibn Khaldūn, who classed them as a Zenati Berber tribe (Ibn Khaldūn, 2012, p. 2016). In his time, they would have been part of a Zenati dialect continuum stretching from the Rif to central Algeria, whose remnants still show substantial mutual comprehensibility. According to Destaing (1907, p. XXVIII), while the Beni-Snous found it difficult to understand Kabyle or Tashelhiyt, they easily understood both eastern Rif Berber, spoken just to its west in Morocco (Kossmann, 2000), and Central Maghrebi varieties spoken further east (Basset, 1895).

Unlike their neighbours, however, the Beni-Snous continued to speak Berber long after most of western Algeria had shifted to Arabic, and well into the twentieth century; in 1903-1905, Destaing was still able to gather extensive Berber data from Beni-Snous villages in the Kef and Oued Khemis regions. By that time, their dialect already contained numerous Arabic loans, like all Northern Berber varieties, and “all the tribe’s inhabitants could speak Arabic” (Destaing 1907, *ibid*). However, language shift proceeded rapidly over the early 20<sup>th</sup> century; the youngest living speakers were born in the 1940s. At present, their daily language is Arabic, and of the Beni Snous daira’s 22,000 inhabitants, hardly a dozen still speak Berber. Only a few Berber terms have been retained in their Arabic – in particular toponyms, traditional or wild foods such as *sesnu* ‘arbutus’, *tabya* ‘blackberries’, and traditional technology such as *tila* ‘couscous sieve’, *azduz* ‘pestle’.

## 2.2 Borrowing and the numeral system

Most Berber varieties from Morocco to Egypt, and all but the southernmost Northern Berber ones, have borrowed all numerals above ‘two’ (or occasionally ‘three’) from Arabic, the language of inter-regional trade, major urban centres, and education (Kossmann, 2013, pp. 306–311; Souag, 2007). Beni-Snous Berber is no exception to this pattern, but it more specifically aligns with neighbouring Rif Berber in having borrowed all numerals except ‘one’ (M *iğ(en)*, F *tict*) from Arabic (Lafkioui, 2007, p. 266); this contrasts with Central Maghrebi varieties, which also retain Berber *sen* (F *sent*) ‘two’ (Basset, 1895, sec. 81). Also as in Rif Berber (Kossmann, 2013), ‘two’, *tnayen*, differs from modern Algerian Arabic *ZUJ*, reflecting borrowing at an earlier stage (cp. Classical Arabic *ITNAYN*). It can thus be regarded as specific to regional Berber, and is not used in Beni-Snous Arabic. According to W. Marçais (1902, p. 156), some rural Arabic dialects near Tlemcen used *TNIN* rather than *ZUJ*, but the Beni-Snous are not familiar with any such dialects.

Above ‘two’, all Beni-Snous Berber numerals are indistinguishable from their Western Algerian Arabic counterparts. Most speakers are unaware that Berber has ever had non-Arabic numerals other than ‘one’<sup>1</sup>, and all living speakers are bilingual in Arabic. The relevant forms, following Destaing (1907:212) with transcription modified to fit modern data [and irreconcilable modern forms between square brackets], are given in Table 1.

**Table 1.** Beni-Snous Berber numerals

1	<i>iġ(en)</i> , f. <i>tict</i> / <i>WAḤED</i> , f. <i>WAḤDA</i>	11	<i>EḤDAεC</i> [ <i>ḤDAεC/ḤDAC</i> ]	10	<i>εECRA</i>	100	<i>MYA</i>
2	<i>tnayen</i> / <i>ZUJ</i>	12	<i>TNAεC</i> [ <i>TNAεC</i> ]	20	<i>εECRIN</i>	200	<i>MEYTIN</i> [ <i>MITIN/MITAYEN</i> ]
3	<i>TLAṬA</i>	13	<i>TELTAεC</i>	30	<i>TLAṬIN</i>	300	<i>TEL-TMYA</i>
4	<i>ARBεA</i>	14	<i>ARBATAεC</i> [ <i>RBA(ε)TA(ε)C</i> ]	40	<i>ERBEIN</i>	400	<i>ARBεA-MYA</i>
5	<i>XEMSA</i>	15	<i>XEMSTAεC</i> [ <i>XEMSTA(ε)C</i> ]	50	<i>XEMSIN</i>	1,000	<i>ALEF</i>
6	<i>STTA</i>	16	<i>STTAεC</i> [ <i>STTA(ε)C</i> ]	60	<i>STTIN</i>	2,000	<i>ALFEYN</i> [ <i>ALFIN/ALFAYEN</i> ]
7	<i>SEBεA</i>	17	<i>SBATAεC</i>	70	<i>SEBEIN</i>	3,000	<i>TEL-ALAF</i>
8	<i>TMENYA</i>	18	<i>TMENTAεC</i>	80	<i>TMANYIN</i>	10,000	<i>εECR-ALAF</i>
9	<i>TESEA</i>	19	<i>TSATAεC</i>	90	<i>TESEIN</i>	15,000	<i>XEMSTAεAC ALF</i>
10	<i>εECRA</i>	20	<i>εECRIN</i>	21	<i>WAḤD U</i> <i>εACRIN</i>	1,000,000	<i>BELYUN</i> [ <i>MELYUN</i> ]

In addition to the forms above, 2-10 have shortened forms which can only be used when the numeral is directly adjacent to a noun, without an intervening genitive: *-AYEN* / *-IN*, *TELT* / *TEL*, *REBE*, *XEMS*, *SETT*, *SEBE*, *TMEN*, *TESE*, *εECR*. This pattern too is widespread – possibly ubiquitous – in Northern Berber (Kossmann, 2013, p. 311), cf. 5.1. In local Arabic these forms can be combined with any nouns, while throughout Berber they are restricted to measure nouns borrowed from Arabic (including the higher numeral units, starting from ‘hundred’); while the distributions are different, these forms too are shared across the two languages.

Distinguishing Berber nouns from Arabic nouns is easier than distinguishing Berber numerals from Arabic ones. The most frequently used nouns are generally retained from Berber, showing no similarity to Arabic. Even Arabic borrowings are often distinguishable from their Arabic source by virtue of having adopted distinctive Berber morphology. In Berber, masculine singular nouns are generally marked with *a-*, and feminine singular with *ta-...-t*; Arabic has no masculine singular affix and marks the feminine singular with *-A*. Arabic loans into Berber are frequently Berberised by imposing the corresponding Berber affixes. Even when an Arabic loan retains its original morphology, it can usually be distinguished from Arabic in some syntactic environments – including numeral-noun combinations – by looking at the article: most loans into Berber enter with an invariant copy of the Arabic definite article *L-*, used even in contexts where Arabic would not permit an article.

## 3. Numeral syntax in Destaing's data

### 3.1 Destaing's analysis

Destaing (1907, pp. 211-213) devotes two pages to the form and syntax of cardinal numerals. As seen in 2.2, only 'one' and 'two' differ from Algerian Arabic, although their Algerian Arabic equivalents, 'one' م *WAḤED* ف *WAḤDA* and 'two' *ZUJ* / *JUJ*, are also used. The remaining numerals are precisely as in Algerian Arabic. 'One' is followed by a singular Berber noun in the construct state.<sup>2</sup> 'Two' through 'nine' (and presumably 'ten') are followed by the Berber noun in the plural, in the construct state (sometimes with the genitive, as illustrated by his examples). All numerals above 'ten' are followed by an Arabic noun in the singular, with an intervening linker *-ER-* for 'ten' (*sic*, presumably 'eleven') through 'nineteen' and *-T-* for 'hundred', and no intervening linker for other numerals. A number of examples are given to illustrate these claims, eg 2, or:

- (3) *XEMSTAEC-ER*      *BEGRA*  
 fifteen-LK              cow<sub>Ar</sub>  
 "fifteen cows"
- (4) *XEMSIN*      *BEḶLA*  
 fifty              mule<sub>Ar</sub>  
 "fifty mules"
- (5) *TEL T ALAF*      *EJDI*  
 three thousands      kid<sub>Ar</sub>  
 "three thousand kids"

Contrast normal Beni-Snous Berber *tafunast* PL *tifunasin* 'cow', *aserdun* PL *iserdan* 'mule', *iyid* PL *iyayden* 'kid (goat)' (Destaing 1914). The only other nouns for which such examples are given are 'woman' (Berber *tamṭtuṭ* PL *tisennan* vs. Arabic *MRA*), 'sheep' (Berber *icerri* PL *icraren* vs. Arabic *KEBC*), and 'mare' (Berber *taymarṭ* pl. *tiymarin* vs. Arabic *EAWDA*). All happen to be generic animates.

Destaing does not state how he obtained the data on which he bases this section of the grammar. However, all of the examples in this section for numerals  $\geq 11$  are absent from, and unparalleled in, his published texts, and the conclusions he reaches are quite different from those to which his texts would have led him (cf. 3.2). His examples in this section therefore most likely represent the results of elicitation, although some might have been jotted down from overheard discourse.

### 3.2 Reanalysing Destaing's data

The texts Destaing transcribed paint a picture rather different to his grammar. In those texts, there occur 213 numeral+noun phrases as outlined in Table 2, more than 70% of them using the numeral 'one'.

**Table 2.** Frequency by numeral and language of numeral+noun phrases in Destaing's texts

	Berber / Berberised Arabic	Arabic or identical to Arabic	Total
one	152	0	152
two – ten	38	11	49
≥ eleven	2	10	12

This already reveals two counterexamples to the claim that numerals above ten always take suppletive Arabic forms: *RBAËIN wussan* (forty days) “forty days” (p. 359), with a Berber noun, and *SEBËA U SEBËIN lqamaṭ* (seven and seventy fathoms) “seventy-seven fathoms” (p. 294), with a noun of Arabic origin but using the article *l-* and the plural *-aṭ* in a position where Arabic would require an article-free singular, forming an unambiguously non-Arabic phrase, and suggesting that *lqama* is best interpreted as a loanword into Berber. These counterexamples prove that Destaing's claim is not valid as an absolute rule. At the same time, forms indistinguishable from Arabic are five times commoner than Berber ones in this range, supporting Destaing's conclusion as a statistical generalisation.

In fact, however, every instance of an Arabic numeral+noun phrase in this data – below or above ten – belongs to a single coherent semantic class: measure words, for which the use of borrowed syntax and loanwords is widespread across Berber, and indeed elsewhere (as in Japanese, where inherited *toshi* ‘year’ contrasts with Chinese loans *ICHI-NEN*, *NI-NEN*, *SAN-NEN*... ‘one year, two years, three years...’). Of the 21 Arabic numeral+noun phrases, 10 use ‘days’ (Berber *ass* PL *ussan*, Arabic *YUM* PL *IYYAM*), 8 use ‘years’ (Berber *asegg<sup>was</sup>* PL *isegg<sup>wasen</sup>*, Arabic *SNA* PL *SNIN*), and the remainder feature the unit of weight ‘quintals’ (Berber *akunṭar*, Arabic *QUNṬAR*) and the units of currency ‘doro’ (Berber=Arabic *DURU*) and ‘soldi’ (Berber=Arabic *ṢULDI*). Unlike in Figuig Berber (see 5.1), the use of Arabic loans with Arabic syntax for such measures is not obligatory, as already illustrated by the two counterexamples above and by other cases such as *SEBËA wussan* (seven days) “seven days” (p. 146), *XEMSA n isegg<sup>wasen</sup>* (five GEN years) “five years” (p. 282). However, Arabic is strongly preferred for such forms: looking at this data set for numerals ≥2, we find that Arabic nouns are used for ‘day’ in 8 out of the 13 instances of ‘*n* days’, and for ‘year’ in 8 out of the 9 instances of ‘*n* years’. In contrast, when not preceded by numerals ‘day(s)’ and ‘year(s)’ are rather consistently expressed by Berber nouns, never Arabic ones (except in the lexicalised compound *kull-a-yum* ‘every day’, p. 251; see Destaing (1914, s.v. ‘jour’)). Measures of time also stand out for allowing the Arabic dual ending *-AYEN* (attested here in *YUM-AYEN* (day-DUAL<sub>Ar</sub>) ‘two days’, p. 302), fallen into disuse even in regional Arabic for most other words. Breaking the data down into measure words (units of time, length, weight, and currency) vs. non-measure words, we find:

**Table 3.** Frequency by measure status, numeral, and language of numeral+noun phrases in Destaing's texts

	Non-measure			Measure		
	Berber	Arabic	Total	Berber	Arabic	Total
one	141	0	141	11	0	11
two – ten	32	0	32	6	11	17
≥ eleven	0	0	0	2	10	12



This observation in turn allows us to explain the much higher frequency of Arabic numeral+noun phrases above ten; in fact, all items counted in the text with numerals greater than ten are measure words. 7 examples use ‘days’, 4 use ‘years’, and the remainder ‘quintals’ and ‘fathoms’. Judging by the texts, the generic animates illustrated by Destaing’s examples are in practice simply not commonly counted with large numerals.

In addition to bearing on the usage of Arabic phrases, the texts also require a slight refinement to Destaing’s claims on number selection. Judging by the two counterexamples above, numerals above ten take the plural when an Arabic noun is not substituted, contrary to Destaing. In general, as in Destaing, numerals up to ten take the plural, but there are two exceptions: *SETTA ŠULDI* (six soldi) “six soldi” (p. 321), *ARBEA DURU* (four doros) “four doros” (p. 292). Algerian Arabic requires the singular after numerals for units of currency (see 5.2), as does Figuig Berber (see 5.1); the same thus appears to be true of Beni-Snous.

Destaing does not comment on the usage of the genitive marker *n*, but his examples indicate that it could be present or absent before Berber nouns, although it never intervenes in Arabic phrases. Textual data bears this out. It is not always possible to determine whether genitive *n* is used, due to conditioned assimilation of *n*, but in *TEMNYA tiwqay* (eight pebbles) “eight pebbles” (p. 340) or *SETTA inejjiwen* (six guests) “six guests” (p. 307) it is unambiguously absent, whereas in far more cases it is unambiguously present, eg *TLATA n iyunam* (three GEN reeds) “three reeds” (p. 263). In no case does *n* unambiguously occur with an Arabic noun.

The corpus also casts unexpected light on the syntactic structure of numeral+noun phrases in Beni-Snous Berber. Conjunct constructions offering a choice of possible numerals for the same noun, using the Berber conjunction *nay* ‘or’, are relatively frequent (17 examples.) Among these, we find examples of phrasal conjunction with noun ellipsis – [[Num (*n*) N] or Num], eg *XAMSA wussan nay SETTA* (five days or six) “five days or six” (p. 306) – and examples with numeral conjunction alone – [[Num or Num] (*n*) N]], eg *TLATA nay SETTA inejjiwen* (three or six guests) “three or six guests” (p. 307). Phrasal conjunction without ellipsis – [[Num N] or [Num N]] – is attested in only four instances, all Arabic, eg *XAMS-SNIN nay SETT-SNIN* (five-years or six-years) “five years or six years” (p. 320). Conversely, for Arabic numeral phrases only phrasal conjunction with or without ellipsis is attested (7 instances), and never numeral conjunction. This suggests that Arabic numeral phrases have a different syntactic status than Berber ones: Arabic numeral phrases are processed by Berber syntax as indivisible units, whereas Berber ones have accessible internal structure into which conjunctions can be inserted. Unfortunately, available examples are limited to measure nouns.

#### 4. Contemporary numeral data

As seen, Destaing’s presumably elicited examples for numeral+noun constructions  $\geq 11$  show no overlap with his textual examples, and both sets combine such numerals only with measures or generic animates. This made further data desirable. In order to provide a wider perspective, one author interviewed a woman speaker from Kef, aged 87 (I), and the last six speakers from the Beni-Ziddaz (formerly Adziddaz): three men, aged 80, 72, and 70 (E, H, A), and three women, aged 85, 76, and 71 (C, B, D), who had all acquired the language one to two generations later than the period of Destaing’s work. These speakers have few occasions to practice the language in daily life, except occasionally when visiting similarly elderly relatives.

The interviews produced a corpus containing 171 numeral+noun phrases. Since such phrases occur so rarely in natural speech, they had to be actively elicited, through translation (eg “How do you say “I have *n* sheep?” in Berber?”) or more indirectly (“How many sheep did you own in the old days?”) Occasionally, the speaker spontaneously provided further non-elicited examples. Measure words were avoided in elicitation (see

3.2, 5.1). This corpus lets us determine the grammar of numeral+noun phrases, in a context strongly favouring monolingual discourse, for the last speakers of Beni-Snous Berber.

In important respects, these results are not directly comparable with Destaing's. Destaing describes the grammar of a community where Berber was still regularly spoken in daily life, based on both elicitation and texts (although, on this particular point, his presumably elicited results in the grammar disagree with his corpus data.) In the course of a century in which near-complete language shift has occurred, grammatical change may be expected, and indeed the results often suggest attrition. Nevertheless, we find an encouraging degree of continuity with inherited Berber patterns. Where the elicited data agrees both with Destaing's corpus data and with wider Zenati Berber norms (see 5.1) but disagrees with his grammar, we may assume that the elicited data reflects continuity rather than change – it is unlikely that Beni-Snous Berber would coincidentally regain an option that reconstruction indicates was present in it prior to Arabic influence, and that Destaing's texts indicate was never actually lost. Conversely, where speakers repeatedly produce Arabic numeral-noun phrases in agreement with Destaing's grammar despite direct prompting to produce Berber and not Arabic, we may assume that this too reflects the continuity of a norm favoring code-switching even in otherwise monolingual discourse in particular syntactic contexts, rather than attrition.

Non-numeral-related code-switching is marginal in this modern elicited data, just as it is in Destaing's corpus. However, as expected given the dominance of Arabic, it occurs here more frequently and in a wider range of contexts than in Destaing's corpus. Whereas code-switches in Destaing's corpus are rare and consistently formulaic (recall 5.1), the much smaller modern data set includes multiple instances of non-formulaic code-switching, as in example 14 below, and the following (in response to the prompt “two boys said they would come”):

- (6) *tnayn irban, RA-HŪM I-GUL-U I-ŽI-W.*  
 two boys, lo-3pl 3IPFV-say-3PL 3IPFV-come-3PL.  
 Two boys, they say they're coming.

#### 4.1 Lower numerals

The usual situation for 1-10 is as described by Destaing, except for the loss of the construct state. For “one”, Num N.SG is dominant; for 2, Num N.PL is almost always used, except for one instance of Num *n* N.PL; for 3-9, speakers A, C, D, E, and H all use Num *n* N.PL, eg:

- (7) *XEMSA n tazarin*  
 five GEN figs  
 “five figs”

Only occasional divergences are found. For “one”, speaker D once used the Arabic numeral *WAḤDA*, and speakers C and E also used the construction N.SG Num, possibly a calque on Arabic but also sporadically attested in nearby Beni Iznacen Berber (Maarten Kossmann pc). Speaker B, whose proficiency is particularly poor (she omits subject agreement on verbs), also used it with “five”.

In two instances (Speakers A, E) an Arabic genitive linker is substituted for the normal Berber one; such a usage is well-attested in local Arabic (cf. 5.2), but contrasts both with Destaing's data and with these speakers' usual practice, and hence may well be a production error:

- (8) *TESĒAD* *iserdan*, *ĒECRA* *NTAĒ izemrawen*  
 nine GEN<sub>Ar</sub> mules, ten GEN<sub>Ar</sub> lambs  
 “nine mules”, “ten lambs”

In all cases, as described by Destaing, the noun remains in Berber.

## 4.2 Higher numerals

4.2.1 *With Berber nouns.* The construction Num (*n*) N.PL using a Berber noun – dominant in mutually intelligible Berber varieties on both sides of Beni-Snous (see 5.1) and attested twice in Destaing’s Beni-Snous texts for numerals  $\geq 11$  (see Table 2) but described by Destaing as restricted to numerals up to ten – is frequently used for all numerals by speakers A, C, D, E, and H:

- (9) *HDAC n irḥan*, *HDAC n iẓerwan*, *XEMSTAC n tazarin*  
 eleven GEN boys, eleven GEN crags, fifteen GEN figs  
 “eleven boys”, “eleven crags”, “fifteen figs”
- (10) *STTA U SETTIN n ixamen*, *ĒECRIN iẓerwan*  
 six and sixty GEN houses, twenty crags  
 “sixty-six houses”, “twenty crags”

For 11-19 this construction is overwhelmingly dominant, while for  $\geq 20$  it is frequent. However, another constructions using Berber nouns but entirely absent from Destaing is also well-attested for A, C, E, and H, namely Num (*n*) N.SG, in which the noun remains singular as in Arabic:

- (11) *HDAC e tameṭṭut*, *ĒECRIN n azru*  
 eleven uh woman, twenty GEN crag  
 “eleven women”, “twenty crags”

This construction is occasionally attested for 11-19, and is about as frequent as the first one for numerals  $\geq 20$ .

A single instance of N.SG Num (by speaker C) is also found, diverging from normal Berber and Arabic practice alike but corresponding to the minor construction already noted for the lower numerals: *azru ĒECRIN* (crag twenty) “twenty mountains”.

The high productivity of constructions using Berber nouns in all cases was confirmed by re-eliciting some of Destaing’s own examples with multiple speakers; in almost all cases Berber equivalents were given, eg:

- (12) *HDAĒC en tmeṭṭutin*, *STTAĒC en tiymarin*  
 eleven GEN women.CST, sixteen GEN mares  
 “eleven women”, “sixteen mares”

On the basis of this data, we can confidently state that modern speakers’ grammar does not require Arabic nouns to be substituted after large numerals. The large majority of elicited examples diverge strikingly from Destaing’s description: whereas Destaing claims that the noun following a number greater than ten is always in

Arabic, these use Berber nouns rather than Arabic equivalents, including for animate generics. However, this is not the whole story.

4.2.2 *With Arabic nouns*. In many cases, across four different speakers – usually for numerals  $\geq 20$ , once with 11 – an Arabic noun, always in the singular, is substituted where a Berber one might have been expected, resulting in a construction Num N<sub>Ar.SG</sub> (with an intervening linker *EL* in the case of 11, as in local Arabic: *HDAC-EL MRA* [eleven-LK woman] “eleven women”). In this elicitation-based context, the speaker’s goal is normally to conform to monolingual norms and to avoid phrase-internal code-switching; closer examination of the context of such usages suggests that code-switches of this specific kind form part of the speakers’ target norm. In the following example, for instance, Speaker E initially prefers to keep the noun in Arabic, using a Berber form only after extensive prompting:

- (13) *zwa-γ*      *ƎECRIN*      *ŽBEL*      *BAC* *ebbɔ-ey*      *dadi*  
 go-1SG      [twenty      mountain]<sub>Ar</sub>      so.that arrive-1SG      here  
 I crossed twenty mountains to get here.

(interviewer repeats prompt)

*ruh-ey* *ƎECRIN*      *EŽBEL*      *BAC* *ebbɔ-ey* –  
 go-1SG [twenty      mountain]<sub>Ar</sub>      so.that arrive-1SG  
 I crossed twenty mountains to get–

*a,*      *ƎECRIN*      *adrar*      *BAC* *ebbɔ-ey*      *dadi*  
 ah,      twenty      mountain      so.that arrive-1SG      here  
 ah, twenty mountains to get here.

For the very familiar concept “cows”, Speaker A explicitly pushes back against the interviewer’s attempts to elicit a Berber form, even though Speakers E and H both used only the Berber noun in similar contexts:

- (14) *yer-i WAHED U TLATINBEQRA, MA-KAN-C*      *ELLA EL-HEDRA HADIK.*  
 at-1SG [one and thirty      cow]<sub>Ar</sub>, [not-exist-NEG2 except the-speech      that.F.SG]<sub>Ar</sub>.  
 I have thirty-one cows, that speech is all there is [=that’s the only way to say it].

No speaker used only or even predominantly Arabic nouns in this environment. However, within this data set, the substitution of Arabic nouns appears to be restricted to this environment: no comparable hesitations or substitutions were observed for numerals below 11. A weaker version of Destaing’s claim thus seems to apply: for numerals  $\geq 11$ , even modern speakers show a tendency to substitute Arabic forms for some nouns. Quantitative analysis gives a clearer picture.

### 4.3 Statistical analysis

The elicited data yields a total of 170 numeral+noun phrases, excluding repetitions of the same phrase within a single recording. Table 4 concisely sums up the range of constructions found. In one instance (*tmanyā u tesein n sekkur* [eight and ninety partridge] “ninety-eight partridges”), the noun is of Berber origin

(*ta-sekkur-t* ‘partridge’) but fails to show Berber morphology, suggesting that it might be a loan into the speaker’s Arabic dialect; it has provisionally been treated as Berber.

**Table 4.** Frequency by numeral, language, and construction properties of numeral+noun phrases in elicited data

Order:	N-Num	Num-(Linker)-N											Total
N lang.:	Berber						Arabic						
N num.:	SG	PL				SG							
Linker:	∅	∅	<i>n</i>	<i>D</i>	<i>NTAÉ</i>	<i>n</i>	<i>-ER</i>	∅	∅	<i>-EL</i>	<i>D</i>	<i>n</i>	
one	4	0	0	0	0	1	0	7	0	0	0	0	12
two – ten	1	6	45	1	1	0	0	0	0	0	0	0	54
eleven – nineteen	0	0	14	0	0	3	1	1	0	1	0	0	20
≥ twenty	0	3	27	0	0	13	0	23	15	0	1	2	84
Total	5	9	86	1	1	17	1	31	15	1	1	2	170
Status:	ML (inc. all ‘one’)			CS		CML (excl. ‘one’)			EL		CEL		
	108			2		41			17		2		

In terms of Myers-Scotton’s (2002) typology of codeswitching, the unmixed Matrix Language construction here is Num (*n*) N.PL, (for ‘one’, Num (*n*) N.SG) attested for a range including all numerals except ‘one’ in Destaing and in Beni-Snous’ closest relatives (see 3.2, 5.1); in this data set, it is the commonest overall and for every applicable range of numerals. Phrase-internal classic codeswitching seems to be exemplified by the rare construction Num *D* / *NTAÉ* N.PL, which uses an Arabic genitive particle, but is only attested for ‘two’–‘ten’. In Arabic, the numerals ‘eleven’–‘nineteen’ take a linker *-EN/-EL/-ER* (see 5.2); *-EN* happens to be indistinguishable from the Berber genitive particle *n*, but *-EL* and *-ER* are attested once each, followed by a singular noun respectively in Arabic and in Berber, and may be grouped with other constructions. For numerals other than ‘one’, Num (*-ER*) N.SG is a Composite Matrix Language construction fully calqued from Arabic, while Num (*-EL / D*) N<sub>Ar.SG</sub>, with or without the Arabic genitive *D*, is a straightforward Embedded Language Island fully copied from Arabic; the former occurs only in contexts where Destaing claimed the latter was obligatory. Both constructions, however, are also, rarely, attested with the Berber genitive *n* intervening between numeral and noun; Arabic does not normally permit an *n* to appear for numerals outside the range 11–19, and some attestations are outside this range. In the two cases where the following noun is Arabic, the result looks like an Arabic phrase structurally influenced by Berber; this unexpected, rare outcome may be called – to use a term not found in Myers-Scotton – a Composite Embedded Language Island. The rare construction N.SG Num has precedents in both Berber and Arabic (see 4.1); since it appears with Berber nouns, it is probably best interpreted as instantiating the Matrix Language. The contrast with Destaing is striking. In Destaing’s texts, only the two possibilities that make up classic codeswitching were found – Matrix Language phrases, and Embedded Language Islands. In this data, however, Composite Matrix Language phrases are substantially more common than Embedded Language Islands, and occur exclusively in the context (numerals ≥11) where, with some exaggeration, Destaing claimed the latter were obligatory.

To prove that, as Destaing suggests, the size of the numeral affects its chances of being followed by an Arabic noun, we must rule out the null hypothesis that the probability of the phrase being in Arabic when the numeral is greater than ten is the same as or less than the probability of the phrase being in Arabic when the numeral is less than ten. We can do this by grouping together the cells in the previous table according to the language of the noun and whether or not the numeral is greater than ten, and applying a Fisher's exact test to the result. (The rare construction Num  $n$  N<sub>Ar.SG</sub> is hard to classify as an island, due to the unexpected  $n$ , but this difficulty does not affect the hypotheses as phrased here; it is therefore included.) This gives us Table 5:

**Table 5.** Frequency by numeral and language of numeral+noun phrases in elicited data

	Berber	Arabic	Total
≤ ten	66	0	66
> ten	85	19	104
Total	151	19	170

This yields a 2-tailed p-value of 0.0001, which is statistically significant at the 5% and even the 0.1% level. (The latter remains true even if Num  $n$  N<sub>Ar.SG</sub> is excluded, and even if the boundary is changed from 'ten' to 'twenty'.) Thus we can safely conclude that, for modern speakers of Beni-Snous Berber, a numeral greater than ten is significantly more likely to form an Arabic numeral+noun phrase than a numeral less than or equal to ten, even in a data set systematically excluding measure nouns. In short, a weaker version of Destaing's claim still holds true.

## 5. Beni-Snous numeral syntax in comparative perspective

### 5.1 Berber

In eastern Rif Berber, according to Kossmann (2000, pp. 160–161), the normal construction above one is Num  $n$  N<sub>PL</sub>, except for a few Arabic loans such as *YYAM* 'days' and *CHUR* 'months' which do not take genitive  $n$ . For Figuig, a closely related variety a bit further south, Kossmann (1997, pp. 209–210) describes the situation in more detail. The normal construction above one is Num  $n$  N<sub>PL</sub>. However, for monetary units (all identical to Arabic), Num N<sub>SG</sub> is used (Num  $L$  N<sub>SG</sub> for 11–19), while for a small set of measure words including 'day', 'month', 'year', and 'hour', with well-attested Berber singulars, Num N<sub>Ar.PL</sub> is used for 2–10, and Num N<sub>Ar.SG</sub> for ≥11.

For the Central Maghreb, Laoust (1912, p. 58) briefly indicates that Chenoua uses Num ( $n$ ) N<sub>PL</sub> for all numerals greater than one; the same is true for closely related Chaoui, according to Penchoen (1973, p. 29), who gives no examples greater than ten. In Bissa, Genevois (1973, p. 67) describes a more complicated distribution: Num ( $n$ ) N<sub>SG</sub> for 'one' and Num ( $n$ ) N<sub>PL</sub> for 'two' through 'ten' (with  $n$  dropped only before vowels), but Num  $d$  N<sub>PL</sub> for numerals greater than ten.

The strategy Num ( $n$ ) N<sub>PL</sub> for numerals greater than or equal to 'two' is thus well-established within Berber among Beni-Snous' closest relatives, and is a retention. The frequent use of Arabic phrases borrowed whole for measures, particularly measures of time, is also widespread in Northern Berber.

## 5.2 Arabic

Regional Arabic uses significantly different syntax for numeral phrases. As in Berber, the numeral normally precedes the noun. Unlike in Berber, the noun is singular whenever the numeral is greater than ten. For two through ten, the noun is plural, except for certain measure nouns which remain singular. In some cases, especially for the teens, a dialectally variable pre-nominal linker intervenes between the numeral and the noun. In Tlemcen, according to W. Marçais (1902, pp. 160–161), *EL* (which he analyses as the definite article) is used as an obligatory pre-nominal linker after ‘one’ and the teens, while genitive *NTAÉ* and *(ED)D(I)* may optionally be used as linkers after any numeral. (This non-Classical usage of the genitive may itself be an early calque from Zenati Berber.) P. Marçais (1977, pp. 176–178) comments more generally that the use of genitive particles as pre-nominal linkers after numerals is common in Northwestern “Oranie”, while the pre-nominal linker used with the teens is usually *EN* (likely from Berber), but is *EL* in Tlemcen and Nedroma and in Moroccan cities, and *ER* in the “Sud oranais” and north of Taza in Morocco.

Beni-Snous Arabic conforms to these general patterns: the noun is plural for two through ten, singular from eleven up. The genitive particle is optionally used preminally after numerals, particularly ‘three’ through ‘ten’; it is typically *(N)TAÉ*, but in Oued Khemis also *D / DYAL*. For ‘eleven’ through ‘nineteen’, the pre-nominal linkers *EL* and *EN* are both attested, and even (only among older speakers) *ER*. The constructions Num (*n*) N.SG and (rare) Num *D / NTAÉ* N.SG in modern Beni-Snous Berber therefore reflect Arabic influence, whether the noun that follows is in Arabic or in Berber.

## 6 Explaining the results

### 6.1 Calquing and attrition

The empirical generalisation that “waning” languages frequently replace structures specific to the minority language with structures from the dominant language has a long history (cf. Dawkins, 1916, p. 198; Myers-Scotton, 2002, p. 164; Thomason & Kaufman, 1988, p. 102). Beni-Snous Berber is no exception. Syntactic differences between Arabic and Berber numeral constructions that were robust in Destaing’s time have been blurred in more recent data, as schematically illustrated by the following table, which excludes Embedded Language Islands:

**Table 6.** Range of numerals with which particular constructions can be used, by language of noun

	Arabic	Beni-Snous (contemporary)	Berber	Beni-Snous (Destaing 1907)	Berber
Num N.PL	2-10	≥2		≥2	
Num GEN N.PL	2-10	≥2		≥2	
Num N.SG	1, ≥11	1, ≥11		1	
Num <i>n</i> N.SG	(11-19)	≥11		–	
Num-LK N.SG	11-19	(11-19)		–	
N.SG Num	1	1 (5?)		–	

The most striking change since Destaing's time is the emergence of two new constructions calqued on Arabic for numerals  $\geq 11$ : Num N.SG, Num *n* N.SG. In terms of Myers-Scottons' (2002, p. 164) Abstract Level model, both constitute the copying of the Arabic morphological realization pattern "nouns counted by numerals  $> 10$  take the singular" into Berber – whereas, in Destaing's time, this pattern could be realised only within an Arabic Embedded Language Island. In Num N.SG, the calque is complete. In Num *n*.SG, however, the genitive *n* is retained from Berber: optional genitive marking is normal throughout Beni-Snous Berber numeral+noun constructions, whereas analytic genitive marking is impossible for speakers we consulted in the corresponding Arabic construction for numerals  $> 10$ , and the linker *-ən* is not used in Arabic with numerals  $> 19$ . In terms of Matras and Sakel's (2007) pivot-matching model, Beni-Snous Berber speakers have taken singular marking as the pivot of the Arabic construction, identifying the Arabic singular with the Berber singular, while making no effort to copy other, non-pivotal features such as the lack of genitive marking. Another feature apparently taken by most speakers as non-pivotal, the Arabic variable linker *-əl / -ər* used with numerals 11-19, is attested only in one non-island token.

## 6.2 Embedded Language Islands and cross-linguistic syntax conflicts

Above two, Beni Snous Berber and Arabic numerals are identical. Up to ten, inherited Berber syntax is harmonious with regional Arabic, as illustrated by Table 6: both use the order numeral-noun, both require the singular after 'one' and the plural after 'two' through 'ten', and both optionally allow a genitive marker between the numeral and the noun. From eleven onwards, however, they are in conflict: Berber demands a plural, while Arabic demands a singular. For the teens, moreover, Berber prefers the linker *n*, while regional Arabic prefers *EL*. With the systematic exception of measure words, bilingual suppletion is attested only with numerals for which Berber and Arabic noun selection requirement are in conflict.

Comparison to the only other reported case of bilingual suppletion, Domari, suggests that this is no coincidence. Domari's original numerals select for the singular, and 1-3 are all inherited. In Domari, bilingual suppletion is described as applying consistently with the nouns 'houses' and 'boys' from four to ten (Matras, 2012, p. 197), and with 'years' from four up. As Matras emphasises, 4-10 is precisely the range for which Arabic numerals are used where inherited Domari and Arabic noun selection requirements would have been in conflict, since Domari numerals normally select for singular nouns; below four, Domari numerals are used, while above ten, Domari and Arabic both require singulars. Measure words again show different behaviour: 'years' exceptionally remains in Arabic above ten in Domari, just as it exceptionally remains in Arabic below ten in Beni Snous Berber.

Recall Myers-Scotton's Embedded Language Island Hypothesis, quoted in Section 1: "When there is insufficient congruence between the lemma underlying an E[mbedded] L[anguage] content morpheme and its M[atrix] L[anguage] counterpart at one or more of the three levels of lexical structure, the only way to access the EL element is in an EL island." In terms of Myers-Scottons' Abstract Level model of lexical structure, numerals' noun selection requirements occupy the morphological realization patterns level within their lemmas. An Arabic numeral is incongruent at this level with its Berber counterpart if and only if the numeral is greater than ten, and with its Domari counterpart if and only if it is between 4 and 10. Substituting the particulars of this situation, her hypothesis thus predicts that: "When a numeral is greater than ten (in Berber), or between 4 and 10 (in Domari), the only way to access it as an EL element is in an EL island." Assume (cf. Section 1) that bilinguals can insert numerals shared by Arabic and Berber/Domari as Embedded Language content morphemes. Her hypothesis then implies that inserting numerals  $> 10$  (in Berber) or 4-10 (in Domari) as EL content morphemes should force the following noun to be in Arabic, while inserting other numerals as



Embedded Language content morphemes need not have any effect on the following noun, which in monolingual discourse should remain in the Matrix Language. This corresponds well to what is observed.

As seen in the previous section, modern Beni-Snous Berber speakers have adopted an optional new construction calqued on Arabic – unattested in Destaing – for numerals  $\geq 11$ : Num (*n*) N.SG. Insofar as this calque is symptomatic of increasing Arabic influence accompanying language attrition, it might be expected that, in elicited data, the use of Arabic numeral-noun phrases would have become even more common than it was for Destaing. Yet, while still present, it seems to have become rather less common: in elicitation Destaing appears to have consistently heard Arabic numeral-noun phrases above 10 (cf. 3.1), whereas in contemporary elicitation these occurred only rarely (cf. Table 5). This is again exactly what would be predicted on the hypothesis that selectional conflict is the driver for such Arabic islands, as Myers-Scotton’s Embedded Language Island Hypothesis would suggest; by making Arabic and Berber morphological realization patterns optionally congruent, calquing creates an alternative way to avoid the selectional conflict.

These two data points suggest that, where bilingual suppletion in numeral+noun combinations emerges, it will occur only following borrowed numerals whose noun selectional requirements in the source language differ from those in the recipient language. Bilingual suppletion with measure words is evidently not subject to this constraint, but such suppletion is categorically distinct, in that it involves only a very limited set of nouns and readily occurs even in languages where bilingualism has never been universal, such as Japanese.

### 6.3 Examining other factors

We have seen that the Embedded Language Island Hypothesis explains both the limited range of contexts for which numeral+noun Embedded Language Islands are attested and the apparent decrease in their frequency following syntactic calquing. It does not explain the widespread borrowing of numeral+measure noun phrases, and has nothing to say about calquing. Can the former facts be explained differently, in a manner that would also explain the latter generalisations?

Two obvious suggestions can be ruled out. The apparent decrease in the frequency of Embedded Language Islands is not the result of any overall decline in code-switching; as described in Sections 2.1 and 4, non-numerical code-switching is practically absent from Destaing’s corpus, and well-attested in the modern elicited data. Nor can it be explained as a side-effect of speakers’ self-monitoring in the relatively unnatural context of elicitation; as seen in 3.1, Destaing’s grammar’s examples are absent from his texts and unlikely in discourse, and thus most likely derive from elicitation too.

Speakers undergoing attrition tend to generalise default constructions – here, presumably, Num (*n*) N.PL – at the expense of range-specific ones; but they also tend to prefer constructions matching their dominant language (cf. 6.1). Since established Num N<sub>Ar.SG</sub> matches Arabic perfectly, it should either be retained or replaced, if at all, by Num (*n*) N.PL, not by a calque, and especially not by a range-specific one. The opposite is observed: above ten, Num (*n*) N.SG has become more frequent than Num N<sub>Ar.SG</sub> (Table 4). Purism would confound this prediction, but, as discussed, modern speakers do not otherwise seem to avoid switches into Arabic any more than Destaing’s consultants did. Attrition thus does not adequately explain the modern Beni-Snous results, much less predict broader generalisations.

The role of frequency might be considered: Destaing’s texts suggest that numerals  $>10$  are in practice more often combined with measures than with non-measure nouns, and Northern Berber in general prefers to express measure phrases in Arabic. Yet Domari, with the same preference for Arabic measure phrases, limits bilingual suppletion to 4-10. Frequency cannot simultaneously explain both outcomes.

The principal difference between both languages' inherited Num N. {SG|PL} constructions and Arabic Num N. {SG|PL} lies in the range restriction (for both languages, none; for Arabic, respectively 2-10,  $\geq 11$ ). One way to derive bilingual suppletion would be through taking this range restriction as pivotal (cf. 6.1) and (variably) copying it, leaving speakers with no way in their own language to express numeral+noun for numerals outside the range and thus inducing them to codeswitch to Arabic. Later, Beni-Snous speakers filled the gap by calquing Arabic's Num N.PL, identifying the noun's number as pivotal. This account captures the relationship between bilingual suppletion and selectional conflicts, and unifies it with calquing, but does not explain the important role of numeral borrowing.

No hypothesis under examination explains the widespread use of Arabic numeral+measure noun phrases; all languages concerned already had countable words for 'day' and 'year'. For the moment, one can only observe that such phrases are more commonly borrowed cross-linguistically, even without general bilingualism (see 3.2). However, this does not contradict any hypothesis under examination either.

The Embedded Language Island Hypothesis is thus the only one examined to explain both why numeral+noun Embedded Language Islands are attested only with borrowed numerals whose noun selectional requirements in the source differ from those in the recipient, and why their frequency should decrease following syntactic calquing.

## 7. Conclusion

Destaing overstated his case in claiming that the grammar of Beni-Snous Berber requires the insertion of embedded Arabic islands to express numeral+noun phrases whenever the numeral is  $>10$  – even in his time, and even for measures of time, Berber nouns could be used. However, in the non-elicited data represented by Destaing's corpus,  $>80\%$  of such phrases are in Arabic, and their embedded island status seems to be confirmed by their impermeability to Berber conjunctions. In elicited data, where the pressure to avoid Arabic might be expected to be more intense, Destaing's consultants seem to have produced exclusively Arabic forms. Ours, a century later, generally prefer Berber forms, yet sometimes refuse to produce versions with the corresponding Berber noun, even under prompting and immediately before or after using the Berber noun. In short, Beni-Snous Berber still allows – and formerly preferred – numeral+noun phrases  $\geq 11$  to be expressed as Arabic Embedded Language islands within discourse intended to be monolingual. Both Beni-Snous Berber and Domari data suggests that the switching is motivated by conflict between Arabic and non-Arabic selectional requirements, as outlined in 6.2.

While this appears to be a necessary condition, however, it is not a sufficient one. Such syntactically conditioned switches to Arabic nouns are not reported for any other Berber variety, even though most Berber varieties have equally extensively borrowed both numerals and certain lexically restricted numeral+noun combinations from Arabic, and even though many smaller Berber communities are almost entirely bilingual. The restricted geographical distribution of this phenomenon and its continuity over more than a century suggest that community-specific conventions play a role: while inter-grammar conflicts motivate the production of islands like these, it takes community acceptance (cf. 13 and 14) to change their social status from production errors to be corrected into a normal part of discourse intended as monolingual.

## Acknowledgements

We are grateful to the speakers interviewed for their help, and to Evangelia Adamou, Maarten Kossmann, and anonymous reviewers for their comments.

## Funding

This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors. It was indirectly supported by the authors' institutions.

## Notes

1. In games, Beni-Snous children still use a distinct counting series up to ten: *maħu* (older speakers' *wahū*), *tanu*, *teltu*, *rabu*, *xamu*, *celta*, *ceeba*, *qejet*, *mejjet*, *miw*. Up to seven, it consists of creatively distorted Arabic forms. Children elsewhere in the region use similar series (Grimme, 1926; Souag, 2010, sec. 4.2.1.1).
2. Berber typically uses a morphologically marked form of the noun, the construct state, for postverbal subjects and the objects of prepositions; cf. Mettouchi & Frajzyngier (2013). According to Destaing, the construct state was used for Berber nouns following all numerals – contrast *fus* 'hand' with *iğ ufus* 'one hand' – but modern speakers often generalise the citation form.

## References

- Basset, R. (1895). *Etude sur la Zenatia de l'Ouarsenis et du Maghreb Centrale*. Paris: Ernest Leroux.
- Dawkins, R. M. (1916). *Modern Greek in Asia Minor; a study of the dialects of Sīli, Cappadocia and Phārasa, with grammar, texts, translations and glossary*. Cambridge: Cambridge University Press.
- Destaing, E. (1907). *Etude sur le dialecte berbère des Beni-Snous*. Paris: Ernest Leroux.
- Destaing, E. (1914). *Dictionnaire Francais-Berbère: Dialecte des Beni-Snous*. Paris: Ernest Leroux.
- Genevois, H. (1973). *Djebel Bissa: Prospections à travers un parler encore inexploré du Nord-Chéelif*. Alger: Fichier périodique.
- Grimme, H. (1926). Nachtrag zu A. Klingenberg's Studie über die Berberischen Zählmethoden. *Zeitschrift Für Eingeborenen-Sprachen*, XVII, 230–234.
- Ibn Khaldūn, A. Z. `Abd al-R. ibn M. (2012). *Tārīkh Ibn Khaldūn*. Alwaraq.net. Retrieved from <http://www.alwaraq.net/Core/AlwaraqSrv/bookpage?book=116&session=ABBBVFAGFGFHAAWER&fkey=2&page=2216&option=1>
- Kossmann, M. (1997). *Grammaire du parler berbère de Figuig (Maroc oriental)*. Paris: Peeters.
- Kossmann, M. (2000). *Esquisse grammaticale du rifain oriental*. Paris: Editions Peeters.
- Kossmann, M. (2013). *The Arabic Influence on Northern Berber*. Leiden: Brill.
- Lafkioui, M. (2007). *Atlas linguistique des variétés berbères du Rif*. Köln: Rüdiger Köppe.
- Laoust, É. (1912). *Etude sur le dialecte berbère du Chenoua, comparé avec ceux des Beni-Menacer et des Beni-Salah*. Paris: Ernest Leroux.
- Marçais, P. (1977). *Esquisse grammaticale de l'arabe maghrébin*. Paris: Maisonneuve.
- Marçais, W. (1902). *Le dialecte arabe parlé à Tlemcen: grammaire, textes et glossaire*. Paris: Ernest Leroux.
- Matras, Y. (2010). Contact, Convergence, and Typology. In R. Hickey (Ed.), *The Handbook of Language Contact* (pp. 66–85). Chichester: Wiley-Blackwell.
- Matras, Y. (2012). *A grammar of Domari*. Berlin: De Gruyter.
- Matras, Y., & Sakel, J. (2007). Investigating the mechanisms of pattern replication in language contact. *Studies in Language*, 31(4), 829–865.
- Mettouchi, A., & Frajzyngier, Z. (2013). A previously unrecognized typological category: The state distinction in Kabyle (Berber). *Linguistic Typology*, 17(1), 1–20. doi:10.1515/lity-2013-0001
- Myers-Scotton, C. (1992). Comparing codeswitching and borrowing. In C. M. Eastman (Ed.), *Codeswitching* (pp. 19–40). Clevedon: Multilingual Matters.
- Myers-Scotton, C. (1997). *Duelling Languages: Grammatical Structure in Codeswitching*. Oxford: Oxford University Press.
- Myers-Scotton, C. (2002). *Contact Linguistics: Bilingual Encounters and Grammatical Outcomes*. Oxford: Oxford University Press.

- Penchoen, T. G. (1973). *Étude syntaxique d'un parler berbère (Ait Fraḥ de l'Aurès)*. Napoli: Centro di Studi Magrebini.
- Poplack, S. (2012). Myths and facts about loanword development. *Language Variation and Change*, 24(3), 279–315.
- Poplack, S., Sankoff, D., & Miller, C. (1988). The social correlates and linguistic processes of lexical borrowing and assimilation. *Linguistics*, 26, 47–104.
- Souag, L. (2007). The Typology of Number Borrowing in Berber. In N. Hilton, R. Arscott, K. Barden, S. Shah, & M. Zellers (Eds.), *Proceedings for the Fifth Cambridge Postgraduate Conference in Linguistics* (pp. 237–245). Cambridge: Cambridge Institute of Language Research. Retrieved from <http://www.ling.cam.ac.uk/camling/>
- Souag, L. (2010). *Grammatical Contact in the Sahara: Arabic, Berber, and Songhay in Tabelbala and Siwa*. School of Oriental and African Studies, University of London.
- Thomason, S. G., & Kaufman, T. (1988). *Language Contact, Creolization and Genetic Linguistics*. Berkeley: University of California Press.

## Author biographies

**Lameen Souag** is a researcher at Lacito (Langues et cultures à tradition orale), a mixed research unit (UMR 7107) within the National Centre for Scientific Research, in France. His research interests include language contact in North Africa and the relationship between lexical and grammatical borrowing.

**Fatma Kherbache** is a teacher at the University of Abou Bekr Belkaid in Tlemcen (Algeria). As her research interests are in the field of sociolinguistics, she works on the Arabic varieties spoken in her birthplace, the Beni Snous valley (extreme west of Algeria), with a particular focus on issues related to dialect contact and accommodation, dialectal variation, and attitudes.

## Appendix: Abbreviations used

1	1 <sup>st</sup> person
3	3 <sup>rd</sup> person
A, Ar	Arabic
B	Berber
CST	construct state
F	feminine
GEN	genitive
Lg	language
LK	linker
M	masculine
N	noun
NEG2	second part of bipartite negation
Num	numeral
PL	plural
SG	singular