Prefazione agli *Occasional Papers del CeSLiC* – ISSN 1973-221X

Local Editorial Board

Full Editorial Committee

La serie degli *Occasional Papers* è una collana collocata all'interno dei *Quaderni del Centro di Studi Linguistico-Culturali (CeSLiC)*, il centro di ricerca del quale sono responsabile scientifico e che svolge ricerche nell'ambito del Dipartimento di Lingue e Letterature Straniere e Moderne dell'*Alma Mater Studiorum* – Università di Bologna.

Gli *Occasional Papers*, nati nel 2005, sono accessibili all'URL
http://amsacta.unibo.it/view/series/Quaderni_del_CeSLiC=2E_Occasional_papers.html

Finora sono stati pubblicati i seguenti saggi:

(2005) Fusari, Sabrina, Il direct mail per le organizzazioni nonprofit: analisi retorica interculturale italiano-inglese

(2005) Louw, Bill, Dressing up waiver: a stochastic collocational reading of 'the truth and reconciliation' commission (TRC)

(2005) Nobili, Paola, 'Saper vivere' con gli altri

(2006) Witalisz, Alicja, English Linguistic Influence on Polish and other Slavonic Languages

(2006) Larisa Poutsileva, Raccontare il mondo in lingue diverse: Sara' lo stesso mondo?

(2007) Mette Rudvin, Stereotypes of 'primitivism' and 'modernity' in immigrant-related discourse in the Italian media

(2007) Ana Pano, Anglicismos en el lenguaje de la informática en español. El "misterioso mundo del tecnicismo" a través de foros y glosarios en línea.

(2007) Sabrina Fusari, Idioletti e dialetti nel doppiaggio italiano de I Simpson

(2008) Alida Maria Silletti, La traduzione del futuro verbale in ottica contrastiva italiana-francese-inglese

(2008) Jane Helen Johnson, Corpus Stylistics and Translation

(2009) Sabrina Fusari, Il filmato turistico multilingue come discorso specializzato: il caso di studio della Val Gardena

(2009) Alida Maria Silletti, Analisi della "futurità" verbale in ottica comparativa italiana-francese-inglese

(2009) Antonella Luporini, *Frames*, *Transitivity Structures* e gerarchie di potere tra personaggi: Harry Potter affronta Lord Voldemort in *The Sorcerer's Stone*

(2009) Jane Helen Johnson, Towards an identification of the authorial style of Grazia Deledda. A corpus-assisted study

(2010) Cinzia Spinzi, 'How this holiday makes a difference': the language of environment and the environment of nature in a cross-cultural study of ecotourism

(2010) Goranka Rocco, Das Gerundium in italienischen Gesetzestexten und seine Umsetzung ins Deutsche

(2010) Sabrina Righi, L'African American Vernacular English: una varietà linguistica sovra-regionale

(2011) Fabio Regattin, Qu'est-ce que la mémétique? Et quel rôle peut-elle jouer pour la traductologie?

(2011) Mette Rudvin, Colonialismo, letteratura per l'infanzia e traduzione. Lo sguardo coloniale: Una lettura del testo e delle illustrazioni del *Libro della Giungla*

(2012) Ilaria Biondi, Traduzione e transfert culturale

(2012) Cinzia Bevitori, How green is "green"? A corpus-assisted analysis of environmental discourse across forms of journalism

(2012) Giulia Grata, Traduire la poésie: l'action du style. Erba et Luzi traducteurs de Michaux

(2013) Davide Vago, Traduire le Tartuffe de Molière. L'inscription de l'oralité

(2013) Sandra Petroni, Advermational territories on the web: Hybridity or resemiotisation?

(2014) Peter R. R. White, The attitudinal work of news journalism images – a search for visual and verbal analogues

(2015) Martina D'Avalos, Vox Populi vs Vox Political: un confronto linguistico tra elettori e candidati francesi per le elezioni europee 2014

(2015) Mariangela Picciuolo, European Identity: a Multimodal Perspective

(2015) Marina Manfredi, The language of popular science from the printed page to the Web: The case of the Table of Contents

(2016) Antonella Luporini, Spotlighting fantasy literature with the tools of Frame Semantics and Systemic Functional Linguistics: A case study

(2017) Alberto Biondi, Framing the Candidate: a Corpus-based Rhetorical Analysis of the 2016 Democratic Primaries in the USA

(2017) Alexanne Don, Negation as part of the Engagement Framework: Explorations in the territory Disclaim: deny

(2019) Angela D'Ambrosio, #BlackLivesMatter and hashtag activism in a critical discourse analysis perspective

(2019) Mona Chenani Saleh e Alexanne Don, Ideology and Attitudinal Shift in Persian Political Interpreting

A questi *papers* si aggiungono le altre pubblicazioni del CeSLiC – ossia, gli **E-Libri** – che includono:

1) la serie di manuali dei Quaderni del CeSLiC: **Functional Grammar Studies for Non-Native Speakers of English**, nata nel 2005, che già vanta sei volumi pubblicati (ISSN 1973-2228), il più recente dei quali è:
Miller, Donna Rose (2017) **"Language as Purposeful: Functional Varieties of Text. 2nd Edition"** a
http://amsacta.unibo.it/5504/

2) gli **Atti dei Convegni** patrocinati dal centro, nati nel 2005 (ISSN: 1973-932X), che comprendono i volumi:

- a cura di D. Londei, D.R. Miller, P. Puccini, Gli atti completi delle giornate di studio del CeSLiC del 17-18 GIUGNO 2005:
"**Insegnare le lingue/culture oggi: Il contributo dell'interdisciplinarità**", a
http://amsacta.cib.unibo.it/archive/00002055,
disponibile anche in versione cartacea:
Londei D., Miller D.R., Puccini P. (eds) (2006) *Insegnare le lingue/culture oggi: Il contributo dell'interdisciplinarità*, Atti di Convegni CeSLiC 1, Bologna: Edizioni Asterisco.

- a cura di Miller D.R. e Pano A., *Selected Papers* di quelli presentati al convegno internazionale CeSLiC del 4-5 dicembre, 2008, dal titolo:
"**La geografia della mediazione linguistico-culturale**/ **The Geography of Language and Cultural Mediation**", a
http://amsacta.cib.unibo.it/2626/
disponibile anche in versione cartacea:
Miller D.R. e Pano A. (eds) (2010) *La geografia della mediazione linguistico-culturale*, Selected Papers, Atti di Convegni CeSLiC 2, Bologna: Dupress.

- a cura di Miller D.R. e Monti E. (2014) *Selected Papers* di quelli presentati al convegno internazionale CeSLiC del 12-14 dicembre, 2012, dal titolo "**Tradurre Figure / Translating Figurative Language**", a
http://amsacta.unibo.it/4030/
disponibile anche in versione cartacea, pubblicata da BUP, Bologna, 2014.

- a cura di Cervini, C. (2016) *Selected Papers* di quelli presentati al seminario CeSLiC del 3-4 febbraio, 2014, dal titolo **"IN A TRA – Progetti per l'apprendimento linguistico: pluralità di obiettivi, metodologie e strumenti"**, pubblicati col titolo: "**Interdisciplinarità e Apprendimento linguistico nei nuovi contesti formativi. L'apprendente di lingue tra tradizione e innovazione"**, a http://amsacta.unibo.it/5069/

- a cura di Regattin, F. e Pano Alamán, A. (2017) *Selected Papers* di quelli presentati al seminario CeSLiC del 18 novembre, 2016, dal titolo **"Giochi di parole e traduzione nelle lingue europee"**, a http://amsacta.unibo.it/5749/

- a cura di Ivancic B., Puccini P., Rodrigo Mora M. J. e Turci M. (2018) *Selected Papers* di quelli presentati alla Giornata di Studi CeSLiC del 13 maggio, 2016, dal titolo **"Il testo letterario nell'apprendimento linguistico: Esperienze a confronto"**, a http://amsacta.unibo.it/5843/

Inoltre gli **E-libri del CeSLiC** comprendono anche i volumi compresi in:

3) la collana di **Studi grammaticali**, dal 2008 (ISSN: 2036-0274);

e

4) la collana di **Altre pubblicazioni** – AMS Acta, nata nel 2010 (ISSN: 2038-7954).

Tutte le pubblicazioni del centro CeSLiC sono disponibili a:

http://www.lilec.it/ceslic/i-quaderni-del-ceslic/

Sono lieta di presentare un nuovo e significativo *Occasional Paper*, di un giovane studioso, Gianni Onnis, nato a San Gavino Monreale nel 1992 e laureato magistrale in *Language, Society and Communication* presso l'Università di Bologna
Lo studioso parteciperà alla quattordicesima edizione della Lancaster University International Postgraduate Conference in Linguistics and Language Teaching (LAEL PG), dove presenterà questo lavoro, tratto dalla sua tesi magistrale dal titolo, "A corpus-based analysis of the influence of Sardinian and Māori in RIS and NZE dailies. Assessing the vitality of two endangered languages".

Il saggio che pubblichiamo ora si intitola:

A corpus-based analysis of the vitality of Sardinian.

A comparison with Māori

Come rileva l'autore, Il Sardo e il Māori sono due lingue a rischio di estinzione, appartenenti a sistemi linguistici simili quanto distanti fra loro. Entrambi i sistemi sono caratterizzati da insularità e contatto linguistico fra lingue autoctone e lingue successivamente importate. In entrambi i casi, la competizione tra codici linguistici ha elevato lo status delle lingue importate a lingue maggioritarie, mentre le lingue autoctone sono state relegate alla posizione di lingue minoritarie. Allo stesso tempo, la diffusione su nuovi territori ha portato le lingue importate a differenziarsi dai propri standard sotto l'influenza dei codici linguistici preesistenti, dando vita a due nuove varietà di lingua: l'Italiano Regionale di Sardegna (RIS) e l'Inglese della Nuova Zelanda (NZE).

Questo articolo presuppone la possibilità di misurare indirettamente la vitalità di due lingue in pericolo, come il Sardo e il Māori, attraverso un'analisi della loro influenza sulle lingue maggioritarie con cui competono, il RIS e il NZE. Sulla base della ricerca condotta da Macalister sul NZE Corpus (Macalister 2001; 2006a; 2006b) è stato quindi creato e analizzato il RIS Corpus. Un confronto fra la presenza lessicale del Sardo sull'Italiano Regionale di Sardegna e i risultati riportati da Macalister sul Māori dimostra come il Sardo venga rappresentato sempre più nei termini legati alla cultura materiale, a differenza del Māori, che trova progressivamente maggiore espressione anche nei termini astratti propri della cultura sociale. Conclusioni che acquistano ulteriore significato se messe in relazione con le valutazioni dell'UNESCO, secondo cui il Sardo è considerato una lingua 'sicuramente in pericolo d'estinzione' mentre il Māori è ritenuta un 'lingua vulnerabile'. Lo studio dei corpora viene così considerato come un valido strumento aggiuntivo nello studio delle lingue in pericolo che, affiancandosi alla linguistica e alla statistica, offre la possibilità di ottenere dati concreti sull'uso delle lingue a rischio e sul valore attribuito alle culture di cui si fanno espressione.

Keywords: language endangerment, Corpus Linguistics, sociolinguistics, Sardinian, Maori

Donna R. Miller

Responsabile scientifica del CeSLiC e *General Editor* dei Quaderni del CeSLiC

Bologna, lì 29 giugno 2019

A corpus-based analysis of the vitality of Sardinian. A comparison with Maori

Gianni Onnis

(University of Bologna)

1. Introduction

Research in language endangerment has notably grown during the last decades, particularly since the 1970s, in parallel with, and as a consequence of, a worldwide concern with the recovery of ethnic identity in an age of advanced globalisation (Fishman et al. 1985). Correspondingly, the study of the actual vitality of endangered languages has received great impulse and has begun to investigate the causes of language endangerment (Crystal 2000), while language surveys providing statistical evidence of the actual condition of endangered languages have multiplied. Within this research area, this study explores language endangerment from a particular perspective, restricting the field of investigation to language-contact situations set on insular settings. Indeed, language interaction and insularity are the two main features characterising the ecosystem of Sardinian and Māori, two endangered languages set on opposite sides of the globe. Both languages are autochthonous of a restricted territory, which ends where the sea begins, where they live side by side with later imported languages – Italian and English respectively – with which they interact on a daily basis. In both cases, the interaction between languages has changed the status of the competing codes, leading the former to become endangered minority languages and the latter to become majority languages. At the same time, the encounter of two linguistic codes has led to the generation of a new variety of the imported majority languages – Regional Italian of Sardinia (RIS) and New Zealand English (NZE) – both varieties displaying a great degree of contamination from the local autochthonous languages.

While research on Māori has bloomed in the last decades, the condition of the Sardinian language remains a less explored research area. This article represents, therefore, an effort to test the vitality of Sardinian by analysing its lexical influence on Regional Italian of Sardinia over time and then compare the obtained results with the findings of the research conducted on the relationship between Māori and New Zealand English. The two endangered languages are seen indirectly – for the information conveyed by the majority languages with which they interact, and by which they are threatened – and from a diachronic perspective. For this reason, Sardinian and Māori are not considered alone, but rather as being pieces of two language

systems, composed by Sardinian, Standard Italian and Regional Italian of Sardinia in Sardinia, and Māori, Standard English and New Zealand English in New Zealand.

The methodology chosen to conduct this investigation is corpus linguistics. Given the existence of a NZE Corpus (Macalister 2001), analysed by Macalister in 2006, a RIS Corpus has been created and analysed according to the same blueprint for the purposes of this investigation. Both corpora consist of newspaper language data collected according to different periods of publication.

Firstly, this article constructs a common base enabling the comparison of the two language systems. The Sardinian and the New Zealand contexts are presented according to three perspectives: language variation, language status and language endangerment. This threefold contextualisation serves as the basis to expound the history of the two settings and the related languages, a consideration of the local language policies ruling the relationship between endangered languages and imported languages in each setting, and the introduction of statistical evidence concerning the actual degree of endangerment of Sardinian and Māori.

The second section focuses on the methodology chosen to undertake this research, which takes Macalister's model of NZE Corpus as its theoretical basis (Macalister 2001, 2006a, 2006b). This section presents Macalister's research, before turning the attention to the composition of the RIS Corpus and the challenges related to its analysis, when applying Macalister's concepts to a different context. Indeed, Macalister's model of analysis is not only applied to a different socio-geographical context, but it is also related to a different theoretical framework, which presents the subject matter from the point of view of language endangerment. Unlike the present research, Macalister's study was originally conceived to investigate the diachronic change of New Zealand English under the influence of Māori, without focusing directly on the vitality of the Māori language.

Finally, the findings are shared, firstly concerning the condition of Sardinian in relation to RIS. They are then discussed together and compared with Macalister's findings, which are recontextualised under the framework of language endangerment.

2. Contextualisation

2.1. Three perspectives

The Sardinian and the New Zealand contexts can be seen under different lights, depending on the language or variety of a language that is chosen as the point of departure for an analysis.

For the purposes of this research, each context can be seen as constituting a system of three languages: a local autochthonous language, an imported language considered in its standard form, and the variety of the imported language resulting from the contact between the autochthonous language and the imported standard. The Sardinian system is composed by autochthonous Sardinian, imported Italian, and the resulting Regional Italian of Sardinia (from here on RIS). The New Zealand system sees Māori, English and New Zealand English (from here on NZE) as the languages playing respectively the same roles in the same type of language-contact situation.

According to this subdivision, three perspectives are supposed to reflect the same threefold composition of each insular context: language variation, language status and language endangerment. The order in which the three perspectives are presented aims to represent each language system from the wide-ranging linguistic relationships which exceed the borders of each insular setting to the less far-reaching interactions which take place within the microcosmos of each insular territory.

Under the perspective of language variation, this study addresses the relationship existing between the language varieties spoken in each context and their standards. In both cases, standards are set outside the borders of each insular territory. As far as language varieties are concerned, they show two main differences: they differ in the size of the linguistic area covered, and in the distance from their standards. Indeed, if RIS is a variety of Italian spoken in a region of Italy, therefore covering a territory of regional scale, NZE is a variety of English spoken by a whole nation-state. At the same time, while RIS has its standard on the nearby Italian mainland, NZE is set on the opposite side of the world compared to the place where Standard English was born. Still, RIS and NZE are products of the same process of language variation, due to geographical factors (importation of a standard in a different territory) and social factors (contact with languages spoken by other social groups). Seen in this light, the Sardinian context is nothing less than a small-scale representation of the New Zealand context. As concerns NZE, the variety can be categorised under the framework of 'global Englishes' (Jenkins 2015), where it counts as a 'new English' variety, since it is the result of the first diaspora of the English language. As Jenkins explains, new Englishes "developed independently of, and differently from, English in Britain partly because of the original mixtures of dialects among the people who settled in these areas, and partly because of the influence of the language of the indigenous populations" (2015: 27). From this point of view, New Zealand English is characterised by the influence of Māori, the language spoken by the native populations of New Zealand before the arrival of the English settlers. A categorisation of RIS must instead be made in relation to the

small-scale system of regional varieties of Italian. According to Cerruti (2009: 34), regional Italian is "a geographical variety of language, connected to the territorial provenance and distribution of (groups of) Italophones" (my translation)[1]. This subdivision holds for any language[2]: still, as in the case of NZE, RIS arises great scholarly interest due to its co-existence with the autochthonous language of Sardinia, the Sardinian language, and their reciprocal influence (Rindler Schjerve 2011).

The Sardinian and New Zealand contexts also differ when analysed from the perspective of language status, which focuses on the legal condition of a given language as majority or minority language, allowing an exploration of the relations existing between autochthonous languages and imported languages within each insular setting. Under this framework, the Sardinian context is constituted by the Sardinian minority language and the Italian majority language, while Māori and English constitute the New Zealand picture in the same order. As far as the Sardinian language is concerned, the Sardinian context must be broadened to account for the language policy of the Italian state, which recognises Italian as its official language (Law 15 December 1999, n. 482). On the other hand, the Sardinian language is one of the linguistic minorities officially recognised by the state, which, with a total amount of 12 linguistic minorities, is considered to be one of the European countries with highest linguistic complexity, as Lubello (2016) reports. According to the scholar, the Sardinian linguistic minority is progressively decreasing, but still counts as the largest linguistic minority in Italy. As concerns the Māori language, the New Zealand context appears to be even more complex than the Italian one. Firstly, because English is not recognised by the New Zealand government as an official language of New Zealand, but still "is the main language of communication in Aotearoa New Zealand and as such acts as a *de facto* official language" (Royal Society of New Zealand 2013: 3). Secondly, because the Māori language is only one of 160 languages spoken in New Zealand, at the same time being the only language which enjoys official status along with New Zealand Sign Language, even if it is spoken by a minority (Royal Society of New Zealand 2013). Still, Cunningham and King (2018) shed some light on the New Zealand reality, as they report that New Zealand can be considered as one of the most monolingual countries in the world, with English as the majority language spoken.

Finally, the perspective of language endangerment restricts the focus to the endangered language of each setting, allowing a consideration of Sardinian and Māori taken alone,

---

[1] Original text: "una varietà geografica di lingua, connessa alla provenienza e alla distribuzione territoriale di (gruppi di) italofoni".
[2] In the context of the United Kingdom for example, Bauer (2002) identifies regional varieties of English in England, Scotland, Wales and Ireland.

according to their degree of endangerment. Indeed, both Sardinian and Māori appear in the UNESCO's *Atlas of the World's Languages in Danger*, where the Sardinian language is labelled as a 'definitely endangered language' while the Māori language is enlisted as a 'vulnerable language' (Moseley 2010). This response might seem to be in contradiction with the number of speakers the two languages can count on – around 1 million for the Sardinian language (Lubello 2016), while only 148,395 for the Māori language (New Zealand Government 2013)[3]. Nevertheless, the number of speakers is only one of nine factors considered by the UNESCO Ad Hoc Expert Group on Endangered Languages[4], therefore more data are needed in order to have a grasp of the actual condition of the two languages. And, if statistics surely tells us much about Sardinian and Māori in terms of figures, this study is supposed to add another piece to the jigsaw, by testing the vitality of the two endangered languages through an analysis of their concrete usage in the two contexts.

2.2. Background information

Each perspective presented in the last section offers a starting point for the introduction of three essential topics related to the Sardinian and the New Zealand contexts. Under the perspective of language variation, the history of the two settings reveals the double value of the process of language change in the two territories: the process of Italianisation of Sardinia and the process of Anglicisation of New Zealand on the one hand, and the formation of RIS and NZE as varieties of the imported standards on the other hand. The language status perspective leaves space to the analysis of the language policies ruling the interaction between majority and minority languages in each context. On the matter of language endangerment, recent statistics provide an insight on the degree of endangerment of Sardinian and Māori.

2.2.1. History of places and their languages

Human history has begun at very different points in Sardinia and New Zealand, in step with the spread of humanity all over the world's landmass (Cavalli-Sforza 1996). Still, it is possible

---

[3] The number of Māori speakers was autonomously obtained through the NZ Stats table viewer tool, by selecting the field "language spoken by birthplace" and the variable "Māori".

[4] Factors are: 1) intergenerational language transmission; 2) absolute number of speakers; 3) proportion of speakers within the total population; 4) shifts in domains of language use; 5) response to new domains and media; 6) availability of materials for language education and literacy; 7) governmental and institutional language attitudes and policies, including official status and use; 8) community members' attitudes towards their own language; and 9) type and quality of documentation (UNESCO Ad Hoc Expert Group on Endangered Languages 2003).

to identify some common patterns typifying the linguistic history of the two territories. Indeed, it seems that language evolution has gone through similar stages in both settings, proceeding sometimes at the same pace.

Human presence in Sardinia is attested since the lower Palaeolithic times (Vona 1997), while a Proto-Sardinian language is supposed to have been spoken in Sardinia since the Bronze Age (Pallottino 2000). Nevertheless, it is possible to identify a Sardinian language only after the end of the Roman rule over the island's territory, during the IX century, when Sardinian independent kingdoms, called Judicates, arose, and Sardinian started to develop as an authentic Romance language, spreading with small variations over the island and at the same time showing a great degree of differentiation from the other Romance languages (Blasco Ferrer 1986). However, it was not a long time after that the Sardinian language came into contact with Italian, as the Maritime Republics of Genoa and Pisa started to forge intense relationships with the Judicates, from the year 1016 (Piredda 2016). According to Piredda, this period marks the beginning of the first Italianisation of Sardinia, with the ruling of the Maritime Republics over the island and the adoption of Italian languages.

While Sardinia was undergoing the first process of Italianisation, on the other side of the world, in New Zealand, human life was only just beginning. According to Harlow (2005), the archipelago of New Zealand might have been uninhabited until the XI or the XIII century, when navigators of Polynesian descent arrived with their *waka* (canoes). They originally called themselves *tangata whenua*, spoke one language, although with small dialectal variations, and spread all over New Zealand adapting to its differences in land and climate (Smith 2012). These people used to identify with their own tribes, and it was not until the arrival of Europeans that they started to call themselves Māori, meaning 'ordinary people of a place', therefore referring to their being natives, in contrast with the European newcomers (Waitangi Tribunal 2001; 2016). As Harlow explains, Māori belongs to the great family of the Austronesian languages, which counts as many as 1,200 languages, and covers an area stretching from Madagascar to Rapa Nui (Easter Island). The scholar also specifies that, within the linguistic family, Māori belongs to the Tahitic subgroup.

While the Māori culture was blooming in New Zealand, isolated from the outside world, Sardinia was going through four centuries of Spanish domination, from 1323 to 1720, when the island was annexed by the Kingdom of Piedmont, and the second and much stronger process of Italianisation of Sardinia slowly took place (Piredda 2016).

About fifty years later, in 1769, the Europeans arrived in New Zealand, led by Captain Cook (Smith 2012). This event radically changed the fate of New Zealand and its native populations.

The Treaty of Waitangi was signed in 1840, establishing the rule of the British Crown on the country, which became a new colony of the British Empire.

From the 1860s on, the history of Sardinia and New Zealand starts to proceed at the same pace, through similar stages.

In 1861, Sardinia became part of the unified Italian State, which saw Italian becoming the official language to be spoken, while Sardinian was relegated to the position of a dialect of Italian. Italianisation reached its peak in the 1920s, during the fascist regime, which pursued an assimilation policy, supporting Italian purism and recognising Italian as the only language of instruction, in spite of the previously used dialects (Farinelli 2009). As Piredda (2016) explains, during the XX century, the use of Italian in mass media communication and the phenomena of migration and urbanisation due to the industrial and economic development of the state created a condition of bilingualism with diglossia, with Italian gradually acquiring more prestige and becoming the only language of culture, as well as the main language used at work and in everyday life.

On the other hand, in New Zealand, the Native Schools Act (1867) marked the beginning of an only-English policy which established English as the only medium of education in Māori village schools (Spolsky 2003). The 1880s saw an exponential growth of the European (later called *Pakeha*) population, at the hand of massive immigration (Gordon et al. 2004). On the other hand, the Māori population critically decreased as a consequence of wars between Māori tribes, exacerbated by the use of European muskets, and the following wars between Māori and *Pakeha* for the ownership of the land (Barnard 2007). Another lethal factor is to be found in the spread of diseases brought by the immigrants, to which the natives had little or no immunity (Barnard 2007). As Lombardo (1995) concludes, at the end of the XIX century, the Māori were supposedly on the brink of extinction. All these factors, followed by the replacement of Māori in institutional and official domains, concurred in making Māori a disappearing language in the 1970s (Grenoble and Whaley 2005).

That said, if the imported language became the majority language of each setting, at the same time, it had to adapt to the new environment and to the languages spoken by its autochthonous populations. For this reason, the importation of a new language in each context gave birth to a long process of negotiation and accommodation between linguistic codes, whose result was the formation of new non-standard varieties of each imported language.
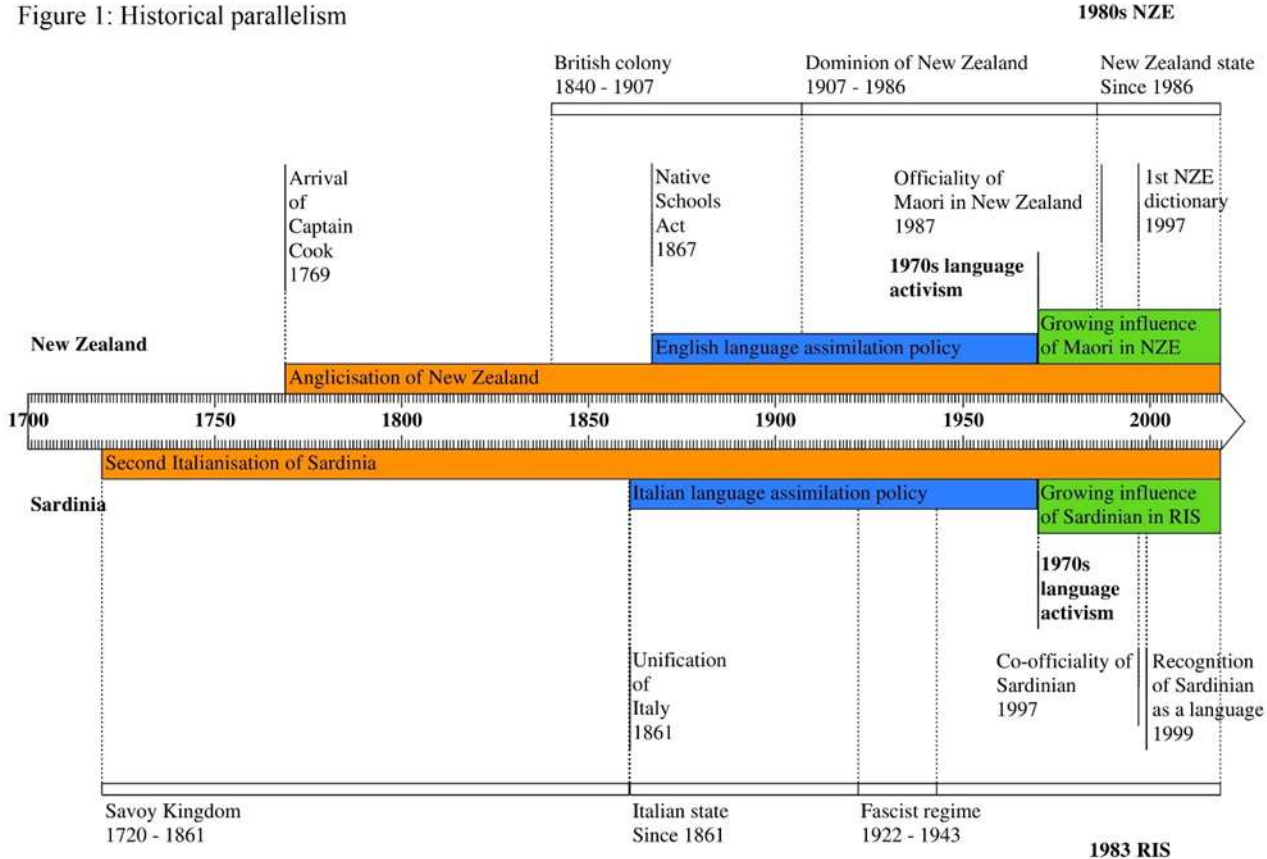
Indeed, as Italian spread all over the insular territory of Sardinia, it adjusted to the variety of Sardinian locally spoken, so that it did not remain the same as Standard Italian, but it rather evolved into RIS, as was baptised by Loi Corvetto (1983), the first scholar to distinguish this

Italian variety from the standard one. Piredda (2017) remarks that this variety acquired more prominence over time and changed according to the different linguistic areas of the island, to such an extent that it is now possible to identify the local provenance of a speaker. As already explained by Loi Corvetto, specific features are to be found in phonetics, morphosyntax and lexis. However, the variety is recognised only at an academic level, but it has not been codified, since there are no dictionaries of RIS. Still, it is interesting to note that the influence of the Sardinian language on RIS has become stronger in the last two decades, as reported by Dettori (2008), who analyses the increasing use of Sardinian lexis in the literary production of Italian starting at the end of the 1990s, probably as a result of linguistic policies which acknowledged the co-officiality of Sardinian together with Italian on the island (1997), as well as the recognition of Sardinian as a language (1999), both the product of previous movements of language activism of the 1970s (a theme which will be further explored in 2.2.2).

As far as the New Zealand context is concerned, Bell and Holmes (1990) report that some new features of the English spoken in New Zealand were already noticed at the beginning of the XX century. According to the scholars, what was denounced was the distortion of vowels and diphthongs, a problem which was meant to be solved by teachers in school, in order to restore the use of Standard English with no distinctive accent, for which the colony had been much applauded until then. According to their account, it was not until the 1950s that this prescriptive approach became descriptive, and NZE started to be recognised as a distinct variety of English, probably as the outcome of new attitudes among the New Zealand population, now mostly composed by New Zealand-born inhabitants, who were severing their sentimental bond with Britain as their 'Home', while accepting a distinct New Zealand identity. Evidence of this shift in identity perception can be found in the change in meaning of the label 'New Zealander', previously used almost exclusively to refer to the Māori, and from the early XX century on being increasingly used to identify the population of European descent (Cormack and Robson 2010). However, as in the RIS case, it was not until the mid-1980s that NZE became the object of real interest for linguists, and many questions began to be raised on matters of standard, geography of dialects, study of diachrony, to cite a few (Bell and Kuiper 2000). Unlike RIS, on the one hand, a standard variety of NZE has been accepted and codified while the first dictionary of NZE was published in 1997 (*The Dictionary of New Zealand English*, edited by H. W. Orsman). On the other hand, as in the case of RIS, NZE has lately been subject to increasing influences of autochthonous Māori, as remarked by Macalister (2006a), in this case again under the circumstances of language activism arisen during the 1970s, whose efforts led to the recognition of Māori as an official language of New Zealand (1987).

In conclusion, since the 1860s, both Sardinia and New Zealand have gone through similar historical periods from a linguistic point of view, with a prior establishment of assimilation policies supporting the use of the imported languages and a following sea change in favour of the local autochthonous languages after the insurgence of language activism movements during the 1970s, as can be seen in Fig. 1.

Figure 1: Historical parallelism

British colony 1840 - 1907	Dominion of New Zealand 1907 - 1986

1980s NZE

New Zealand state Since 1986

Arrival of Captain Cook 1769

Native Schools Act 1867

Officiality of Maori in New Zealand 1987

1st NZE dictionary 1997

1970s language activism

Growing influence of Maori in NZE

New Zealand

English language assimilation policy

Anglicisation of New Zealand

1700   1750   1800   1850   1900   1950   2000

Second Italianisation of Sardinia

Sardinia

Italian language assimilation policy

Growing influence of Sardinian in RIS

1970s language activism

Unification of Italy 1861

Co-officiality of Sardinian 1997

Recognition of Sardinian as a language 1999

Savoy Kingdom 1720 - 1861

Italian state Since 1861

Fascist regime 1922 - 1943

1983 RIS

2.2.2. Language policy

When talking about the relationship between Māori and English and their influence on each other, Spolsky (2003: 553) affirms:

> seen in its widest social, political, and economic context, this process can be understood not as a colonial language loss followed by postcolonial reversing language shift activities, but as the continuation of a long process of negotiation of accommodation between autochthonous Māori and European settlers.

It is easy to widen this vision to the Sardinian context, given the same situation of language contact, even if colonial discourse is less obvious, but still not missing[5]. In addition, in both contexts, the autochthonous language is the language of a minority within the state, and has recently received a new impulse fostered by movements of language activism and a revision of language policies in both cases.

Spolsky (2004) defines a language policy as the sum of three interrelated components: language practices, which refer to the habitual patterns according to which speakers choose among the varieties of a linguistic repertoire; language ideology, concerning the beliefs about languages and their use; and language management, which englobes all the efforts made by individuals or institutions with the aim of modifying language practices or ideology of other people. The scholar had used the same distinction when analysing the process of language shift from Māori to English in New Zealand, from the arrival of Europeans until the 1970s, concluding that

> the changed political and demographic situation had led to a change in language practice, with increasing use of English both inside and outside the Māori community. This was reinforced by governmental-imposed educational management with its commitment to monolingual English education. Some ideological support for the Value of the Māori language continued within the Māori community, but lacking support in language management activities – particularly policies for teaching the language to Māori children – language practice was moving rapidly in the direction of English (Spolsky 2003: 559)

The scholar also identifies a set of extralinguistic factors concurring to language shift, such as immigration and urbanisation, as has been previously explained in the historical section.

On the other hand, applying Spolsky's subdivision to the analysis of the Sardinian context, Blackwood and Tufi (2012) identify different causes for the language shift from Sardinian to Italian: the scholars highlight the historical lack of language management initiatives on the part of Italian state[6], while searching the causes of language shift in the language ideology advanced by the Italian scholar system, which adopted a punitive approach towards dialectophone pupils, therefore dismissing the use of Sardinian, previously considered a dialect of Italian. According

---

[5] Colonial discourse is present in publications concerning Sardinian literature and literature about Sardinia, as in Serra (2017) and Benvenuti (2015), or concerning the representation of the island reality on screen, as in Urban (2013). For a full account see: Serra (2017); Benvenuti (2015); Urban (2013).

[6] Blackwood and Tufi (2012) rather talk about Italian 'non-policies'. The only exception to this political attitude is to be found in the assimilationist and Italian purist measures pursued by the fascist regime (Farinelli 2009).

to the scholars, dialect was perceived in schools as an indicator of 'poor' Italian, an attitude which has been accepted by Sardinians, who in turn have changed their linguistic practices in favour of Italian, possibly in order to distance themselves from Sardinian as a marker of backwardness and unscholarliness.

Still, as has been previously explained, both Sardinian and Māori have later enjoyed a renewed interest, acquiring more prestige thanks to language activism pressures during the 1970s. From this perspective, it is interesting to note that phenomena of language activism are not exclusive to Sardinia and New Zealand but are rather part of a worldwide concern with ethnic identity, which arose from the mid-1960s to the mid-1970s, as Fishman et al. (1985) conclude. Therefore, it seems that this trend has acquired different names and forms according to place. In Sardinia, the revival efforts of the 1960s are known as *neosardismo*, and were coupled during the 1970s with the nationalist claims of the Action Sardinian Party, an autonomist movement (Tufi 2015). In New Zealand, the phenomenon has been called Māori Renaissance: as Mullaney (2010: 59) explains, the term refers to "the resurgence in indigenous cultural expression and production from the 1970s onwards, informed by the galvanizing influence of a range of indigenous political and cultural movements aimed at reversing the depredations of land rights, language and culture".

As concerns the outcome of language activism pressures in changing the language policies of Sardinia and New Zealand, evidence has to be searched in government official statements[7], therefore in language management efforts. On the other hand, the statistical analysis presented in 2.2.3 certainly provides some more concrete clues of the language practices and language ideologies surrounding Sardinian and Māori today.

Regarding the New Zealand context, Spolsky (2003) reports four grassroot efforts of language management aimed at changing the status of Māori in contrast to English during the 1970s: adult relearning of Māori, the *Kōhanga Reo*, which is a programme allowing Māori-speaking grandparents to pass their proficiency on to their grandchildren, the institution of Māori grassroots schools (*Kura Kaupapa Māori*), and the institution of Māori state schools, with the permission of the Ministry of Education. As the scholar reports, grassroot pressure payed off two decades later, around the 1990s, with the adoption of an explicit language policy, which the New Zealand government had managed to avoid enforcing until then: the Māori Language Act (1987), on the basis of which Māori was recognised as an official language of New Zealand and the Māori Language Commission was instituted, the latter assuming over

---

[7] As Spolsky (2004: 11) affirms: "the easiest to recognise are policies that exist in the form of clear-cut labelled statements in official documents".

time the role of a language academy and dealing with the problem of preserving the purity of standard Māori as well as implementing the Māori language policy with dispositions such as the Māori Language Act of 1989, the Education Act of 1989, the Broadcasting Act of 1989, and the Bill of Rights Act of 1990.

On the other hand, as far as the Sardinian and Italian context is concerned, Blackwood and Tufi (2012) also remark the vagueness of the Italian language policy before the 1990s: as they report, the 1948 Italian Constitution only mentioned linguistic minorities proclaiming the government's involvement in their protection, but did not even list the minorities to be protected. Unlike the New Zealand context, the movements of language activism of the 1970s did not lead to grassroot efforts for the promotion of the Sardinian language but, as Wells (2018) reports, the *neosardismo* movement rather found the support of experts, with the first aim to obtain the recognition of Sardinian as a language, since it was still officially considered as a dialect. In the account of Vacca (2017), it is clear how language activism finally payed off, this time after almost three decades, with the proclamation of a more explicit language policy: firstly, with the Sardinian Regional Law 26 of 1997, which acknowledged the co-officiality of the Sardinian language on the region, allowing its use in local and regional administration, followed by the state Law 482 (1999) on the matter of protection of historic linguistic minorities, among which the Sardinian minority was listed. As Blackwood and Tufi (2012) report, Law 482 was then followed by two attempts by the Region to provide a written Standard form of Sardinian: the Limba Sarda Unificada (Unified Sardinian Language, 2001), which raised great popular concern since it was mainly based on one variety of Sardinian, abandoned in favour of the Limba Sarda Comuna (Common Sardinia Language, 2006), a variety of compromise adopted for internal communication, but which has not yet enjoyed a diffused acceptance among the population.

2.2.3. Statistics

As has been said in 2.1., speaker figures taken alone are not consistent enough to show the level of endangerment of a language. As Crystal (2000: 11-12) explains:

> How many speakers guarantee life for a language? It is surprisingly difficult to answer this
> question. One thing is plain: an absolute population total makes no sense. The analysis of
> individual cultural situations has shown that population figures without context are useless.
> In some circumstances, such an isolated rural setting, 500 speakers could permit a

reasonably optimistic prediction; in others, such as a minority community scattered about the fringes of a rapidly growing city, the chances of 500 people keeping their ethnic language alive are minimal. 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. Speaker figures should never be seen in isolation, but always viewed in relation to the community to which they relate.

In line with this reasoning, this research has chosen two statistical studies on the basis of comparability and recency, in order to create a more comprehensive picture of the vitality of Sardinian and Māori today. Since wide statistical research can be found on the Māori language, while only a few statistical investigations are available concerning the Sardinian language[8], the choice was restricted to the last survey on the Sardinian language (Oppo 2007), which reports data collected in 2006, to be compared with the survey of the Ministry of Māori Development, or *Te Puni Kōkiri* (2008), likewise carried out in 2006. The two studies differ in the way data was collected, since research on Sardinian language use was carried out by interviewing a sample of 2,437 Sardinian citizens from municipalities scattered all over the Sardinian territory, while research on the Māori language was based on data collected from six different sources: censuses, administrative data and interview-based surveys. Furthermore, the two research studies adopted a different evaluation scale when assessing the proficiency of interviewees. In the case of Sardinian, the respondents were categorised as: capable to speak, capable to understand, none. In the case of Māori, all language skills were evaluated separately: speaking proficiency, listening proficiency, writing proficiency and reading proficiency, all of which were furtherly normalised in order to assess the overall degree of proficiency.

Nevertheless, the available data can be compared according to three categories: language knowledge and acquisition, language use and language status. As far as language knowledge is concerned, Sardinian appears to be more prominent than Māori, given the fact that 68% of Sardinians are able to speak Sardinian while only 23% of the Māori population are able to speak Māori. The two languages share a greater diffusion among the older population (over 50 years old), in terms of rate and level of competence[9], and their knowledge among the speakers is

---

[8] Statistical research on the Māori language has been carried out continuously in the last decades through autonomous projects or through the censuses of the New Zealand government, taking place every five years, while only two comprehensive studies on language use statistics are available on the Sardinian language since 1995, precisely a language use survey on minority languages carried out in 1995 on behalf of the European Union (Institut de Sociolingüística Catalana 1995), and a latter sociolinguistic research made on behalf of the Region of Sardinia in 2007 (Oppo 2007).

[9] *Te Puni Kōkiri* (2008) remarks the fact that young Māori speakers (under 35 years old) are more numerous than older speakers, since the younger age group constitutes a bigger slice of the Māori population.

rather passive, since speakers are overall more skilled in listening and reading tasks than in the spoken or written communication. A second divergence is related to the gender distribution of the two languages, with Sardinian being more typical of males, particularly in the younger age groups, while Māori female speakers count as the majority in all age groups. In the area of language acquisition, the two languages show opposite pictures. Oppo (2007) points out that the Sardinian language is mostly learned very early and in a familiar context, which involves parents or older relatives, while for children born to families with no Sardinian speakers, language acquisition takes place in interaction with childhood friends and schoolmates. Māori language acquisition appears to be rather entrusted to formal education and language learning programmes. Still, *Te Puni Kōkiri* (2008) also identifies TV and radio broadcasting, and the interaction with friends, relatives and neighbours as means of language acquisition. That said, it seems that Sardinian can still count on some sort of intergenerational transmission (Fishman 1991), even if language knowledge is being passed with decreasing frequency from generation to generation, while clear language management measures have been taken in order to restore the intergenerational transmission of Māori through the education system and the mass media.

The domains of language use clearly reflect the language acquisition practices of each language. Sardinian has its usual habitat in family interactions, while it is gradually less spoken in the public sphere, and even less in institutional or official contexts such as the school or the church, on a continuum. In this respect, Oppo (2007) denounces the lack of any support given by the school or any socialising agencies in the acquisition and use of the Sardinian language. On the other hand, the Māori language is taught in schools and used in governments documents, at the same time enjoying dedicated radio stations and television programmes. *Te Puni Kōkiri* (2008), however, complains about the limitedness of Māori's public domains of use, to be found in *hui* (Māori meetings), *marae* (Māori ceremonial places), and other religious activities where the Māori cultural element is dominant, still acknowledging the increase in language use in households.

As concerns language status, the term is used in *Te Puni Kōkiri* (2008: 7) to refer to "the position of the language within society". Therefore, this time, its meaning is not related to the legal condition of a language, but to the attitudes towards – and acceptance of – a language within a society. Against this backdrop, statistics provides evidence of the effect of recent language policies on language ideology in both settings. In both cases, language policies seem to have had positive effects on language ideology, but not without contradictions. Oppo's study on Sardinian (2007) reports a widespread consensus towards the study of Sardinian at school: still, views are divided on matters like using Sardinian as a medium for teaching or simply

studying it as a subject. In the same way, if interviewees agree on the use of Sardinian in public offices, different solutions are chosen to face the problem of intelligibility between different local varieties of Sardinian. As Oppo concludes, although 63.4% of Sardinians would like their language to be taught at school and used in public offices, they do not agree on the way these projects can be accomplished. These goals have already been achieved in the case of Māori, where *Te Puni Kōkiri* documents a call for the use of Māori in public domains outside the Māori dominant ones, as well as giving voice to a widespread urge for governmental initiatives able to promote the revitalisation of Māori. Still, the study denounces a disproportion between what is referred to as 'active engagement' and 'passive support', words which acquire concrete meaning when remarking the fact that 77% of Māori agree with the statement that all Māori should learn to speak Māori, but only 66% consider the language as a high personal priority.

2.3. The two varieties

This section provides a brief description of the two language varieties spoken in Sardinia and New Zealand, since they are the basis on which research has been conducted.

2.3.1. RIS

RIS is an Italian language variety resulting from the interaction between national language (Standard Italian) and autochthonous language varieties (Sardinian varieties) in Sardinia (Piredda 2017). As Loi Corvetto (1983) explains, on the island, Italian and Sardinian have not been kept separated, causing the permeation between language systems, so that linguistic items and patterns typical of Sardinian have entered the Italian language use. On the other hand, this overlap of linguistic systems has caused Italian elements, in turn, to come into use in Sardinian, as Piredda (2017) remarks.

Loi Corvetto (1983) provides a full account of the linguistic features which mark the distinctiveness of RIS as opposed to other regional Italian varieties, classifying them according to three categories, as mentioned in 2.2.1: phonetics, morphosyntax and lexis.

At a phonetic level, the scholar notes the phenomenon of metaphony, according to which tonic vowels are pronounced with a different degree of openness in relation to the following syllable in a word: for example, the word *collo* is pronounced [kɔllo] in its singular form, and [kolli] in its plural form. A second pattern occurring in RIS is the reinforcement of consonants as in *dico* and *levato* which are respectively pronounced as ['dːikːo] and [lːeˈvːatːo]. The scholar

further remarks a widespread voicing of fricatives and the dental affricate as in *zucchero* which is spelled as ['dzuk:ero].

At a morphosyntactic level, Loi Corvetto (1983) denounces a redundant use of *tutto* as in *chi tutto ha parlato* or *chi tutto c'era*. Emblematic of this variety is also the postposition of the verb, as in *arrivato sei?* or *sete ho*, as well as the postposition of adjectives when used as qualifiers, as in *c'è dell'acqua buona*.

As Piredda (2017) states, the peculiarity of RIS is to be found in lexis, more than in phonetics and morphosyntax. At the lexical level, Loi Corvetto (1983) talks about a positive or negative interference of Sardinian on lexical choices in RIS. This leads to the use of some Italian variants at the expense of others, which means that some lexical possibilities of the Italian language are preferred because of their similarity to Sardinian linguistic structures (positive interference), while others are discarded for the same reason in order to signal a distance from structures typical of the Sardinian language (negative interference). The scholar then turns her attention to what she refers to as dialecticisms, i.e. borrowings from the Sardinian language. Her list includes names of local Sardinian food (such as *civrasciu*, *cordula*, *fregula*, *guefus*, *malloreddus*, *piricchittus*), names of the typical Sardinian knife (*leppa*, *pattadesa*), together with more generic words such as *barracello* (armed guard), *ficchetto* (curious), *sussa* (beating), and *trassa* (affair). Loi Corvetto (1983) affirms that dialecticisms are rather rare in RIS and, for this reason, this study will inquire into the validity of this statement today. Indeed, Dettori (2014) expresses the need for new considerations on RIS, proposing new examples of the occurrence of Sardinian borrowings in RIS, particularly among the younger generations, as observed also by Colella and Blasco Ferrer (2016).

2.3.2. NZE

As has been previously explained, NZE is the variety of English spoken in New Zealand and can be categorised as a new English, according to Jenkins' (2015) model. According to the scholar, the variety originated from a mixture of English varieties and was at the same time subject to the influence of the indigenous Māori, particularly at a lexical level.

Deverson (2000) observes that the distinctive features of NZE in comparison to the other international varieties of English are to be found mainly in phonology and lexis. At a phonological level, Warren (2012) summarises the key features of the NZE vowel system according to two phenomena: a front vowel shift, which acts clockwise, leading the *trap* vowel to become close to the *dress* vowel, while the vowel of *kit* is centralised; and the merger of *near*

and *square*, so that the pronunciation of the two diphthongs is converging to a close variant [iə]. As far as the pronunciation of consonants is concerned, the scholar affirms that NZE does not show any very distinctive features. Still, the scholar notes a widespread non-rhoticity, accompanied by a recurring glottal stop as a variant of the /t/ sound, as well as the vocalisation of the /l/ sound when occurring in coda position.

At a lexical level, Deverson (2000) identifies a group of lexical items to be considered as specific of NZE. According to the scholar, these items, which he calls New Zealandisms, amount to approximately 5% of the total NZE vocabulary. The scholar subdivides New Zealandisms into four categories. The first category includes exclusively NZE words for exclusively New Zealand referents, such as *hongi* (traditional Māori greeting), *rimu* (a native tree of New Zealand), *marching girl* (a member of a team of girls who march in parades). The second category regards exclusively NZE words for objects which are not exclusive of New Zealand, such as *kai* (food), *cobber* (companion, friend), *waterside* (a worker appointed to the load and unload of a ship's cargo). A third category is composed of general English words which have acquired a supplementary meaning in NZE, such as *rifleman* (a bird native to New Zealand), *mob* (flock or herd of animals), *hotel* (pub). The last category refers to general English words which have a substitutive sense in NZE, as is the case for *football* (rugby), *creek* (stream), *paddock* (field or plot).

As can be seen from the previous examples, New Zealandisms include words of Māori origin. As Macalister (2006a) affirms, they are "the most distinctive feature of New Zealand English". The extent of their diffusion in NZE use will be further investigated in section 4.

3. Methodology

This study adopts a corpus-based approach to the study of language variation. Among the many fields of linguistic research on which the study of corpora has been applied for linguistic purposes, this research can therefore be categorised as being of sociolinguistic nature. As Mair (2009) reports, together with corpus linguistics, the field of sociolinguistics has experienced a notable expansion since the 1960s, particularly with Labov's development of the variationist approach. The scholar concludes: "the successive widening of the database both in corpus linguistics and in sociolinguistics has led to a blurring of formerly fixed boundaries and the emergence of a contact zone between the two subfields" (Mair 2009: 8). And if the blend of corpus linguistics and sociolinguistics represents a well-established theoretical common ground for the study of language varieties such as RIS, on a practical level, the need for sources in the

creation of a representative diachronic corpus finds one viable answer in the gathering of language data from the media.

Macalister's NZE Corpus (2001) constitutes a valid blueprint to follow when creating a corpus of RIS. In one of his works (Macalister 2001), the scholar tests the efficiency of a media language corpus which is supposed to be representative of the NZE variety in the year 2000. The scholar bases the collection of data on two assumptions: the aptness of newspaper language to represent a language variety, and the reliability of the constructed-week sampling method in the selection of the language data to analyse.

The same principles are followed in the creation of the RIS Corpus for the purposes of this analysis. At the same time, they assure the comparability of results with Macalister's (2006a) ones, allowing a comparison between the vitality of Sardinian and Māori.

3.1. Corpus Linguistics and Journalism

Bell (1991) makes a list of the reasons for studying media language[10]. Among them, this study focuses on the propensity of the media language to reflect many elements of the language naturally occurring in the wider society, at the same time reproducing its changes over time, besides the concrete accessibility of the data it supplies. Following Macalister's (2001) model, the choice of sources has been further restricted to the newspaper genre only, since it constitutes an older instance of media language, therefore allowing the study of a wider timespan, at the same time guaranteeing the same advantages of media language in general.

Therefore, on a first analysis, newspaper language data is needed since it allows the creation of representative corpora, capable to represent the RIS variety as it naturally occurs in Sardinia. Indeed, in this respect, Bell (1984) proved the tendency of the press genre to reflect the language of a society in his study on the phenomenon of determiner deletion[11]. According to Conboy (2010: 1), this trend must be understood as a marketing strategy:

---

[10] The scholar identifies eight main reasons: 1) accessibility of the media language data; 2) pure interest in some aspects of media language; 3) interest in the increasing occurrence of typical speech features in media language; 4) interest in the uncovered manipulation of language in media communication; 5) interest in how media communication can influence the language of society; 6) interest in the information conveyed by language about the media's structure and values; 7) interest in what elements of the society's language are reflected in the language of media; and 8) interest in the way media communication affects attitudes and ideology of the society (Bell 1991).
[11] Considered as a typical practice of the American press, the phenomenon of determiner deletion has spread in the language of British and New Zealand newspapers during the last decades, acting therefore as a display of the cultural shift towards American trends in the two countries. As Bell (1991: 110) concluded: "the determiner deletion variable is diagnostic of the social stratification of a news outlet's audience. In another country and another medium, the same pattern reappears".

the language of newspapers has always encapsulated what would sell to audiences and how information could best be packaged and presented to achieve this commercial end at any particular time. Newspapers have therefore always attempted to fit into the tastes of their readerships and sought ways to echo these within their own idiom, thereby reconstructing the 'original' audience in the process.

Nevertheless, the scholar points out that many other factors shape the language of newspapers: "readership, new technologies, politicians, journalists themselves all are capable of significantly altering the discourse of newspapers and their language has had to accommodate aspects of all of them" (Conboy 2010: 1). Therefore, acknowledging the fact that the language of newspapers is not a simple reproduction of the language of a society, Conboy's claim at the same time widens the perception of newspaper language as representing a whole historical period with its political and social events. This cannot but enrich the relevance of newspaper language for the analysis of long-term status negotiations between two languages in the same territory, as in the Sardinian case.

On a second analysis, Bell's (1984) study on the phenomenon of determiner deletion at the same time highlights the tendency of media language to act as an indicator of the changes occurring in a society's language. On the matter of newspaper language, Mair and Hundt (1999) remark its innovativeness, defining it as an 'agile' genre. As the scholars conclude: "newspaper prose is still first-rate material for linguists interested in ongoing change precisely because it is a written genre unusually receptive to (and in a good many instances also productive of) innovations" (Mair and Hundt 1999: 13).

Finally, newspaper language data are easily accessible and good in quality, as well as being preserved from any observer's alteration, as Bell (1991) affirms. Still, as the scholar suggests, the press supplies a huge amount of linguistic data per day and a sampling method is needed in order not to "drown in data" (Bell 1991: 13).

3.2. Constructed-week method

In order to work on a manageable quantity of data and at the same time assure the representativity of corpora, Macalister (2001) relies on the constructed-week method of sampling. The method was conceived and tested by Stempel (1952)[12]: starting from a random

---

[12] The scholar analysed the frontpage photographs appearing in a Wisconsin newspaper which was issued six days a week, with no Sunday edition. He made ten samples, each of 6, 12, 18, 24, 48 issues, on a total population of a year, in order to determine the representativeness of each sample size in relation to the whole population.

point in time, Stempel selected every Nth issue of a newspaper, at the same time establishing a gap which guaranteed the choice to move forward of one weekday at each selection. His conclusion was that 12 days, i.e. two constructed weeks, were enough to prove representative of a year of publication, while affirming that "increasing sample size may be poor investment of the researcher's time" (Stempel 1952: 333). Since then, the method has been widely adopted in media language research. Riffe et al. (1993) further tested the efficacy of the constructed-week method comparing it to other sampling methods, such as the simple random and the consecutive day methods. The scholars concluded that the constructed-week method is more reliable and more representative compared to the other sampling methods, and they also demonstrated that even only one constructed-week can be representative of a whole year of publication, depending on the type of analysis a researcher wants to conduct.

3.3. The NZE Corpus

Macalister's corpus of NZE was built according to two criteria: 1) it had to be representative of the written NZE variety; 2) it was intended to allow a diachronic analysis of the NZE lexicon.

As concerns the first requirement, the scholar selected four New Zealand dailies issued in different parts of the archipelago, in order to avoid any corruption of the gathered data by means of local colourings in NZE use, which could have occurred if the scholar had collected all the data from newspapers published only on the North Island or only on the South Island.

On the other hand, for the purpose of diachronic analysis, all the chosen newspapers had a long history of publication, all of them having been published with continuity since at least 1880. The scholar defined six indicator years (1850, 1880, 1910, 1940, 1970, and 2000) and sampled a constructed-week for each daily in each indicator year. As Macalister (2006a) explains, the thirty years gap between each indicator year was established on the assumption that it corresponds to the measure of a generation. For this reason, the selected constructed-weeks were supposed to create a synchronic picture of NZE use in each indicator year, therefore allowing a diachronic analysis of the language on an overall time span of six generations.

The content of the selected newspapers was further restricted to the news genre, according to Bell's (1991) model of media genres, which distinguishes between advertising and news as the main genres appearing in all media. Within the news genre, Macalister selected all articles which were supposed to contain NZE words, i.e. "any piece of news or opinion that was written in New Zealand about New Zealand" (Macalister 2001:37).

3.3.1. Macalister's analysis

Macalister's analysis departs from the assumption that "at some point between 1970 and 1990 an event, or events, triggered a significant shift in the relationship between the Māori language and New Zealand English" (2006a: 3). To test this hypothesis, the scholar created a corpus of 3,773,149 tokens, which he analysed alongside data collected from Parliamentary Debates and School Journals, obtaining a total amount of 5,422,546 tokens. As the scholar explains, adding data from different sources was needed for purposes of representativeness of written NZE. At the same time, a larger quantity of data increased the probability with which Māori borrowings could occur, given the fact that "words of Māori origin in New Zealand English are, by and large, low frequency words" (2006b: 90). Furthermore, Macalister highlights the fact that the majority of Māori words in NZE newspapers are actually proper names, as a consequence of the nature of media language, which is used to report facts and events taking place in a given place and in a given society.

In order to investigate the Māori lexical presence in NZE, the scholar makes a first distinction between proper names and non-proper nouns, subsequently subdividing the collected words according to 6 semantic domains:

1) place names;
2) personal names;
3) other proper nouns;
4) flora and fauna;
5) material culture;
6) social culture.

As he explains, the category of material culture is composed of nouns identifying concrete and tangible objects, while under the label of social culture are to be understood all the words, in this case non-proper nouns, referring to non-material elements of a culture, such as concepts, actions and relationships.

Firstly, Macalister notes that Māori word tokens have increased throughout the analysed time span: according to the combined sources, the frequency of Māori words increases from 3.29 to 8.8 Māori words per thousand tokens from 1850 to 2000. However, from a close analysis of each source, the scholar observes that the Māori lexical presence in newspapers and

parliamentary debates experiences a peak in 1970, while showing a downwards trend in 2000. Particularly, in the case of newspaper language, the frequency of Māori words tokens in 2000 is lower than in 1940. Indeed, as the scholar points out, the presence of Māori words in NZE newspapers amounts, overall, to 7.7 Māori words per thousand tokens.

According to the scholar, a second conclusion to be drawn from the NZE corpus analysis is that words of Māori origin occurring in NZE have gradually shifted in semantic domain. Indeed, Macalister demonstrates that the balance between proper names and non-proper noun tokens changes over time: more particularly, as concerns newspapers data, the use of non-proper nouns decreases from 1850 to 1970, while an upwards trend characterises the indicator year 2000. On the other hand, proper names picture a fluctuating trend, at the same time always outnumbering non-proper nouns in frequency figures.

This finding leads the scholar to a third conclusion: there has been an increase in the number of Māori words types during the 150 years span, which means that the Māori word vocabulary in NZE has grown. Indeed, an increase in the use of non-proper nouns reflects a more frequent use of words related to flora and fauna, material culture and social culture. Macalister observes that, among the three categories, social culture has recently found more space in newspaper language, as demonstrated by the peak in tokens figures in 2000, following a decreasing trend since 1850. But, most significantly, Māori words referring to social culture have not simply increased in number as a consequence of the repetition of the same words, but their number has also increased in type, together with nouns for flora and fauna and material culture. As the scholar concludes, the range of Māori word types being used in NZE has significantly increased from 1970 to 2000. The scholar reports the example of the increase of word types related to social culture in parliamentary debates: the social culture category displayed the second highest number of Māori word types in 1910, the indicator year in which only 8 Māori types were found, while in 2000, the number of types for this category reached its peak, with 63 types. A similar trend was found in newspapers. Nevertheless, given the fact that the trend describing social culture reached the lowest level in all sources in 1970, the growth in word types in this category from 1970 to 2000 was not only remarkable, but the fastest to be observed in all other time frames.

3.3.2. Two waves of borrowings

Macalister finally analyses the Māori types appearing in his NZE Corpus. He identifies 175 Māori types referring to material and social culture categories, 103 of which appear in DNZE

(Orsman 1997), while the other have no DNZE entry. By comparing the words reported in the DNZE, the scholar is able to retrieve information about the first introduction of each recorded word type, while the first use of the words missing in the dictionary is attributed to the first appearance of each of them in the corpus. According to this analysis, Macalister draws a model of the relationship between NZE and Māori from a diachronic perspective. The scholar distinguishes three periods: a first period of progressive colonisation, which lasted until the 1880s, followed by a second period of recolonisation, extending from the 1880s and the 1970s, and a final period of decolonisation, starting from the 1970s.

According to Macalister's results, the first period of colonisation is characterised by an openness towards Māori borrowings in English: as the scholar found out, during this period, 40% of the selected Māori types entered the NZE lexicon. He identifies this period as the first wave of borrowing from Māori. The period of recolonisation rather saw "a resistance to, or at least a stabilisation of, borrowing" (Macalister 2006a: 17). Indeed, the results show that only 13% of the Māori types came into use during this period. The third period of decolonisation is accompanied by a second wave of borrowing, which saw the introduction of another 32% of the selected Māori types in NZE.

Macalister's model of Māori-NZE interaction certainly tells us much about the changes in the New Zealand context over time. His categorisation acquires even more meaning when compared to an account of the history of New Zealand literature. In this respect, Lombardo (1995) identifies four periods characterising the NZE literature, which seem to perfectly fit into Macalister's model. According to Lombardo, written literature in New Zealand started very early, after the arrival of the European, with travel reports of missionaries and travellers. According to the scholar, this kind of tales set the basis for the development of a pioneer or early colonial literature, which was produced in the time span between 1860 and 1899. This first literary period can be seen as the correspondent of Macalister's first wave of borrowing, and indeed, as Lombardo reports, the literature of this period was characterised by a great interest in the Māori culture as well as in the biological peculiarities of the New Zealand Isles. The scholar further identifies two literary periods, which together can be associated to the period of recolonisation of New Zealand: late colonial literature (1890-1934) and New Zealand realism (1935-1965). According to the scholar, the literary production of this period showed a shift in focus to the problems characterising the new-born New Zealand society, without leaving room for Māori literature. The 1970s mark the start of the Māori Renaissance period, which corresponds to Macalister's second wave of borrowing. As Lombardo explains, the literature of this period was marked by a renewed interest in Māori culture, and transmits a widespread

awareness of being part of a bicultural society, among both *Pakeha* and Māori writers. Particularly, among the second group, some writers start to use the Māori language again, while others opt for the use of English and the typically European genres in order to communicate Māori values outside New Zealand borders.

After this reflection, it becomes clear how Macalister's NZE corpus acts as a testimony of the history of New Zealand. The phases identified by the scholar reflect the major changes in New Zealand society from a linguistic point of view. Indeed, the first wave of borrowing pictures the discovery of New Zealand by Europeans. The first contact with the new land and the Māori autochthonous tribes inevitably led to the introduction of new words in English: they were needed to describe the new environment and the Māori populations and to deal with them. This first period lasted until the 1880s, accompanied by the progressive inhabitation of the archipelago and the outnumbering of Māori.

The second period of recolonisation was characterised by the stagnation of Māori borrowings, which grew only slightly as a consequence of the process of urbanisation and, most importantly, of the institution of only-English policies. Their outcome is documented in the literature of the time, which saw the disappearance of the Māori language, a process which first took place in schools, and which contributed to a severe endangerment of Māori. As has been previously said in 2.2.1., during the 1970s, the Māori language was expected to face sure death soon.

The period of decolonisation reflects the inversion of this trend at the beginning of the 1970s. The following years would have seen the rise of grassroot movements of language activism alongside initiatives to revitalise the disappearing Māori language. This is testified in the NZE corpus by the increase in Māori tokens in the corpus, while the diffusion of more Māori word types for social culture portrays the renewed interest in Māori culture notified in the New Zealand literature of the time.

3.4. The RIS Corpus

The RIS corpus was created along the lines of Macalister's corpus: therefore, it was constructed with the aim of being representative of RIS and to provide a diachronic picture of its change over time. However, some obstacles were posed by the small-scale dimension of the Sardinian context.

More precisely, on the matter of representativeness, there are only two dailies with a long history of publication: *L'Unione Sarda* and *La Nuova Sardegna*. Still, it is assumed that the two

newspapers offer a reliable amount of data for a thorough analysis of RIS. Indeed, both newspapers are diffused all over the island territory, therefore their language should reflect some of the uses of RIS occurring within the whole Sardinian society. What is more, *L'Unione Sarda* was founded in Cagliari (on the southern part of Sardinia), while *La Nuova Sardegna* was first issued in Sassari (on the northern part): therefore, combining content sampled from both newspapers was assumed to provide representative data on the RIS variety, at the same time avoiding the overrepresentation of a local variety of RIS, compared to others.

On the other hand, the choice of the two dailies imposed some restrictions on the diachronic perspectives. Indeed, none of the two available dailies boasts publications before the 1880s. More particularly, *L'Unione Sarda* was first issued as a daily in December 1889 (Cecaro 2015), while *La Nuova Sardegna* became a daily newspaper in March 1892 (Cossu 2018). For these reasons, six indicator years were defined: 1893, 1910, 1940, 1970, 2000, 2018. 1893 was chosen as indicator since it is the first year in which both newspapers were issued with daily cadence for the whole year. The following four indicator years were defined following Macalister's method, for reasons of comparability with the NZE corpus. As for the last indicator year, although it cannot be representative of a new generation of RIS speakers, it is supposed to provide a further prospect on the most recent developments of RIS, at the same time representing a synchronic picture of this language variety as it is spoken today. As in the case of NZE, one constructed-week has been sampled for each newspaper in each indicator year. The selected population, counting 2,572,006 tokens, does not cover a time span of six generations but, covering a time period of 125 years, it is assumed to allow the drawing of some trends on the evolution of the RIS variety.

In this case too, the newspapers' content was restricted to all news written in Sardinia about Sardinian matters, since these are the types of news in which Sardinian borrowings have a higher probability to occur.

Table 1 shows the composition of both corpora.

Table 1.

	NZE Corpus	**RIS Corpus**
N° of tokens	3,773,149	2,572,006
Indicator years	1850; 1880; 1910; 1940; 1970; 2000	1893; 1910; 1940; 1970; 2000; 2018
Timespan covered	150 years	125 years
Sampling method	1 constructed-week for each newspaper in each indicator year	1 constructed-week for each newspaper in each indicator year

3.5. The RIS Corpus analysis

The RIS Corpus was analysed in accordance with Macalister's method of corpus investigation, nevertheless, this inquiry proposes a shift in focus from a study on language variation, which best suits the analysis proposed by Macalister, to a study on language endangerment. Indeed, Macalister assessed the extent of the Māori influence on the NZE variety, still, his attention was aimed to the drawing of a picture of the process of diachronic language variation characterising NZE. In this case, although this study is an attempt to measure the influence of Sardinian on the RIS variety, therefore being analogue to Macalister's research, the attention is directed to the study of language endangerment. It is therefore assumed that the influence of Sardinian on RIS is directly proportional to the vitality of the Sardinian language. For this reason, the Sardinian borrowings occurring in the RIS Corpus are considered as evidences of the vitality of the Sardinian language.

Furthermore, on the matter of comparability, it was necessary to reconsider some features of the RIS Corpus in order to apply Macalister's method of analysis.

Firstly, as far as Macalister's categorisation of tokens is concerned, the category of place names has shown a resistance to analysis, given the widespread Italianisation of toponyms and the rare existence of a perfect correspondence between the actual name of each municipality and its Sardinian version. In this regard, Podda affirms that variations have affected place names in Sardinia as a result of Italianisation or for propagandistic or economic purposes, "processes which have often led to the radical modification of a toponym's semantic value, together with the consequent loss of its primary reference, which was probably at first connoted by a strong

man-environment relationship" (2014: 996, my translation)[13]. Indeed, many examples of this phenomenon can be found in Pittau's (2018) study on the etymology of Sardinian place names. For these reasons, a decision was made not to include municipalities' names in the count of Sardinian borrowings, unlike place names referring to zones, localities and rural areas, which have been included in the counts of token and types, given the preservation of their Sardinian original name.

Secondly, as concerns Macalister's category for words referring to flora and fauna, no borrowings from Sardinian were found in the RIS Corpus. It is assumed that the main reason why this type of loanwords is missing is to be found in the analogy between the biological nature of the Sardinian island and the Italian landmass, a condition which has led to the development of linguistic equivalents in Sardinian and Italian. And even if some endemic species of flora and fauna do exist in Sardinia, they are often referred to with Italian names and the classifier 'Sardinian'.

Therefore, the analysis of the RIS Corpus has been based on five of Macalister's categories: place names, personal names, other proper names, material culture and social culture.

4. Findings

This analysis essentially makes four claims about the relationship between RIS and Sardinian since 1893:

1) there has been an overall increase in the use of Sardinian loanwords over time;
2) there has been a shift in the semantic domain of Sardinian borrowings;
3) there has been a recent increase in material culture and social culture types;
4) material culture types are acquiring more value in the RIS culture.


4.1. Diachronic increase in Sardinian tokens

Following Macalister's reasoning, a first investigation has focused on the frequency of Sardinian words used in the RIS Corpus according to each indicator year. Frequencies were normalised to the common base of a thousand tokens, as in the case of Māori.

---

[13] Original text: "Processi che hanno spesso determinato una modificazione radicale del valore semantico del toponimo, con conseguente perdita del riferimento primario che magari, inizialmente, poteva essere connotato da un forte rapporto uomo-ambiente".

The resulting frequencies picture an initial influence of Sardinian on the RIS vocabulary, with the occurrence of 2 Sardinian loanwords on a text-stretch of a thousand words. The trend reaches the lowest point in 1910, before gradually moving upwards until it reaches the peak in 2000, when 7 Sardinian borrowings occur for every thousand tokens. The rising trend changes its course in 2018, when the frequency of Sardinian loanwords is of 6.2 per thousand tokens, as in Fig. 2.

Figure 2: Number of Sardinian words per thousand tokens
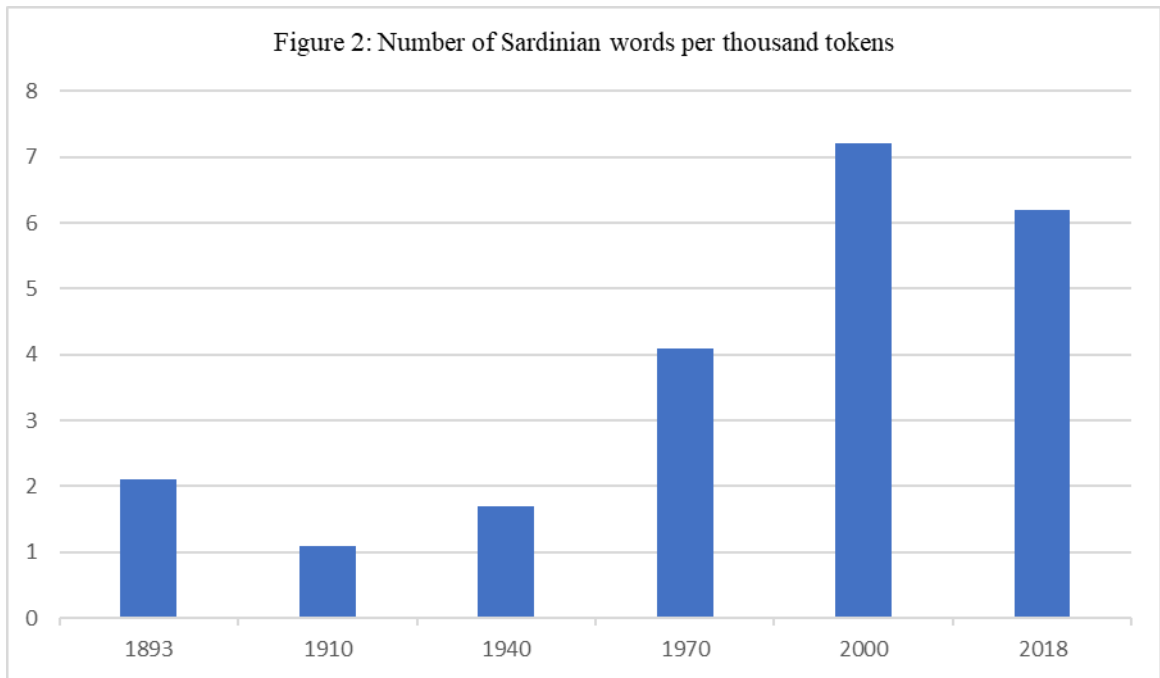
If the higher occurrence of Sardinian borrowings in 1893 finds an explanation throughout this research, it is difficult to establish a connection between the Sardinian historical context and the lowest occurrence of Sardinian words in 1910. Indeed, according to the historical background, the indicator year 1940 should display a process of Italianisation of the island at its peak, given the only-Italian polices initiated by the fascist movement during the 1920s. An explanation of these results can possibly be found in the pedagogic role played by newspapers, and by the media in general, in the diffusion of Italian in Sardinia at that time, as reported by Piredda (2016). Nevertheless, a comparison with different sources of data might shed some light on the matter.

On the other hand, as far as the distribution of frequencies after 1940 is concerned, the significant increase in the use of Sardinian borrowings in 1970, and again in 2000, clearly echoes the insurgence of movements of language activism during the 1970s and the following recognition of Sardinian as a language as well as its co-officiality with Italian at the end of the

XX century. The decrease in Sardinian lexical presence in 2018 might rather suggest, at a first glance, an exhaustion of the linguistic fervour in favour of the Sardinian language.

4.2. Shift in the semantic domain of Sardinian borrowings

An investigation on the balance between proper names and non-proper nouns found among the group of Sardinian borrowings occurring in the RIS Corpus clearly pictures the informative nature of newspaper language. Indeed, the most frequent Sardinian words appearing in the corpus are proper names, i.e. names of places and personal names – in other words, all references to places and people involved in the facts and events reported in each newspaper article. For this reason, the percentage of tokens for Sardinian proper names never goes under 94% in all indicator years, as shown in Fig. 3.

Figure 3. Percentual distribution of proper/non-proper nouns
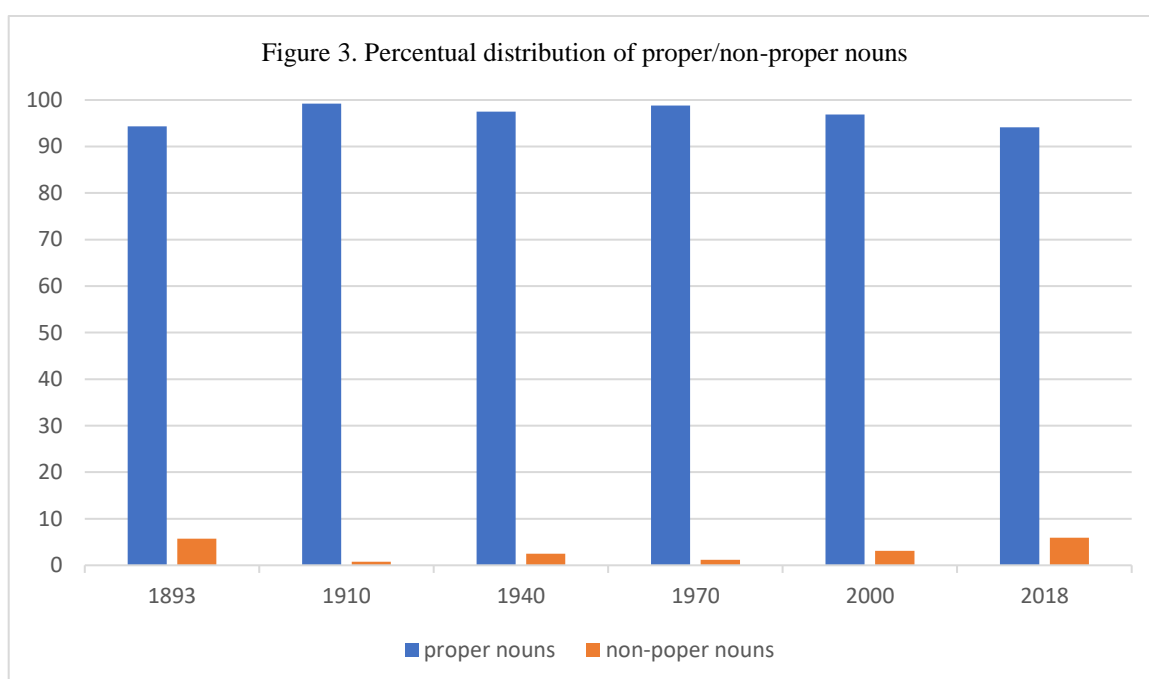
An analysis of the distribution of non-proper noun tokens rather shows the indicator years 1893 and 2018 as representing two peaks at the opposite sides of a fluctuating trend. As Macalister (2006a) explains, non-proper nouns constitute the categories of flora and fauna, material culture and social culture. A first question, considering the results of this type of analysis on the RIS Corpus, concerns the relatively high presence of non-proper nouns of Sardinian origin in the newspaper language of 1893. An explanation of this issue is to be found in the period in which these words have been first introduced in RIS.

Given the absence of a RIS dictionary, research in the introduction of such words has been conducted on two Italian dictionaries: De Mauro (2000) and Sabatini-Coletti (2008). The two dictionaries jointly report seven entries for common words of Sardinian origin, six of which have been introduced before or around 1893: *nuraghe* (1854), *barracello* (1886), *orbace* (1895), *nasco* (1895), *tanca* (1895), *canonau* (1896), while the term *launeddas* has been introduced in 1933. All seven words occur throughout the corpus: more specifically, the terms *nuraghe* and *barracello* account for more than half of the tokens for non-proper nouns registered in the indicator year 1893. It is therefore possible to interpret the high presence of Sardinian non-proper nouns in this year as marking an openness of the Italian culture towards the peculiarities of Sardinia, as testified by the introduction of most Sardinian borrowings in Italian during these years.

On the other hand, the highest occurrence of non-proper noun tokens throughout the corpus is registered in 2018. Despite the lower distribution of Sardinian loanwords in all categories in this year (4.1.), a higher frequency of non-proper nouns might suggest a shift in the semantic domain of use of Sardinian loanwords, as the reflection of a renewed interest in the Sardinian culture during the last decade.

4.3. Recent increase in material culture and social culture types

After having investigated the frequency of Māori non-proper nouns in NZE, Macalister moves to the description of their composition in terms of types for each indicator year. The same analysis has been carried out on the RIS Corpus, where non-proper nouns belong to the two categories of material culture and social culture.

The distribution of non-proper noun types all over the RIS Corpus clearly reflects the token frequencies shown in Fig. 2, with the only exception that the number of types occurring progressively also increases in 2018, following an upwards trend starting in 1910.

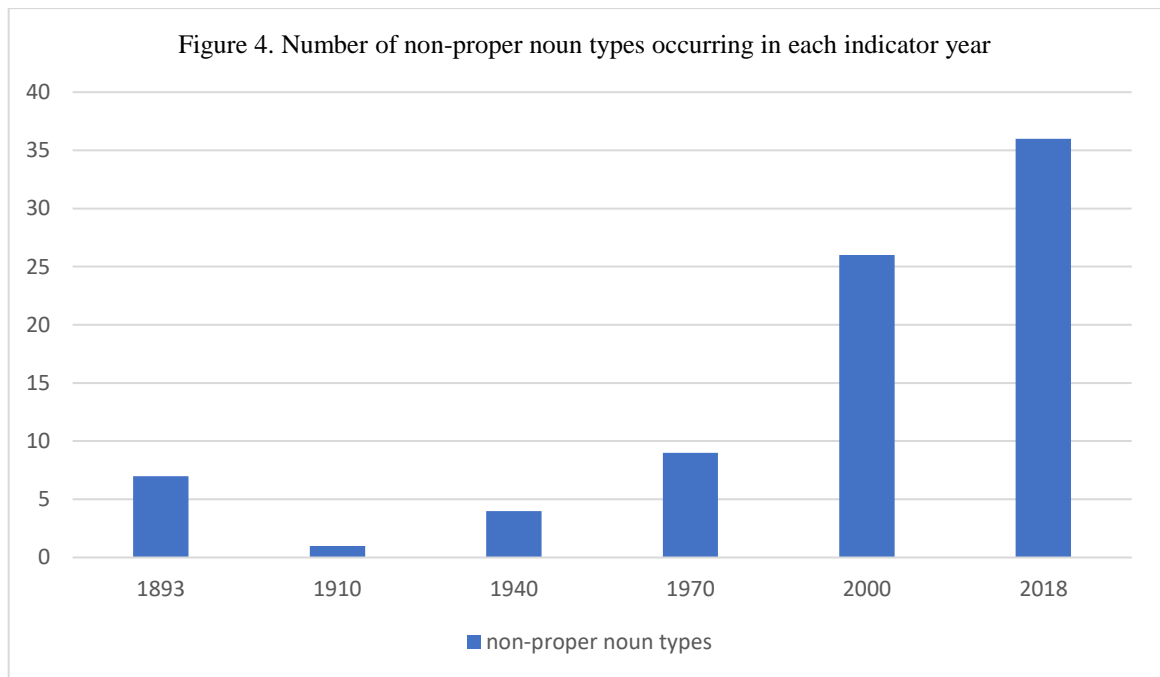Figure 4. Number of non-proper noun types occurring in each indicator year

Fig. 4. reports the number of non-proper noun types occurring in each indicator year. 42 types were identified in the whole corpus, 36 of which are represented in the indicator year 2018, the year with the highest type diversity.

In this case too, a first peak in the resulting trend is registered in 1893, meaning that the group of non-proper nouns occurring in this year was characterised by a relatively high internal diversity. Still, the lexical composition of RIS since 1970 shows higher degrees of variation concerning non-proper noun types, particularly in 2000 and in 2018, suggesting that higher frequencies of non-proper noun tokens in 1893 (seen in Fig. 1) might reflect the repetition of a more restricted set of Sardinian borrowing. This reasoning cannot but confirm a stronger influence of Sardinian in RIS, particularly in the last two generations.

4.4. Growing weight of material culture

A fourth investigation of the collected data focuses on the type of culture which is finding expression in RIS today.

In the case of NZE, the non-proper noun group has been further explored according to Macalister's categories of flora and fauna, material culture and social culture. In this case, the Sardinian non-proper noun group, composed by 42 types, has been subdivided into types referring to material culture and types referring to social culture. The representation of each category has been calculated in percentage of occurrence among the 42 types total.

The results of this analysis show that the category of social culture is less sensitive to change over time, compared to the material culture category, as can be seen in Fig. 5.

Figure 5. Percentual distribution of non-proper noun types according to category
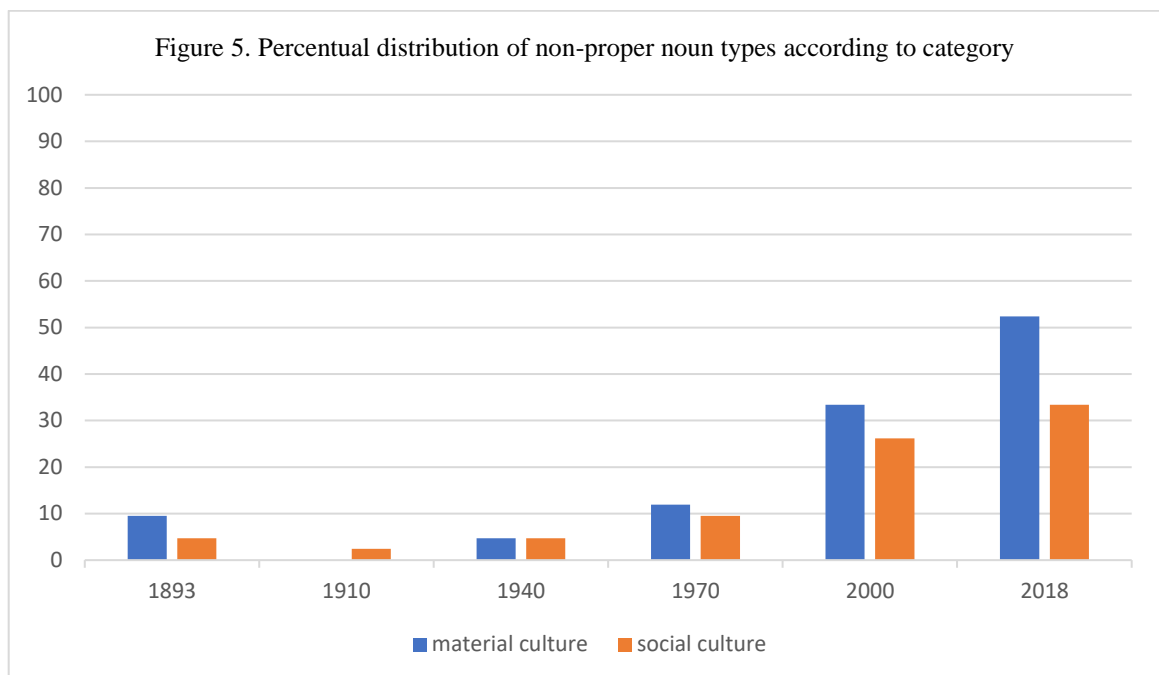
Interestingly, Figure 5 shows a lack of material culture types in the 1910 newspapers, while the representation of the category grows significantly after the 1970s. The social culture category, instead, is always represented throughout the corpus, but it does not grow much compared to the other category.

According to the results, it seems that the Sardinian material culture is progressively gaining more importance in RIS, particularly in the last decade. Indeed, if this category was overrepresented in 1910, it reached the same level of distribution of social culture in 1940, before growing exponentially until 2018.

This cannot but lead to the conclusion that what is most highly valued in the Sardinian culture on the island today are mainly concrete objects of culture. The majority of types in the material category are related to the Sardinian cuisine (words such as, *malloreddus*, *culurgiones*, *seadas*, *fregola*) and Sardinian traditional garments (*bertula*, *cambales*). Social culture types rather refer to traditional celebrations (such as *ardia*, *sartiglia*) or traditional music (*canto a tenores*) and traditional dance (*ballu tundu*). Regarding the social culture, what is underrepresented today of the Sardinian culture are all the rites, superstitions and ways of thinking, some of which had found expression in the earlier newspaper language: words such as *mutos*, a genre of songs in rhymes based on improvisation, or *gosos*, ancient chants imported from Spain, which narrate

the story of saints and religious figures. Other words related to the Sardinian social culture have not even been recorded in the whole corpus.

This might suggest that the Sardinian culture has been emptying, particularly in the last decades, since the preservation of culture-specific objects is only partially paralleled by the diffusion of social culture concepts. More specifically, in this respect, the relationship between material culture types and social culture types in 2018 shows the highest level of disproportion between the two categories in favour of the former. Future research might investigate on the future developments of this trend.

5. Conclusion

By taking a glance at the results of the investigations conducted in both corpora, it is possible to identify some common patterns of behaviour shared by NZE and RIS. Indeed, the two languages seem to change in the same direction, under a growing influence of the autochthonous ones, Māori and Sardinian. Both majority languages draw new lexical content from the co-territorial minority languages in order to express a set of cultural elements characterising the society in which their speakers live.

As has been demonstrated, words of Māori and Sardinian origin have been borrowed in NZE and RIS at an increasing rate, particularly from the 1970s. A first look at the minority language loanwords appearing in both corpora shows an overwhelming presence of proper names, which is a natural consequence of the mainly referential function of newspaper communication. Still, common Māori and Sardinian words have recently found more space in newspaper articles. Furthermore, their presence in both corpora is not the result of the repetition of a restricted group of loanwords but can be detected in a growing range of word-types which, interestingly, has reached the highest degree of internal diversity during the same years in which research on the two corpora has been carried out.

More interestingly, a classification of the non-proper nouns which have recently entered the corpora leads to different assessments of the state of affairs affecting NZE and RIS today. Indeed, if both categories of material culture and social culture have recently been more represented in the two corpora, in the case of NZE, the number of borrowings related to social culture has drastically increased, while the RIS Corpus shows a recent growth in the number of borrowings related to material culture. This means that NZE has recently been borrowing more lexical contents regarding the Māori perception of the world, life and kinship relations, while RIS leaves more space to the expression of concrete lexical contents regarding Sardinian

traditional objects and products. This is the most important difference defining the nature of the relationship between NZE and Māori, and RIS and Sardinian today.

Macalister (2006a) does not provide an explanation for the growing relevance of social culture borrowings in his corpus. The scholar rather suggests some historical and political factors which might have supported the general phenomenon of introducing Māori words into the NZE lexis. Still, for the purposes of this study, the growing relevance of Māori social culture in NZE assumes a greater value in the light of language endangerment. This can be interpreted as the reflection of a renewed interest in core elements of Māori culture, which is spreading within the whole New Zealand population. More specifically, Macalister's findings regarding a shift in the representation of lexical categories, from a predominant focus on material elements until 1970 to more abstract aspects of the Māori culture in 2000, might suggest a recently deeper digging into the essence of Māori culture, from concrete manifestations to their social value.

On the other hand, the permeation of Sardinian borrowings related to the material culture in RIS suggests a recent tendency to place more value on physical objects proper to the Sardinian culture. Indeed, as has been seen, the borrowings which have recently been used in the corpus are mainly related to traditional Sardinian costumes and local food. An interpretation of this result could be that, in this case, an attachment to the concrete manifestations of culture is not leaving space to the expression of the values and concepts lying in the material culture of Sardinia. In other words, only the surface of Sardinian culture seems to be finding expression in RIS.

This difference in the realisation of each autochthonous culture in the respective majority language variety acquires even more meaning when seen under the markedly anthropological framework of folklore theory. On this point, Gramsci's (1950) reflections on the concept of folklore can be considered as the basis for an analysis of the attitude of each society towards its autochthonous culture and the expression of this sentiment in their majority language. Gramsci considers folklore as a conceptualisation of the world shared by a society, comprising religious beliefs, moral values, the understanding of natural right, together with its concrete manifestations in customs and practices. According to this view of folklore, the Sardinian scholar denounced the erudite approach used by many schools of thought of his time in the study of folklore, based on the pure collection and classification of data regarding the concrete expressions of folklore in a community. Gramsci's approach to the study of folklore is rather unveiled in his description of the value of folksongs: "what characterises folksongs […] is neither the artistic fact, nor the historical origin, but their way of conceiving the world and life

[…]" (Gramsci 1950: 220, my translation)[14]. Adopting this theoretical framework, the different expression of autochthonous content in RIS and NZE can be interpreted as a marker of different attitudes to local culture in each society. Therefore, a conclusion might be drawn that if New Zealanders have recently been rediscovering the Māori culture by contemplating Māori ways of living and thinking, as seen also in Lombardo (1995), Sardinians are rather increasingly valuing aesthetic manifestations of local culture, from an erudite perspective, putting their social meaning in the background. More precisely, it seems that the Sardinian population is now clinging on what remains of a local folklore, which has lost the link with the conceptions of the world typical of the Sardinian culture.

On the other hand, seen from the perspective of majority-minority languages interaction, the findings of the analysis of the RIS Corpus might rather suggest an orientalisation of the Sardinian culture in Italian, to borrow Said's (1978) theory. As the scholar explains, the term Orientalism has three interrelated meanings: it can be referred to the ways in which the Orient is researched, described and taught in the academic field, or it can be associated with general ways of thinking about the Orient in the western ideology, or, it can also relate to the occidental ways of dealing with the Orient from a more political and institutional point of view. According to the scholar, all these factors have contributed in creating a constructed image of the East based on the contraposition with the West, inevitably leading to a widespread misrepresentation of the Orient which is found in "theories, epics, novels, social descriptions, and political accounts concerning the Orient, its people, customs, 'mind', destiny, and so on" (Said 1978: 3). The scholar is rather concerned with the study of Orientalism in the context of colonial and post-colonial France, Great Britain and America. Still, the results of this analysis support an application of Orientalism to the Sardinian context as well. Indeed, the fragmentary and unbalanced representation of Sardinian cultural elements in the RIS Corpus might suggest a distorted representation of the Sardinian autochthonous culture in Italian. Most importantly, this assumption must not be widened to include the whole Italian territory, since the RIS Corpus does not provide any information on this issue, but it can be associated to a fragmentary interpretation of Sardinian culture within the same Sardinian population. Indeed, the greater value placed on material culture elements suggests a conception of the Sardinian culture eradicated from its deeper values and meanings, therefore leading to a distortion of its representation in the research field, in general ideology, and on political ground.

---

[14] Original text: "ciò che contraddistingue il canto popolare […] non è il fatto artistico, né l'origine storica, ma il suo modo di concepire il mondo e la vita […]".

Nevertheless, a greater valorisation of social culture in NZE might not necessarily mean that the Māori culture has been better preserved than the Sardinian one. Indeed, among the many scholars who have turned their attention to the Māori question in New Zealand, some voices seem to be in contrast with the widespread revitalisation of Māori culture. Among them, Hanson (1989) presents the New Zealand context from a perspective which can easily be included in the orientalist framework. More precisely, the scholar talks about an 'invention' of the Māori culture and acknowledges the fact that "anthropological interpretations and misinterpretations have joined the contributions of other scholars, government officials, and Māoris themselves (including some Māori anthropologists) in the inventive process" (Hanson 1989: 890). In Hanson's account, the construction of Māori culture took place in two different phases according to different historical periods. According to the scholar, during a first phase, starting around the turn of the XX century, a Māori image was created according to "the political desire to assimilate Māoris to Pakeha culture" (Hanson 1989: 897), a statement which surely shares much with Said's (1978) Orientalism. A second phase of the invention of Māori culture is dated in the 1980s and is supposed to be caused by an attempt to place more political power in Māori's hands. According to Hanson, during this period, an image of Māori culture was created by identifying Māori qualities to be considered in contraposition with less positive aspects of *Pakeha* culture. More interestingly, the two phases described by Hanson seem the match the two waves of borrowings proposed in Macalister's (2006a) model: indeed, the first phase of Māori cultural invention corresponds to the first wave of Māori borrowings in NZE, while to the second phase corresponds a second wave of borrowing.

Waves of borrowing can also be identified in the RIS Corpus, with a first wave starting after the unification of Italy, given the date of the first use of Sardinian loanwords in the Italian, while a second wave seems to be the result of the linguistic activism of the 1970s and of the consequent language policies during the 1990s. In this case, the period of stabilisation of borrowing from the turn of the XX century until the 1970s might coincide with what Piredda (2016) refers to as the second wave of Italianisation of Sardinia.

That said, what results from a last analysis of NZE and RIS is that the two languages share similar patterns of variation over time and are progressively giving expression to the autochthonous cultures of New Zealand and Sardinia. The lexical presence of Māori and Sardinian in the two majority languages supports an interpretation regarding the value which has recently been placed on material and social culture in the two contexts, while at the same time leaving unanswered a variety of questions about the fair representation of minority cultures

in both settings. For this reason, the collection of more data from different sources is considered to be the first step needed to obtain a more systematic understanding of the matter.

Conclusions can be drawn also on a more practical level. Indeed, the difference in representation of autochthonous cultures in RIS and NZE can at this point be interpreted as a demonstration of UNESCO's (2003) statement on the different degree of endangerment of Māori and Sardinian, which considers the former as vulnerable and the latter as definitely endangered. Might there exist a one to one relationship between the representation of the two endangered languages in NZE and RIS and their degree of endangerment? This author proposes that the relation is true in this case. Nevertheless, more sources and different types of investigation are needed to make a statement on the condition of the two endangered languages today. Still, what is clear from the results obtained is that, despite their endangerment, Māori and Sardinian still live within NZE and RIS.

On a last analysis, it is beyond doubt that more data, and particularly more type of sources, are needed in order to create a complete picture of RIS today and to study the representation of Sardinian and the Sardinian culture. Despite the limits of the collected data, the corpus offers a number of ideas for consideration. Among them, one of the most interesting examples is the presence of phenomena of Sardinian-English interaction such as eat&buffas (eat&drink) and ConCheHipHop (Hip Hop head). Another topic that deserves further research is the use of Sardinian in politics, such as in the name of political parties (Fortza Paris, Sardigna Libera or Sardigna Natzione) or Sardinian-sounding names of organisations (Lavoras, Forestas). Probably, future research will shed some light on this and other issues surrounding RIS, an Italian variety, which has recently received more academic attention but is still almost a new ground in systematic research. At the same time, these cues can serve as the starting point for a more thorough understanding of the attitudes of the Sardinian society towards its autochthonous culture.

References

Barnard, R., (2007) "Language-in-Education in New Zealand: Policies and Practices", in Siegel, J., Lynch, J. and Eades, D., (Eds) *Language Description, History and Development: Linguistic Indulgence in Memory of Terry Crowley*, Amsterdam: John Benjamins Publishing, 407-418.

Bauer, L., (2002) *An Introduction to International Varieties of English*, Edinburgh: Edinburgh University Press.

Bell, A., (1984) "Language Style as Audience Design", *Language in Society*, 13 (2), 145-204.

Bell, A., (1991) *The Language of New Media*, Oxford: Blackwell Publishers.

Bell, A. and Holmes, J. (1990) *New Zealand Ways of Speaking English*, Clevedon: Multilingual Matters.

Bell, A. and Kuiper, K., (2000) *New Zealand English*, Amsterdam: John Benjamins Publishing.

Benvenuti, G., (2015) "Raccontar fole: La Sardegna esotica dei viaggiatori tra Sette e Ottocento riletta da Sergio Atzeni", in Camilotti, S., Crotti, I. and Ricorda, R., (Eds) *Leggere la lontananza: Immagini dell'altro nella letteratura di viaggio della contemporaneità*, Venezia: Edizioni Ca' Foscari, 119-134.

Blackwood, R. and Tufi, S., (2012) "Policies vs Non-Policies: Analysing Regional Languages and the National Standard in the Linguistic Landscape of French and Italian Mediterranean Cities", in Gorter, D., Marten, F.H. and Van Mensel, L., (Eds) *Minority Languages in the Linguistic Landscape*, New York: Palgrave Macmillan, 109-126.

Blasco Ferrer, E., (1986) *La lingua sarda contemporanea: grammatica del logudorese e del campidanese*, Cagliari: Edizioni Della Torre.

Cavalli-Sforza, L., (1996) *Geni, popoli e lingue*, Milano: Adelphi.

Cecaro, R., (2015) *I Giornali sardi dell'Ottocento*, Cagliari: Regione Autonoma della Sardegna.

Cerruti, M., (2009) *Strutture dell'Italiano regionale: morfosintassi di una varietà diatopica in prospettiva sociolinguistica*, Frankfurt am Main: Peter Lang.

Colella, G. and Blasco Ferrer, E., (2016) "Per lo studio del parlato giovanile in Sardegna", in Bardel, C. and De Meo, A., (Eds) *Parler les langues romanes/ Parlare le lingue romanze/Hablar las lenguas romances/Falando línguas românicas*, Napoli: Il Torcoliere, 161-179.

Conboy, M., (2010) *The Language of Newspapers: Socio-historical Perspectives*, London: Continuum International Publishing Group.

Cormack, D. and Robson, C., (2010) *Ethnicity, national identity and 'New Zealanders': considerations for monitoring Māori health and ethnic inequalities*, Wellington: Otago University Press.

Cossu, C., (2018) "La Nuova Sardegna cambia, fedele alla sua storia", *La Nuova Sardegna*, [online],
<http://www.lanuovasardegna.it/tempo-libero/2018/04/29/news/la-nuova-sardegna-cambia-fedele-alla-sua-storia-1.16774178> (18/01/2019).

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

Cunningham, U. and King, J., (2018) "Language, Ethnicity, and Belonging for the Children of Migrants in New Zealand", *International Journal of Bilingualism*, 8 (2), [online], <https://journals.sagepub.com/doi/abs/10.1177/2158244018782571> (20/12/2018).

De Mauro, T., (2000) *Il dizionario della lingua italiana per il terzo millennio*, Torino: Paravia.

Dettori, A., (2008) "Lingua sarda in movimento: dal parlato all'uso letterario", *La Linguistique*, 44 (1), 57-72.

Dettori, A., (2014) *Dalla Sardegna all'Europa: Lingue e Letterature regionali*, Milano: FrancoAngeli.

Deverson, T., (2000) "Handling New Zealand English lexis", in Bell, A. and Kuiper, K., (Eds) *New Zealand English*, Amsterdam: John Benjamins Publishing Company.

Institut de Sociolingüística Catalana (1995) "Sardinian Language Use Survey", *Euromosaic*, [online], <https://www.uoc.edu/euromosaic/web/document/sard/an/e1/e1.html> (18/01/2019).

Farinelli, M.A., (2009) "Il fascismo ad Alghero: Italianizzazione alla periferia del regime", *Insula. Quaderno di cultura sarda*, 6, 67-92.

Fishman, J.A., (1991) *Reversing language shift: Theoretical and empirical foundations of assistance to threatened languages*, Clevedon: Multilingual Matters.

Fishman, J.A., Gertner, M.H., Lowly, E.G. and Milan, W.G., (1985) *The rise and fall of the ethnic revival: Perspectives on language and ethnicity*, Berlin: Mouton de Gruyter.

Gordon, E., Campbell, L., Hay, J., Maclagan, M., Sudbury, A. and Trudgill, P. (2004) *New Zealand English: Its Origins and Evolution*, Cambridge: Cambridge University Press.

Gramsci, A., (1950) *Letteratura e vita nazionale*, Torino: Einaudi.

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

Hanson, A., (1989) "The Making of the Maori: Culture Invention and Its Logic", *American Anthropologist*, 91 (4), 890 – 902.

Harlow, R., (2005) "Maori: Introduction", in Bell, A., Harlow, R. and Starks, D., (Eds) *Languages of New Zealand*, Wellington: Victoria University Press, 59-66.

Jenkins, J., (2015) *World Englishes: A Resource Book for Students*, London: Routledge.

Lindquist, H., (2009) *Corpus Linguistics and the Description of English*, Edinburgh: Edinburgh University Press.

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

Lombardo, A. (1995) *Verso gli antipodi. Le nuove letterature di lingua inglese: India, Australia, Nuova Zelanda*, Roma: La Nuova Italia Scientifica.

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

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

Macalister, J., (2006a) "The Maori presence in the New Zealand English lexicon, 1850 – 2000", *English World-Wide*, 27 (1), 1-24.

Macalister, J., (2006b) "The Maori lexical presence in New Zealand English: Constructing a corpus for diachronic change", *Corpora*, 1 (1), 85-98.

Mair, C., (2009) "Corpus linguistics meets sociolinguistics: The role of corpus evidence in the study of sociolinguistic variation and change", *Language and Computers*, 69 (1), (. 7-32.

Mair, C. and Hundt, M., (1999) ""Agile" and "Uptight" Genres: The Corpus-based Approach to Language Change in Progress", *International Journal of Corpus Linguistics*, 4 (2), 221-242.

Moseley, C., (2010) *Atlas of the World's Languages in Danger*, Paris: UNESCO Publishing.

Mullaney, J., (2010) *Postcolonial Literatures in Context*, London: Continuum International Publishing Group.

New Zealand Government (2013) "2013 Census", *NZ Stats table viewer*, [online], <nzdotstat.stats.govt.nz/wbos/Index.aspx?DataSetCode=TABLECODE8052&_ga=2.1218 81353.304506557.1542740873-1194352557.1542138608#> (02/12/2018).

Oppo, A., (2007) *Le lingue dei sardi: Una ricerca sociolinguistica*, Cagliari: Regione Autonoma della Sardegna.

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

Pallottino, M., (2000) *La Sardegna nuragica*, Nuoro: Ilisso.

Piredda, N., (2016) "L'italiano in Sardegna", *Italienisch: Zeitschrift für italienische Sprache und Literatur*, 38 (75), 85 - 97.

Piredda, N., (2017) "L'Italiano regionale di Sardegna", in Blasco Ferrer, E., Koch, P., and Marzo, D., (Eds), *Manuale di linguistica sarda*, Berlin: De Gruyter, 495-507.

Pittau, M., (2018) *Luoghi e Toponimi della Sardegna*, Sassari: Ipazia Books.

Podda, C., (2014) "Sistemi Informativi Geografici per la conservazione e valorizzazione della toponomastica della Sardegna", *Atti della 18a conferenza nazionale ASITA*, Firenze 14-16 ottobre 2014, [online], <http://atti.asita.it/ASITA2014/Pdf/193.pdf> (25/01/2019).

Riffe, D., Aust, C.F. and Lacy, S.R., (1993) "The Effectiveness of Random, Consecutive Day and Constructed Week Sampling in Newspaper Content Analysis", *Journalism Quarterly*, 70 (1), 133-139.

Rindler Schjerve, R., (2011) "Sardinian: Italian", *Trends in Romance Linguistics and Philologhy*, 5, 271-294.

Royal Society of New Zealand (2013) "Languages in Aotearoa, New Zealand", *Royal Society Te Apārangi*, [online],
<https://royalsociety.org.nz/assets/Uploads/Languages-in-Aotearoa-New-Zealand.pdf> (10/12/2018).

Sabatini, F. and Coletti, V., (2008) *Il Sabatini – Coletti. Dizionario della Lingua Italiana*, Firenze: Sansoni.

Said, E.W., (1978) *Orientalism*, New York: Vintage Books.

Smith, P.M., (2012) *A Concise History of New Zealand*, Melbourne: Cambridge University Press.

Spolsky, B., (2003) "Reassessing Māori Regeneration", *Language in Society*, 32 (4), 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), 333-334.

Te Puni Kōkiri (2008) *The Health of the Maori Language in 2006*, [online],
<https://www.tpk.govt.nz/documents/download/67/tpk-2008-health-maori-language-2006.pdf> (15/12/2018).

Trudgill, P., (2006) *New-Dialect formation: The Inevitability of Colonial Englishes*, Edinburgh: Edinburgh University Press.

Tufi, S., (2015) "Insularity in the Linguistic Landscapes of Sicily, Sardinia, and Corsica", in Blackwood, R.J. and Tufi, S., (Eds) *The Linguistic Landscape of the Mediterranean: French and Italian Coastal Cities*, Liverpool: Palgrave Macmillan, 104-148.

UNESCO Ad Hoc Expert Group on Endangered Languages (2003) "Language Vitality and Endangerment", in *International Expert Meeting on the UNSECO Programme Safeguarding of Endangered Languages*, Paris 10 – 12 March 2003, [online],
<http://unesdoc.unesco.org/images/0018/001836/183699E.pdf> (06/12/2018).

Urban, M.B., (2013) *Sardinia on Screen: The Construction of the Sardinian Character in Italian Cinema*, Amsterdam: Rodopi.

Vacca, A., (2017) *Rights to Use Minority Languages in the Public Administration and Public Institutions: Italy, Spain and the UK*, Torino: G. Giappichelli Editore.

Vona, G., (1997) "The peopling of Sardinia (Italy): History and effects", *International Journal of Anthropology and Ethnology*, 12 (1), 71-87.

Waitangi Tribunal 2001, (2016) *Rekohu: A Report on Moriori and Ngati Mutunga Claims in the Chatham Islands*. [online], <https://forms.justice.govt.nz/search/Documents/WT/wt_DOC_68595363/Rekohu%20Rep ort%202016%20Re/print.pdf> (17/12/2018).

Warren, P. (2012) "Origins and development of New Zealand English", in *International Journal of Language, Translation and Intercultural Communication*, 1, 87-102.

Wells, N., (2018) "State recognition for 'contested languages': a comparative study of Sardinian and Asturian, 1992–2010", in *Language Policy*, 17, 1-25.