Borrowing and Language Vitality Assessment: Sardinian as a Donor Language in Comparison with Māori

Gianni Onnis

University of Bologna

Abstract

This study tests the relevance of borrowings for the assessment of the vitality of two endangered donor languages: Sardinian and Māori. The vitality of these two languages has been analysed by quantifying their influence on the majority languages with which they interact, respectively Regional Italian of Sardinia (RIS) and New Zealand English (NZE). A RIS Corpus has been created by combining the methodologies employed by Macalister (2006) and Rankine et al. (2009), two corpus-based studies which analyse the presence of Māori borrowings in NZE. Sardinian borrowings have been quantified and different borrowing patterns have been used to speculate on UNESCO's (2003) vitality factors in order to assess the vitality of Sardinian. The results of Macalister (2006) and Rankine et al. (2009) in NZE have been reinterpreted accordingly, in order to obtain evidence of the vitality of Māori. The two languages show a positive response to new media and might have increasing intergenerational transmission rates. Nevertheless, Sardinian shows more limitations in domains of use and lower prestige than Māori. The results obtained therefore confirm UNESCO's evaluations, which consider Sardinian as a severely endangered language and Māori as a vulnerable language, proving the relevance of borrowing for language vitality assessment.

Keywords: borrowing, donor languages, UNESCO, language vitality assessment, corpus linguistics

1. Introduction

Nowadays, the spread of new media in a globalised market has multiplied the chances for languages to come in contact with each other. Language contact can be a source of enrichment for a language, by contributing to its linguistic repertoire, but at the same time, it can act as a source of threat for its vitality, changing the linguistic habits of its speakers (Crystal 2000). Linguistic borrowing can play a role in the endangerment of a language, by substituting its words and structures, leading to a progressive reduction of its domains of use (Thomason and Kaufman, 1992). The literature on borrowing has traditionally focused its attention on recipient languages, probably as a natural consequence of the fact that the effects of borrowing are only visible in recipient languages. But what value has borrowing for a donor language? Can borrowing provide evidence of the vitality of a donor language?

The aim of this study is to test the relevance of borrowing for the assessment of the language vitality of two endangered donor languages: Sardinian and Māori. The two languages are spoken on opposite sides of the world, Sardinia (Italy) and New Zealand, but are involved in very similar language contact situations. In both cases, language contact has taken place between a local indigenous language, Sardinian and Māori, and a later imported language, respectively Italian and English. Over time, in both contexts, the imported languages have acquired the status of majority languages. Still, they have changed under the influence of the local indigenous languages, to the extent that new majority language varieties have formed: Regional Italian of Sardinia in Sardinia and New Zealand English in New Zealand. On the other hand, Sardinian and Māori have over time become endangered minority languages.

The methodology chosen to carry out this research is based on two corpus-based studies which investigate the use of Māori borrowings in New Zealand English: Macalister (2006) and

Rankine et al. (2009). Macalister (2006) focuses on the diachronic change of New Zealand English under the influence of Māori, while Rankine et al. (2009) seek for evidence of the support granted to the endangered Māori in New Zealand English newspapers. The methodologies developed in these two research works have been employed to conduct an investigation of the Sardinian context by the creation of the Corpus of Regional Italian of Sardinia.

The analysis of the corpus has focused on lexical borrowings, given the fact that this type of borrowing occurs with higher frequency in language contact situations (Winford 2008) and generally precedes other forms of borrowing (Field 2002). Sardinian borrowings occurring in the corpus have been analysed in terms of frequency and semantic category. Their features have been associated with different factors of language vitality proposed by UNESCO (2003), allowing speculation on the vitality of Sardinian. Finally, the results obtained by Macalister (2006) and Rankine et al. (2009) on Māori have been reinterpreted according to the same language vitality framework, allowing a comparison between the language vitality of Sardinian and the language vitality of Māori.

2. Literature review

The theoretical basis of this research consists of three main topics, which will be discussed in the following order. A first section will consider language vitality and its components, as presented in UNESCO (2003). A second section will deal with lexical borrowing and its relevance for the study of language vitality, while a third section will provide a characterisation of the Sardinian and the New Zealand context, introducing the investigations carried out by Macalister (2006) and Rankine et al. (2009).

2.1. UNESCO's vitality assessment

As Roche points out, "vitality is not a property of a language itself, nor of a population that speaks a language, but rather a description of the relationship between a language, its speakers, and its wider linguistic, social, and political context" (2017: 193). A definition of language vitality is therefore necessarily operational and depends on the sociolinguistic factors one decides to take into account. UNESCO (2003) proposes nine factors to analyse language vitality; they are shown in Table 1.

Table 1. UNESCO's language vitality factors

Factor 1	Intergenerational language transmission
Factor 2	Absolute number of speakers
Factor 3	Proportion of speakers within the total population
Factor 4	Trends in existing language domains
Factor 5	Response to new domains and media
Factor 6	Materials for language education and literacy
Factor 7	Governmental and institutional language policies, including official status and use
Factor 8	Community members' attitudes toward their own language
Factor 9	Amount and quality of documentation

Factor 1 is based on Fishman's (1991) concept of intergenerational language transmission, which refers to speakers' ability to pass on their mother tongue to the next generation. According to this factor, the vitality of a language depends on its diffusion among all generations in a community of speakers. Factors 2 and 3 are of a statistical nature since they refer respectively to the number of speakers of a language taken in isolation and to the number of speakers of a language within the ethnic, religious, regional or national community they belong to. Factors 4 and 5 are both concerned with language domains of use: the former concentrates on the topics and purposes for which a language is used, the latter is mainly

oriented to the space granted to a language in new mediums, such as broadcast media and the Internet. Factor 6 can be related to factor 9, since both factors refer to the availability of materials in a language, in the first case for the sake of education and literacy (established orthography, grammars, dictionaries and literature), in the second case in terms of documenting material (grammars, dictionaries, as well as extensive texts and audio-visual recordings). Factor 7 considers the legal status enjoyed by a language while factor 8 refers to the social status enjoyed by a language in a community. In this sense, the two factors seem to fit Spolsky's (2004) framework of language policy, where factor 7 represents the language management component, as it includes all explicit language planning efforts made by a government in respect to a language, while factor 8 coincides with the language ideology component since it is concerned with the attitudes of a community towards its language(s).

In conclusion, it must be remarked that "none of these factors should be used alone" (UNESCO 2003: 7). A factor, taken alone, does not provide enough information to assess the vitality of a language. Indeed, even speaker's totals can be misleading, as Crystal (2001: 12) explains: "In many Pacific island territories, a community of 500 would be considered quite large and stable; in most parts of Europe, 500 would be minuscule".

2.2. Lexical borrowing, language vitality and language prestige

Lexical borrowing can be defined as the process involving the transfer of lexical material from a donor language to a recipient language (Zenner and Kristiansen 2013). As Kuteva (2017) reports, lexical borrowing is traditionally opposed to structural borrowing. The former affects the lexicon of a language by importing foreign sound-meaning pairs (loanwords; e.g., English *caribou* from Mikmaq *qalipu*) or through the morpheme-by-morpheme translation of foreign words, phrases or idiomatic expressions (loan translations/calques; e.g. French *presqu'île* from

Latin *paen-insula* 'almost-island'). The latter implies the importation of phonetic, phonological, morphological and syntactic features, therefore leading to changes in the corresponding language components. Given the greater sensitiveness of lexis to innovation, compared to the other language components, lexical borrowing tends to occur more frequently than structural borrowing (Winford 2008). Furthermore, it tends to precede structural borrowing diachronically, to the extent that "a significant amount of lexical borrowing is expected before one finds evidence of other 'interferences', i.e., actual structural borrowings" (Field, 2002: 3).

Literature on borrowing deals with language vitality in relation to two main aspects of borrowing: social motivation and language domains. Myers-Scotton (2006) analyses the social motivations causing the introduction of borrowings in a language and introduces the dichotomy between cultural borrowings and core borrowings. As the scholar explains, cultural borrowings are introduced in order to fill gaps in the lexical store of a language when new objects or concepts are imported in a language's culture. This is the case of computer-related words, which have spread to other languages together with the introduction of computers. On the other hand, core borrowings duplicate meanings already existing in the recipient language. This is the case of the use of greetings from French into Tunisian Arabic or from Spanish into northern Moroccan Arabic. According to Myers-Scotton (2006), these borrowings occur as the result of cultural pressure from the donor language, which occupies a dominant position compared to the recipient language. As the scholar explains, by importing core borrowings, the recipient language "loses some of its vitality" (2006: 230) to the donor language. The effects of cultural pressure on language vitality are also taken into consideration by Thomason and Kaufman (1992), who rather focus on the language domains affected by borrowing. They propose a borrowing scale establishing different phases in the borrowing process, as a fixed order where "features lower on the scale will not be borrowed before features higher on the scale are

borrowed" (1992: 73-74). As the scholars explain, the lower the position occupied by a language feature on the scale is, the greater the intensity of language contact and cultural pressure needed to cause its introduction is. At the same time, the scholars remark that overwhelming cultural pressure can have a negative impact on the vitality of a language. Therefore, it can be concluded that borrowings on lower positions on the borrowing scale might act as indicators of different degrees of endangerment of the recipient language since their introduction is due to higher cultural pressure from the donor language. Particularly, in the case of lexical borrowing, the scholars propose that non-basic vocabulary is borrowed before basic vocabulary, followed by different types of functional words.

As has been explained, when considering the relationship between borrowing and language vitality, both Myers-Scotton (2006) and Thomason and Kaufman (1992) analyse borrowings as instances of the cultural pressure exerted by the dominant donor languages on the subordinate recipient languages. Still, as Haspelmath (2009) points out, the concept of cultural pressure is too vague and intangible. The scholar rather prefers to talk in terms of 'language prestige', still acknowledging the difficulty in quantifying this factor influencing borrowing. A framework to analyse this element is introduced by Mühleisen (2002), who identifies three components of language prestige: language status (the legal condition of a language), language attitudes (social evaluations of the language within its community) and language function (what actions can be performed with a language). According to the scholar, these three components are interrelated as in a continuum in which prestige change can proceed according to opposite directions. It can take place as a top-down change triggered by political turns, therefore starting as a change in language status and then spreading to language attitudes and functions. The scholar understands this process as 'planned change', as opposed to 'natural change', which first occurs as a change in the functions of a language, triggered and reflected

by language attitudes, before having an effect on its status. Within the context of endangered languages, the revitalisation of French in Quebec, through the stipulation of the *Charter of the French Language* (Bill 101) in 1977, is certainly a case of planned prestige change, while an example of natural prestige change can be found in the increasing use of Occitan in music after the 1970s. As the scholar points out, planned change is a radical and fast process. Conversely, natural change proceeds at a slow pace, since it takes time for (micro-)function changes to affect language status, which is "the most static language prestige component" (Mühleisen, 2002: 33).

As a conclusion, language prestige can change over time, shifting the balance between dominant and subordinate languages in a contact situation and affecting the borrowing habits of a language. In the same way, changes in language prestige can have different effects on the vitality of a language and linguistic borrowings can act as indicators of these changes.

2.3. The two contexts

Sardinia and New Zealand offer two similar cases of language contact on a different scale. Both cases show the competition between an autochthonous language and an imported language, respectively Sardinian and Italian in Sardinia and Māori and English in New Zealand. In both cases imported languages have had the upper hand, acquiring the status of majority languages, but at the same time, changing under the influence of indigenous languages, to the extent that new language varieties have emerged, respectively, Regional Italian of Sardinia (RIS) and New Zealand English (NZE). On the other hand, indigenous languages have lost ground becoming endangered minority languages. The full picture shows two tripartite systems formed by endangered autochthonous languages, imported languages and new varieties of the imported languages. The difference between the two contexts derives from their geographical dimension.

Sardinia is a region of Italy where RIS has developed as a regional variety of Italian, having its standard on the nearby Italian mainland. On the other hand, NZE is a national variety of English which has originated far away from its standard's homeland.

Still, language contact has proceeded according to similar stages in both contexts, particularly from the 1860s on. The entering of Sardinia into the Italian State in 1861 (Piredda, 2016) and the establishment of the Native Language act (1867) in New Zealand (Grenoble and Whaley, 2005) paved the way for the pursuit of assimilation policies in favour of the imported languages in both territories, leading to a progressive endangerment of the indigenous languages. It was not until the 1970s that the course of events was changed by the insurgence of language activism movements, which favoured the recognition of autochthonous languages: Sardinian gained official status in 1999 (Vacca, 2017) while Māori was officially recognised in 1987 (Spolsky, 2003). Despite their legitimation, Sardinian and Māori still face endangerment: they both appear in the UNESCO's Atlas of the World's Languages in Danger, where Sardinian is labelled as definitely endangered language and Māori as a vulnerable language (Moseley, 2010).

On the other hand, in both contexts, the influence of autochthonous languages on the imported majority languages has attracted academic interest only in recent times, during the 1980s. Loi Corvetto (1983) was the first scholar to acknowledge the existence of a different non-standard variety of Italian in Sardinia, which took the name of Regional Italian of Sardinia. Similarly, during the 1980s New Zealand English attracted real scholarly interest; successively, *The Dictionary of New Zealand English* (Orsman, 1997) was published. And if the Sardinian influence on RIS has not become a topic of great interest in literature, if not in terms of descriptive studies, such as Loi Corvetto (1983) and Lubello (2016), the Māori influence on NZE English has attracted the attention of many scholars, particularly in the last decades.

Among them, Macalister (2006) studied the Māori lexical presence in NZE from 1850 to 2000. Differently, Rankine et al. (2009) focused their attention on intentional uses of Māori borrowings in NZE newspapers, in a representative sample from 2007. Both investigations rely on newspaper language corpora collected by sampling news data from different New Zealand dailies. They rely on Stempel's (1952) constructed week sampling method, which allows the creation of one or more constructed weeks through the selection of every Nth issue in a year of publication, starting from a random point in time and respecting a gap which dictates a one-day incrementation at each choice.

Macalister (2006) employs a diachronic perspective on the study of Māori borrowings in NZE. Data was collected from four New Zealand dailies with a long history of publication. The scholar established a set of six indicator years (1850, 1880, 1910, 1940, 1970, 2000) and then sampled a constructed week from each newspaper in each indicator year. The choice was further restricted to all news and opinions written in New Zealand about New Zealand matters. The corpus was then analysed together with data collected from parliamentary debates and school journals according to the same set of indicator years. The scholar found an overall increase in the use of Māori loans, which grew from 3.29 Māori words per thousand tokens in 1850 to 8.8 words per thousand in 2000. Proper nouns made up more than 90 per cent of the Māori word total in each indicator year. Non-proper nouns were divided according to the categories of flora and fauna, material culture (words referring to tangible objects of a culture, such as instruments, weapons and buildings) and social culture (words referring to non-material aspects of a culture such as abstract and religious concepts, relations of kinship and actions). Their frequency in the NZE Corpus increased particularly from the 1970s onwards, suggesting a recent expansion of the Māori vocabulary used in NZE. Indeed, the use of non-proper nouns grew in terms of tokens as well as in terms of types. Most particularly, from 1970 to 2000,

social culture types grew exponentially compared to the other non-proper noun types. The scholar finally identified three phases characterising the interaction between Māori and NZE: a progressive colonisation phase (until 1880), a recolonisation phase (roughly from 1880 to 1970) and a decolonisation phase (from 1970 onwards). As Macalister (2006) explains, the progressive colonisation phase and the decolonisation phase showed openness to borrowing, with greater numbers of Māori types occurring, while the recolonisation phase showed resistance to borrowing, therefore showing a change in attitudes during a period of English assimilation policy.

Rankine et al. (2009) analysed the intentional use of Māori words in NZE newspapers from a synchronic perspective. The scholars gathered 740 news articles by sampling two constructed weeks from New Zealand newspapers issued from February to March 2007. The selection of news was more focused on Māori lexis and included only articles containing at least one word or phrase from an arranged set of keywords about Māori topics. The analysis of Rankine et al. (2009) focused on the occurrence of Māori borrowings with an alternative in English. The scholars identified 199 Māori types, the majority of which belonged to the Social culture category (words such as *haka* 'war dance', *hapū* 'sub-tribe' and *tino rangatiratanga* 'Māori control and self-determination'), therefore, confirming Macalister's (2006) findings. On the other hand, the results obtained by Rankine et al (2009) "showed little evidence of support for New Zealand's endangered indigenous official language" (2009: 187).

3. Data and methods

This investigation applies the research methods developed by Macalister (2006) and Rankine et al. (2009) to the study of RIS. A RIS Corpus was created on the model of Macalister's (2006) NZE Corpus. Data was collected from the only two Sardinian dailies with such a long history

of publication to allow a diachronic analysis of the type of Macalister (2006): *L'Unione Sarda* and *La Nuova Sardegna*. Six indicator years were defined (1893, 1910, 1940, 1970, 2000, 2018) in order to analyse the widest time span available, at the same time, allowing comparisons with Macalister's (2006) data, at least from 1910 to 2000. Indeed, the only exceptions are 1893, which is the first year in which both newspapers were issued on a daily basis for the entire year, and 2018, which was chosen in order to obtain information on most recent trends characterising the use of Sardinian borrowings in RIS. A constructed week was sampled from each newspaper in each indicator year and the newspaper's content was then restricted to news and opinions produced in Sardinia and concerning the Sardinian context. The corpus obtained counts 2,572,006 word tokens and covers a time span of 125 years.

The analysis of the corpus focuses on native borrowings both from a diachronic point of view, as in Macalister (2006), as well as in terms of intentional use, as in Rankine et al. (2009). The RIS Corpus was processed with the software AntConc, a corpus analysis tool which allowed the creation of word lists and concordances, i.e. lists of all the instances of a word and its immediate context, where Sardinian borrowings could be firstly identified and where their contexts of use could be furtherly examined when needed. Code-switches have been counted and analysed together with borrowings since the two transfer types have been considered as being part of a continuum where code-switching as individual behaviour facilitates the normalisation of foreign forms as borrowings accepted by the whole community (Field 2002). Sardinian words have been analysed with regard to tokens, types and the semantic categories of proper nouns, flora and fauna, material culture and social culture. Material culture loans included names of Sardinian food (*malloreddus*, *casizolu*, *civraxu*), traditional clothes (*berritta*, *bertula*) or cultural monuments (*nuraghe*, *domus de janas*). Social culture types included traditional dances (*durdurinu*, *pass'e dusu*) and traditional rituals (*ardia*. *candelaria*). A

distinction of material and social culture types was not always straightforward as in the case of the term *pandela*, which has a material meaning ('standard', 'flag') as well as a social meaning ('role in traditional horserace'), therefore, requiring a deeper analysis of its concordances. The focus has also turned to Sardinian words with an alternative in Italian, which have been considered as core-borrowings, according to Myers-Scotton's (2006) framework.

The difference from the other two studies on NZE consists in the way borrowings are considered. Borrowings are not regarded as factors of language change, as in Macalister (2006), nor as evidence of the support granted to an endangered language, as in Rankine et al. (2009). Borrowings are here considered as pieces of evidence of the vitality of endangered donor languages. For this reason, different borrowing patterns have been associated to different UNESCO's factors of language vitality, as shown in table 2.

Table 2. Language vitality factors and corresponding borrowing patterns

UNESCO's factors	Parameters for Corpus Analysis
1) Intergenerational Language Transmission	Borrowing frequency (hypothesis) Semantic categories (hypothesis)
2) Absolute Number of Speakers	-
3) Proportion of Speakers within the Total Population	-
4) Trends in Existing Language Domains	Semantic categories
5) Response to New Domains and Media	Borrowing frequency
6) Materials for Language Education and Literacy	-
7) Governmental and Institutional Language Attitudes and Policies Including Official Status and Use	Radical change in borrowing rates Presence of cultural/core borrowings
8) Community Members' Attitudes toward their Own Language	Moderate change in borrowing rates Presence of cultural/core borrowings
9) Amount and Quality of Documentation	-

A preliminary remark should be made about factors 2 and 3, whose purely statistical nature shows resistance to a borrowing-based analysis. These two factors are rather supposed to be studied through the use of censuses and for this reason they have not been considered here. The other UNESCO's vitality factors have been associated with four borrowing patterns: frequency, semantic categories, change in borrowing rates and presence of cultural/core borrowings.

The frequency with which borrowings appear in the corpus can be used to provide evidence on three UNESCO's factors. It can show the acceptance of indigenous languages in media (factor 5): indeed, the use of native borrowings in newspapers shows the extent to which an indigenous language is accepted by the press (as in Rankine et al. 2009), suggesting similar trends in other media. Native borrowings can also play a role in literacy (factor 6), contributing to the creation of an established orthography, when missing in the donor language, through the choice of one spelling alternative among others. They might also be considered as material for language documentation (factor 9) as they constitute records of the indigenous language. Still, these last two factors have not been considered here since the relevance of native borrowings for language education and documentation is supposed to be minimal.

Semantic categories are supposed to provide evidence of the language domains in which the donor language is used (factor 4) and where it should be best preserved. On this matter, borrowing frequencies and semantic categories can be combined to formulate hypotheses on the extent of intergenerational transmission (factor 1) of the donor language. Indeed, when language maintenance allows language transmission (Fishman 1991), the introduction of higher frequencies of native borrowings with the involvement of more semantic categories, in a specific period of time, might indicate that a bilingual community still preserves and transmits some kind of knowledge of the donor language to allow the understanding of such borrowings, particularly in the case of code switches. Therefore, a hypothesis can be formulated, which must be necessarily tested on the donor language, to make sure that the knowledge of autochthonous words in a bilingual community corresponds to an actual competence in the donor language.

Changes in the rate of borrowings can be associated with factors 7 and 8, based on Mühleisen's (2002) prestige change framework. According to this framework, the planned

change reflects changes in governmental attitudes (factor 7) and since it is a radical process, it should bring about dramatic changes in the occurrence of native borrowings. On the other hand, natural change is a gradual process which affects language attitudes (factor 8) before having an effect on language status, therefore, it should manifest itself by means of moderate changes in the native borrowing rate.

Finally, Myers-Scotton's (2006) dichotomy between cultural borrowings and core borrowings can be used to investigate on the prestige of the donor language, where higher frequencies of core borrowings are associated with greater institutional (factor 7) and social (factor 8) recognition of the donor language.

4. Results and discussion

The analysis of Sardinian borrowings in the RIS Corpus shows four main results:

- the Sardinian language has been increasingly used in news communication;

- intergenerational transmission might have received a new impulse after 1970;

- domains of language use are restricted;

- the government and the community have recently supported the use of Sardinian, but its prestige is still low.

As in Macalister (2006), the frequency of indigenous borrowings registered in each indicator year has been combined to create a diachronic picture showing the occurrence of Sardinian borrowings in RIS during the analysed time span (Figure 1.). The frequency of Sardinian borrowings decreases from 1893 to 1910, the year in which the lowest numbers of Sardinian tokens have occurred. 1910 also marks the starting point of a rising trend which reaches its peak in 2000. During this time lapse, the frequency of Sardinian borrowings increases from 1.1 to 7.2 words per thousand tokens, before decreasing again in 2018, where

6.2 Sardinian words per thousand tokens occur. On the one hand, the distribution of Sardinian tokens along with the corpus clearly reflects the phases which have characterised the Sardinian language policy, showing low borrowing frequencies during the period of Italian assimilation policy, a first dramatic increase in 1970, during the period of Sardinian language activism, and again in 2000, soon after the official recognition of Sardinian. On the other hand, considering the use of Sardinian in RIS newspaper language, the overall increase in the use of indigenous borrowings shows that the Sardinian language is coming into use in press communication and probably in other media as well.
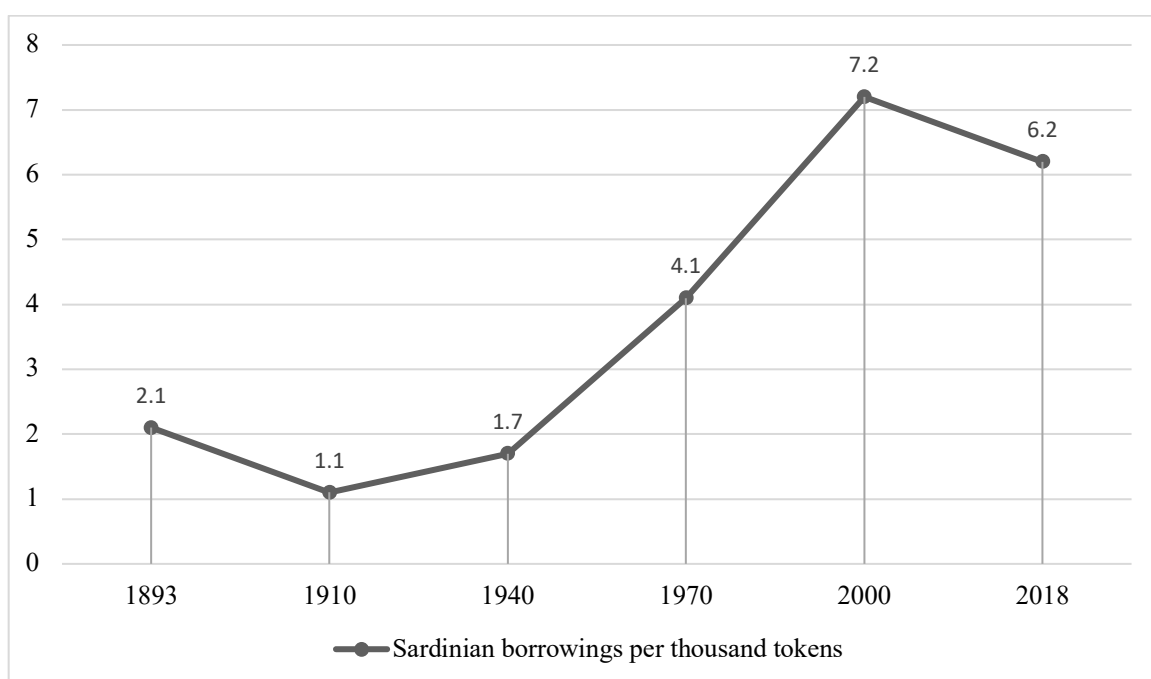

Figure 1. *Sardinian borrowings per thousand tokens*

The total amount of Sardinian borrowings in each indicator year has then been decomposed in terms of proper noun tokens and non-proper noun tokens (Fig. 2). It is not surprising to note an overwhelming presence of Sardinian proper nouns, which make up more than 94 per cent of the total Sardinian tokens in each indicator year, as this is a natural consequence of the referential style of news language. On the other hand, an analysis of the

distribution of Sardinian non-proper nouns, i.e. content words referring to flora and fauna, social culture and material culture, allows the formulation of a hypothesis on the phases which might have characterised the intergenerational transmission of the Sardinian language. If the presence of Sardinian content words is very low until 2000, making it difficult to obtain a clear picture of the process of intergenerational language transmission during the period of Italian assimilation policy, the dramatical increase of Sardinian non-proper nouns (both in terms of percentual distribution and absolute token numbers) in 2000 and in 2018 calls for considering an increment in the intergenerational transmission of Sardinian after 1970. But if intergenerational language transmission might have been strengthened, in which contexts might it be at work?
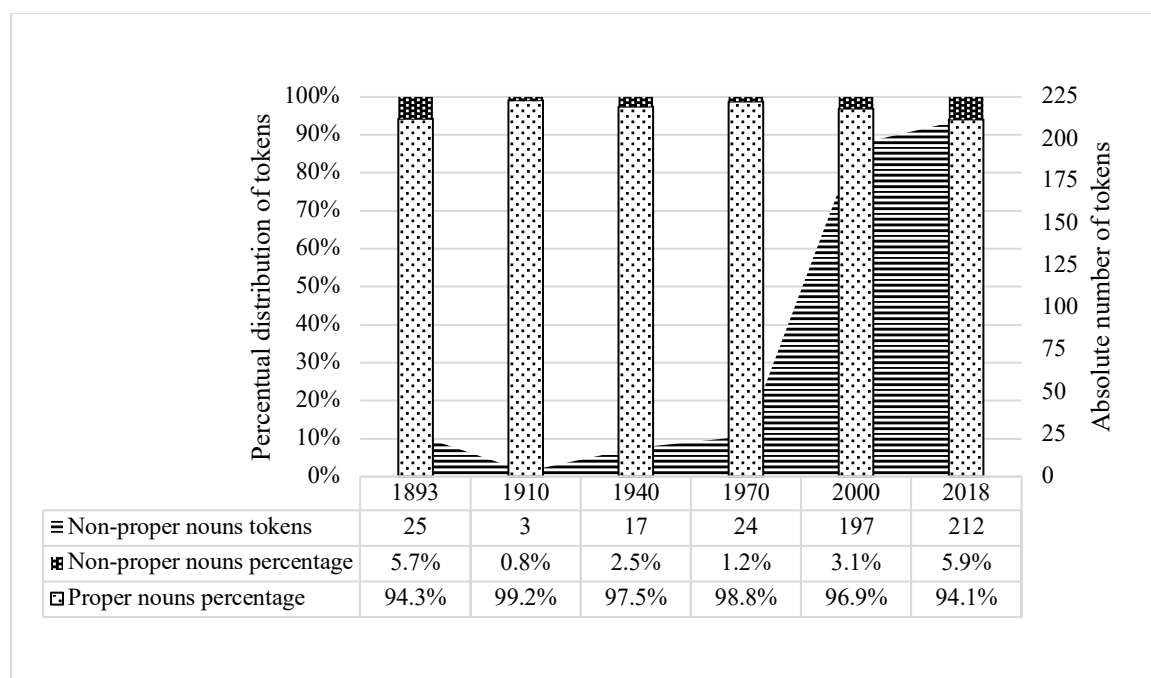

	1893	1910	1940	1970	2000	2018
≡ Non-proper nouns tokens	25	3	17	24	197	212
⊞ Non-proper nouns percentage	5.7%	0.8%	2.5%	1.2%	3.1%	5.9%
▢ Proper nouns percentage	94.3%	99.2%	97.5%	98.8%	96.9%	94.1%

Figure 2. Percentual distribution of Sardinian proper nouns and non-proper nouns in each indicator year

Table 3 shows the total amount of Sardinian non-proper noun types occurred in each indicator year and their distribution according to the semantic categories of flora and fauna, material culture and social culture. Altogether, the number of Sardinian non-proper noun types

found in the RIS Corpus totals to 88 types. If only 10 of them appear in 1970, their amount almost quintuples in 2000, where 49 types occur, while a moderate increase characterises the year 2018, which counts 64 types. Seen from the perspective of Mühleisen's (2002) prestige change framework, these variations in type rates can be associated with different processes of change in the Sardinian language prestige. The radical upsurge of type figures in 2000 can be considered as the result of planned change, acting as an indicator of the positive effects of governmental efforts to promote the Sardinian language during the 1990s. On the other hand, the moderate increase in the Sardinian type rate in 2018 would suggest a switch to a bottom-up natural change, fuelled by positive community attitudes.

Indicator years	Types total	Flora and fauna	Material culture	Social culture
1983	7	-	4	3
1910	1	-	-	1
1940	5	-	4	1
1970	10	-	7	3
2000	49	-	30	19
2018	64	1	38	25

Table 3. *Distribution of non-proper noun types in the RIS Corpus*

The semantic classification of Sardinian types occurred in each indicator year provides insights on the language knowledge which is supposed to have been transmitted throughout the analysed time span. The lack of Sardinian loanwords belonging to the category of flora and fauna is not surprising, given the great similarity between the Sardinian territory and the Italian mainland. On the other hand, if the categories of material culture and social culture appear to be almost equally represented in 1893, their proportion changes as the total amount of

borrowing types increases. Indeed, the indicator year 2018 shows the most unbalanced representation of these two categories: of the 88 types found in the whole corpus, 38 appear in the material culture category (43 per cent on the total) and only 25 appear in the social culture category (28,4 per cent on the total). This implies that the knowledge which has been transmitted from generation to generation has impoverished since the social culture vocabulary has been increasingly less transmitted. Indeed, some words seem to have been lost on the way: for example, the social culture words *gosos* ('religious chants in honour of saints') and *mutos* ('a genre of songs in rhymes') have been recorded only in 1893 and do not occur in the rest of the corpus. An analysis of the main topics touched in 2018 shows that almost all the types occurring in this year can be related to a few topics, which are: the Sardinian cuisine (names for typical dishes such as *coccoi*, *fregula*, *lorighittas*), Sardinian traditional garments (*berritta*, *cambales*), traditional celebrations (*iscravamentu*, *intregu*), traditional music (*canto a tenores*) and traditional dance (*annanza*, *pass'e dusu*). This means that the Sardinian language knowledge and use have been preserved only in restricted language domains, losing ground to Italian in others.

It is not a coincidence that the only flora and fauna type found in the corpus has been recorded in the indicator year 2018. This Sardinian word (*babaiole*, "ladybugs") is a clear example of core borrowing, according to Myers-Scotton's (2006) framework, and indicates higher prestige of the donor language both in terms of legal condition and social position enjoyed. Only 14 core borrowing types were found in the corpus. They all appear in the indicator years 2000 and 2018 and have a minimal proportion within the total word tokens of each year (Table 4). They show that, although the prestige of Sardinian has increased in the last decades, this language still occupies a subordinate position, both legally and socially, compared to the Italian majority language. The higher presence of social culture types is

probably due to the fact that Sardinian words related to abstract elements of culture are more prone to carry semantic nuances missing in Italian alternatives, compared to Sardinian loanwords referring to material objects of culture.

Indicator year	Types total	Semantic categories			Proportion of tokens within the total of the year
		Flora and fauna	Material culture	Social culture	
2000	10	-	4	6	0.004 %
2018	11	1	5	8	0.005 %

Table 4. Intentional use of Sardinian words in 2000 and in 2018

By comparing the results of this study to the results obtained by Macalister (2006) and Rankine et al. (2009) it is possible to identify two main differences between the vitality of Sardinian and the vitality of Māori. By a reinterpretation of Macalister's (2006) results, it is possible to state that, like Sardinian, Māori has been recently used more frequently in newspapers and its intergenerational transmission might have improved particularly after 1970, as a result of governmental support and positive community attitudes. Indeed, Macalister (2006) noticed an overall increase of Māori tokens throughout the corpus with a dramatic upsurge of non-proper noun types in 1970. Still, the higher presence of Māori social culture types detected in the NZE Corpus represents an opposite trend in respect to the distribution of Sardinian borrowings in the RIS Corpus. This suggests that, over time, the abstract aspects of the Māori culture have been better preserved than in the Sardinian case. In other words, it appears that, to a certain extent, Māori has maintained the link between material objects of culture and their deeper social meaning. Differently, the knowledge of Sardinian which has been handed down from one generation to the other has progressively impoverished so that what remains today of the Sardinian language are mostly material elements of culture which act as symbols of an

emptying Sardinian culture. The same happens to the language domains in which the two languages are used: a glance at the glossaries reported by Macalister (2006) and Rankine et al. (2009) suggests that Māori is increasingly used in institutional contexts, such as politics and the school, as well as in private households. On the other hand, Sardinian appears to be relegated to traditional and religious contexts, and its use outside these domains is sporadic.

Finally, a reinterpretation of the results obtained by Rankine et al. (2009) in terms of the presence of indigenous core borrowings shows that Māori enjoys more support at an institutional level and a better reputation in New Zealand society. Indeed, even if the rate of Māori core borrowings is very low in their corpus, it is far higher than the frequency of Sardinian core borrowings found in the RIS Corpus. The higher specificity of the data selected by Rankine et al. (2009), through the use of Māori keywords, certainly plays a role in the divergence between the results found in the Sardinian case. Still, this divergence must be related also to a difference in the prestige enjoyed by the two endangered languages.

In conclusion, the divergences existing between the use of native borrowings in the two contexts confirm UNESCO's statement of the vitality of Sardinian and Māori, with Sardinian being severely endangered and Māori being vulnerable.

5. Conclusions

The results show that the study of borrowing from the perspective of donor languages can provide evidence on at least five UNESCO's (2003) factors of language vitality: allowing the formulation of a hypothesis on the condition of a language's intergenerational transmission, defining domains of language use, proving its response to new media, and identifying the polarity of governmental and community attitudes towards it. In this sense, the study of

borrowings can provide practical data to support censuses' data when assessing the vitality of endangered languages involved in contact situations.

As far as Sardinian and Māori are concerned, the results obtained in this study are coherent with UNESCO's judgment on the vitality of these two languages. Nevertheless, more data gathered from oral sources in different contexts, from private to public domain, might provide much more evidence on the extent to which these languages are maintained and on their domains of use in more informal contexts.

Furthermore, the Sardinian and the Māori contexts offer very similar cases of language contact. Therefore, this method of analysis should be applied to different languages and different language contact situations, in order to measure the extent of its applicability, given the unpredictable ways in which languages borrow and land linguistic material from one another.

References

Crystal, D. (2000). *Language Death*. Cambridge: Cambridge University Press.

Field, F. W. (2002). *Linguistic borrowing in bilingual contexts*. Amsterdam: John Benjamins Publishing Company.

Grenoble, L. A., & Whaley, L. J. (2005). *Saving Languages: An Introduction to Language Revitalization*. Cambridge: Cambridge University Press.

Haspelmath, M. (2009). Lexical borrowing: Concept and issues. In M. Haspelmath, & U. Tadmor, (Eds.), *Loanwords in the World's Languages: A Comparative Handbook* (pp. 35-54). Berlin: De Gruyter Mouton.

Loi Corvetto, I. (1983). *L'italiano regionale di Sardegna*. Bologna: Zanichelli.

Lubello, S. (2016). *Manuale di linguistica italiana*. Berlin: Walter de Gruyter.

Kuteva, T. (2017). Contact and Borrowing. In A. Ledgeway, & I. Roberts (Eds.), *The Cambridge Handbook of Historical Syntax* (pp. 163-186). Cambridge: Cambridge University Press.

Macalister, J. (2001). Introducing a New Zealand newspaper corpus. *New Zealand English Journal*, 15, pp. 35-41.

Macalister, J. (2006). The Maori presence in New Zealand English lexicon, 1850-2000. *New Zealand English Journal,* 7 (1), pp. 1-24.

Mühleisen, S. (2002). *Creole Discourse: Exploring prestige formation and change across Caribbean English-lexicon Creoles*. Amsterdam: John Benjamins Publishing Company.

Myers-Scotton, C. (2006). *Multiple Voices: An Introduction to Bilingualism*. Malden: Blackwell Publishing.

Orsman, H. (1997). *The Dictionary of New Zealand English.* Auckland: Oxford University Press.

Piredda, N. (2016). L'italiano in Sardegna. *Italienisch: Zeitschrift für italienische Sprache und Literatur*, 38 (1), pp. 85-97.

Rankine, J., Moewaka Barnes, A., Borell, B., Kaiwai, H., Nairn, R., McCreanor, T., & Gregory, A. (2009). Intentional use of te reo Māori in New Zealand newspapers in 2007. *Pacific Journalism Review*, 15 (2), pp. 174-190.

Spolsky, B. (2003). Reassessing Māori Regeneration. *Language in Society*, 32 (4), pp. 553-578.

Spolsky, B. (2004). *Language Policy*. Cambridge: Cambridge University Press.

Stempel, G. H. (1952). Sample Size for Classifying Subject Matter in Dailies. *Journalism Quarterly*, 29 (2), pp. 333-334.

Thomason, S. G., & Kaufman, T. (1992). *Language Contact, Creolization, and Genetic Linguistics*. Berkeley: University of California Press.

UNESCO: United Nations Educational, Scientific and Cultural Organization. (2003). *Language Vitality and Endangerment*. Paris: UNESCO.

Winford, D. (2008). Processes of Creole Formation and Related Contact-Induced Language Change. *Journal of Language Contact*, 2 (1), pp. 124-145.

Zenner, E., & Kristiansen, G. (2014). Introduction: Onomasiological, methodological and phraseological perspectives on lexical borrowing. In E. Zenner, & G. Kristiansen (Eds.), *New Perspectives on Lexical Borrowing: Onomasiological, Methodological and Phraseological Innovations* (pp. 1-18). Boston: De Gruyter Mouton.