

Language Variation – European Perspectives VIII

Edited by

Hans Van de Velde

Nanna Haug Hilton

Remco Knooihuizen

John Benjamins Publishing Company

Studies in Language Variation

25

Language Variation – European Perspectives VIII

Studies in Language Variation (SILV)

ISSN 1872-9592

The series aims to include empirical studies of linguistic variation as well as its description, explanation and interpretation in structural, social and cognitive terms. The series will cover any relevant subdiscipline: sociolinguistics, contact linguistics, dialectology, historical linguistics, anthropology/anthropological linguistics. The emphasis will be on linguistic aspects and on the interaction between linguistic and extralinguistic aspects — not on extralinguistic aspects (including language ideology, policy etc.) as such.

For an overview of all books published in this series, please see benjamins.com/catalog/silv

Editors

Peter Auer
Universität Freiburg

Frans Hinskens
Meertens Instituut &
Radboud Universiteit Nijmegen

Paul Kerswill
University of York

Editorial Board

Arto Anttila
Stanford University

Johannes Kabatek
Universität Zürich

Gaetano Berruto
Università di Torino

Pieter Muysken
Radboud University Nijmegen

Jenny Cheshire
University of London

Pia Quist
University of Copenhagen

Katie Drager
University of Hawai'i at Mānoa

Anne-Catherine Simon
Université catholique de Louvain

Katarzyna Dziubalska-Kołaczyk
Adam Mickiewicz University in Poznań

Sali A. Tagliamonte
University of Toronto

Jürg Fleischer
Philipps-Universität Marburg

Øystein Alexander Vangsnes
UiT The Arctic University of Norway

Peter Gilles
University of Luxembourg

Juan A. Villena Ponsoda
Universidad de Málaga

Brian D. Joseph
The Ohio State University

Volume 25

Language Variation – European Perspectives VIII

Selected papers from the Tenth International Conference on Language Variation in Europe (ICLaVE 10), Leeuwarden, June 2019

Edited by Hans Van de Velde, Nanna Haug Hilton and Remco Knooihuizen

Language Variation – European Perspectives VIII

Selected papers from the Tenth International
Conference on Language Variation in Europe
(ICLaVE 10), Leeuwarden, June 2019

Edited by

Hans Van de Velde

Fryske Akademy & Utrecht University

Nanna Haug Hilton

University of Groningen

Remco Knooihuizen

University of Groningen

John Benjamins Publishing Company

Amsterdam / Philadelphia



The paper used in this publication meets the minimum requirements of the American National Standard for Information Sciences – Permanence of Paper for Printed Library Materials, ANSI Z39.48-1984.

DOI 10.1075/silv.25

Cataloging-in-Publication Data available from Library of Congress:
LCCN 2021010868 (PRINT) / 2021010869 (E-BOOK)

ISBN 978 90 272 0885 9 (HB)

ISBN 978 90 272 5982 0 (E-BOOK)

© 2021 – John Benjamins B.V.

No part of this book may be reproduced in any form, by print, photoprint, microfilm, or any other means, without written permission from the publisher.

John Benjamins Publishing Company · <https://benjamins.com>

Table of contents

Introduction	1
<i>Remco Knooihuizen, Nanna Haug Hilton and Hans Van de Velde</i>	
CHAPTER 1	
The volatile linguistic shape of ‘Town Frisian’/‘Town Hollandic’	11
<i>Arjen P. Versloot</i>	
CHAPTER 2	
Is there an interlanguage speech acceptability deficit?	35
<i>Rias van den Doel and Adriaan Walpot</i>	
CHAPTER 3	
Revisiting the vowel mergers of East Anglia: Correlations of MOAN, MOWN and GOOSE	53
<i>Kerri-Ann Butcher</i>	
CHAPTER 4	
Modeling regional variation in voice onset time of Jutlandic varieties of Danish	79
<i>Rasmus Puggaard</i>	
CHAPTER 5	
“Organically German”? Changing ideologies of national belonging	111
<i>Janet M. Fuller</i>	
CHAPTER 6	
Exploring an approach for modelling lectal coherence	135
<i>Karen V. Beaman</i>	
CHAPTER 7	
“I’m dead posh in school”: Attitudes and linguistic behaviour of Merseyside adolescents	161
<i>Rachel Byrne</i>	
CHAPTER 8	
<i>Benim</i> : A new pronoun in Swedish	181
<i>Nathan J. Young</i>	

CHAPTER 9

- Identification of clusters of lexical areas using geographical factors:
A case study in the Occitan language area **209**
Clément Chagnaud, Guylaine Brun-Trigaud and Philippe Garat

CHAPTER 10

- (Il)lteracy and language change: Non-standard relative constructions
in historical Basque **227**
Dorota Krajewska and Eneko Zuloaga

CHAPTER 11

- Dialect contact in the vowel system of Mišótika Cappadocian **247**
Nicole Vassalou, Dimitris Papazachariou and Mark Janse

CHAPTER 12

- Leaders of language change: Macro and micro perspectives **269**
Meredith Tamminga

CHAPTER 13

- Ethnic variation in real time: Change in Australian English diphthongs **291**
James Grama, Catherine E. Travis and Simon Gonzalez

- Index **315**

(II) literacy and language change

Non-standard relative constructions in historical Basque

Dorota Krajewska and Eneko Zuloaga

University of the Basque Country UPV/EHU

In this paper we examine a corpus of sixteenth- to nineteenth-century Basque private letters and administration documents from a historical sociolinguistic point of view. Because of the diglossic situation, such texts are rare but valuable in the historical corpus of Basque. In particular, we analyse relative clauses with the pronoun *zein* ‘which’. The construction was borrowed from Romance by highly literate bilinguals, but then spread, especially through formulaic language, to less literate writers too. We focus on the development of non-standard variants of this relative clause, which are common in our corpus, but not found in printed literary texts. We argue that the sociolinguistic context was crucial in the emergence, spread and syntactic change of the relative construction.

Keywords: historical sociolinguistics, syntax, relative clauses, literacy, Basque

1. Introduction

In this paper we adopt the viewpoint of historical sociolinguistics to reflect on Basque writing, written language, and syntactic change under language contact. Our analysis is based on a corpus of sixteenth- to nineteenth-century Basque letters and administrative documents. Processes described here have parallels in other European languages, but the Basque case offers the opportunity to observe them in a particularly interesting diglossic context, where a prestigious and standardised language (French or Spanish) was the usual means of written communication, and another language (Basque) had not yet developed a standard.

We focus on relative clauses (RC), which are complex constructions, typical of higher registers, and prone to cause problems for not-so-proficient writers. The most common relativisation strategy in modern Basque and in most historical sources is prenominal, and differs from relative clauses found in the surrounding

- (4) *gizon zuhur bat-en* pare [zeiñ-ek bastitu bai-duke bere
man wise one-GEN like which-INDEF.ERG build SUB-AUX.3SG>3SG his
etxe-a *harri-a-ren* *gaiñean*]
house-DEF.ABS rock-DEF-GEN on
'like a man who built his house on the rock' (Haraneder, 1740)

Table 1. Features of Basque RCs

Feature	Prenominal	Postnominal	Appositive	<i>zein</i>
Embedded	+	+	–	–
Relative left to the head	+	–	–	–
Gapping	+	+	+	–
Subordinator	<i>-en</i>	<i>-en</i>	<i>-en</i>	<i>-en/bait-</i>

The prenominal construction, the least marked one, differs from Romance relative clauses. Broadly speaking, Romance languages employ relative pronouns (relativisers which can express semantic features) or uninflected particles usually identical to complementisers (Stark 2016). In non-standard varieties, particles are more widespread: pronouns which inflect in the standard variety may not inflect, and uninflected particles may relativise more arguments (Blanche-Benveniste 1990; Giacalone Ramat 2008; Murelli 2011; Cerruti 2016; Stark 2016). For example, in restrictive clauses Standard French uses the complementiser *que* (object relatives), *qui* (subject relatives), and invariant *dont* (possessives (5a)). With prepositions, inflected *lequel* (5a), invariant *qui* and *quoi* are used (Stark 2016: 1031). In non-restrictive clauses, *lequel* can relativise subjects and, less frequently, objects (Stark 2016: 1032). In non-standard French, *que* relativises more syntactic positions (5b) (see, among others, Guiraud 1966; Auger 1995; Stark 2016).

- (5) a. *le livre dont/duquel* *je t'ai* *parlé*
the book whose/of.the.which I you=have talked
'the book that I have talked to you about' (Stark 2016: 1030) (French)
- b. *le livre que je t'ai* *parlé*
the book that I you=have talked
'the book that I have talked to you about'
(Stark 2016: 1030) (Colloquial French)

The syntax of the Basque *zein* RC (4) resembles the Romance counterparts, especially structures with *lequel* in French or (*el*) *cual* in Spanish. The relativised constituent is pronominalised, usually with *zein* 'which' (*non* 'where' and *noiz* 'when' are also common, and *nor* 'who' and *zer* 'what' are attested, but scarce). The pronoun takes the case required by the syntax of the subordinate clause. Unlike in Romance, in Basque the verb in the subordinate clause is usually marked with one of the

subordinators (*-en* or *bait-*), and the subordination is doubly marked (with the subordinator and the relative pronoun).

Attested since the earliest texts, the *zein* RC was frequent until the nineteenth century: taking into account prenominal, postnominal, appositive and *zein* relatives, it accounts for 20–30% of relatives in many texts, especially in translations, catechisms, and, in general, production heavily influenced by Romance models (Krajewska 2017). In the nineteenth century the frequency of the *zein* RC dropped. Nowadays it is not employed in informal registers, but occasionally appears in formal writing or translations. This decline can be linked to a rise in linguistic awareness and linguistic purism. The construction was first explicitly rejected in Azkue's grammar (1891) as a calque from Romance. This view is repeated in many twentieth-century prescriptive dictionaries and grammars (Azkue 1969 [1905–1906]; Azkue 1923–1925; Lafitte 1991 [1944]).

3. Writing in Basque before the twentieth century

3.1 The historical sociolinguistics of Basque

Since the Middle Ages, Basque has coexisted with Latin and Romance languages in a diglossic situation (Ferguson 1959; Fishman 1972). Following Madariaga (2014: 734–35), in 1600 78% of the inhabitants of the Basque Country were Basque speakers. This proportion decreased to 69% in 1800 and to 52% in 1868. Basque has also lost large geographical areas and urban space due to industrialization processes, population movements, and repression. Basque became an official language in the Basque Autonomous Community and parts of Navarre (Spain) in 1979 and 1982, but it lacks official status in the Basque-speaking territories in France.

Despite being the language of the majority of the population, Basque was rarely used in official spheres until the late twentieth century. Since the fifteenth century, diglossia is reflected for example in legal proceedings where witnesses' testimonies are often given in Basque, even though the rest is written in Romance. Noblemen only wrote Basque for private purposes, for instance in letters sent to family members, or sometimes in literary drafts.

The standardisation of Basque, understood as a development of a common standard accepted by the speakers, began in 1964 (Amorrortu 2003; Salaburu 2018; Zuazo 2019). Before the twentieth century, there were attempts to create partial standards for some dialects. In the Northern Basque Country, a Classical Labourdin literary dialect developed in the seventeenth century through the publication of religious literature. The most important codification effort in the Southern Basque

Country was the work of Larramendi, with an “apology of Basque” (1728), the first published grammar of Basque (1729) and the first Basque dictionary (1745). The common trait of all the pre-twentieth-century codification endeavours was that their influence was limited to learned authors. For example, the impact of Larramendi is clear in the literature produced in the western and central Basque dialects, but it did not affect the writings of lay people (Urgell 2018).

3.2 Literacy and biliteracy

In the Basque Country, as elsewhere in Europe, alphabetisation was more common in urban areas, higher social strata and among men, and less common in rural areas, lower social strata, and among women (Elosegi 2019).

For the Southern Basque Country, first official registers of literacy come from the second half of the nineteenth century: in 1877 38% of inhabitants of Álava were literate, and around 54% in Bizkaia, Gipuzkoa and Navarre (Dávila, Eizagirre & Fernández 1994). In the Northern Basque Country, marriage signatures analysed by Elosegi (2019) show that in Labourd in the mid-eighteenth century 11% of women and 33% of men could sign (though there were differences between areas, with more literates in cities). Illiteracy rates were lower in France than in Spain: France achieved generalised literacy by the end of the nineteenth century, and Spain only in the mid-twentieth (Lyons 2012: 10).

The authors of texts we examine here were Basque-Romance bilinguals, though the level of knowledge of Romance varied. This brings us to the issue of biliteracy, and the way it was acquired. Schooling in Basque was limited. Before the twentieth century it only existed – to a limited extent – in the Northern Basque Country. *Petites écoles*, basic level schools, are documented there since the seventeenth century (Oyharçabal 1999), although we do not know what proportion of the population could access them. Their goal was to propagate the Catholic faith, mainly through memorising the catechism, and reading and writing were subordinate to that objective (Chartier, Julia & Compère 1976; Carter 2011). Reading was taught earlier than writing, and some people could only read, especially those who left school early to work (Grosperriin 1984; Chartier 1997). Thus, those who attended these schools might have had basic writing skills, but not necessarily enough to compose, for example, a letter, which was a specialised skill, as observed by Lyons (2014: 256). As Basque schooling was non-existent apart from the *petites écoles*, authors of texts in our corpus were most probably schooled in the dominant language (Latin or Romance), and not taught to write Basque. The outcome is that they had to exploit the knowledge about writing in one language to write in another.

4. *Zein* RCs in administrative texts and letters

4.1 Non-canonical texts in Basque historical corpus

Mounole and Lakarra (2018: 358) describe the corpus of historical Basque as *small, asymmetric* (no attestations for some varieties until after 1850) and *homogeneous* (mostly religious texts). This is an effect of diglossia: until the last decades of the twentieth century the members of higher social strata wrote almost only in Latin or Romance (Castilian, Occitan or French), depending on the period and the territory. Basque was mostly employed in writing to spread Christian ideas to monolingual people, and that explains why religious works prevail. Sarasola (1975: 109) counted 101 original Basque books, mostly about religion, between 1545 and 1900 (194 when also including translated books, and 588 when adding re-editions). Nevertheless, there are exceptions, such as oral materials transcribed by historians, private letters, and administrative documents.

For this study we have analysed all sixteenth- to nineteenth-century Basque texts available from the following genres: administrative letters and documents, notaries' texts, and private letters. Such texts can be considered "non-canonical" in the history of Basque, as opposed to "canonical" sources: printed texts, usually treating with religion. Canonical texts, extensively studied with reference to relativization strategies (e.g. in Krajewska 2017), will be used here as a basis for comparison, and the grammar used in those texts will be considered standard. The goal is to analyse the variation across textual genres, and to identify differences between canonical and non-canonical sources. This issue has not been the focus of Basque linguistics, which, because of the scarcity of texts, tends to analyse all sources together without considering the genre.

Preserved fifteenth- to eighteenth-century correspondence in Basque comprises only around 23,000 words (Padilla-Moyano 2015). Of these, 15,000 pertain to private letters. The biggest part (around 9,000 words) comes from fifty missives sent in 1757 on the ship *Le Dauphin* from Labourd to Louisbourg (Canada) (Lamikiz, Padilla-Moyano & Videgain 2015). Since many were signed by members of the lower social strata, Padilla-Moyano (2015) proposed that a significant proportion of the Labourdin society, including women, was literate. However, Elozegi (2019), after analysing church registers and notarial documents, concluded that most letters were written by intermediaries. Among the 42 senders he could identify, only ten (seven men and three women) were literate and they belonged to upper or upper-middle social classes. Delegated writing was the usual way through which illiterate or semi-literate people could access writing until literacy became universal (Lyons 2014). Lyons (2014) lists three categories of intermediaries: the professional

writer, the local notable (such as a cleric or a schoolteacher) and a family member or friend. Data from the *Le Dauphin* letters suggest that the latter two categories were involved (Padilla-Moyano 2015; Elosegi 2019).

Administrative documents were scarce before 1900. Trebiño (2001) gathered 79 texts (including letters) from between 1616 and 1935. Among them we find municipal laws, requests, or notifications. Many were translated from Romance and signed by civil servants of different ranks: mayors, secretaries, clerks, commissioners, or accountants. Even though we have little information about authors, it is unlikely that a monolingual Basque speaker could work in administration, as the language normally used in administration was Spanish or French.

Finally, Elosegi (2018) discovered Basque writings by eight notaries from Labourd (France), among whom Martin Harismendy stands out with 164 pages (11,200 words) in Basque. However, Harismendy's Basque documents represent a small proportion of his production of 13,167 pages.

Thus, in this paper we analyse texts which, because of the diglossic situation, are exceptional in the historical corpus of Basque. We have explained why the sociolinguistic context did not favour their production, but it is in order to reflect on why there were exceptions:

- a. Basque could be used for personal reasons (for instance, between two native Basque mayors who knew each other well).
- b. Basque could be a bridge language employed, for example, in correspondence between towns located on opposite sides of the border between France and Spain.
- c. The great part of the society was monolingual. Because of that the authorities considered Basque useful to spread their decisions. Courts sometimes prepared interrogatories in Basque for monolinguals.

4.2 The frequency of *zein* RCs

In this section we analyse the outcomes the sociolinguistic situation had on the Basque syntax. We gathered 227 occurrences of the *zein* RC, from 106 texts.¹ Of these 18 are administration documents (16 different authors), 13 notaries' text (3 authors), 32 administrative letters (17 authors), and 43 private letters. Table 2 presents the distribution of tokens across text types and centuries. Most examples from private correspondence (42 out of 65) come from the *Le Dauphin* collection, where *zein* relatives appear in 24 out of 50 letters.

1. More texts were analysed, but only those featuring the construction in question are included.

Table 2. Tokens of the *zein* RC, split up by genre and century

Century	Adm. doc.	Adm. letters	Notaries	Private letters	Total
16	0	0	0	9	9
17	0	33	0	3	36
18	22	10	37	48	117
19	33	27	0	5	65
Total	55	70	37	65	227

To assess the frequency of *zein* RCs we have also extracted all other headed finite relative clauses from the texts. Table 3 shows the frequency of prenominal, postnominal, appositive and *zein* relatives as well as the overall frequency of relative constructions across text types and in the whole corpus.² Additionally, we compare these figures with a sample of 17 seventeenth-to-nineteenth-century canonical Basque texts: catechisms and other types of religious prose.

Table 3. Normalised frequencies (per 1,000 words) of relative constructions

Text type	Tokens	Frequency (per 1,000 words)				
		Prenom.	Postnom.	Appos.	<i>zein</i>	All RCs
Administrative texts and letters	658	11.0	1.2	1.0	7.1	20.2
– Adm. documents	289	15.9	1.0	1.1	4.5	22.5
– Adm. letters	160	7.6	2.2	1.0	8.5	19.2
– Notaries	95	23.2	0.9	2.8	17.1	44.0
– Private letters	114	4.3	0.5	0.4	7.0	12.4
Religious texts	3,241	7.9	0.7	1.1	2.0	11.6

As expected (see Biber 1988), there are differences between text types: the highest rate of relatives per 1,000 words is found in notarial documents (44), followed by administrative documents and letters (around 20). Texts produced by notaries and clerks have more relatives than Basque religious prose has (with about 12 per 1,000 words), and private letters contain a similar number of relative clauses as religious prose.

The proportion of postnominal and appositive constructions does not vary across text types, but there are differences in the use of the *zein* RC. It accounts for 17% of all relatives in religious texts. This proportion is similar in administration documents, but higher in private and administration letters (56% and 44%, respectively) and notarial texts (39%). Consequently, the prenominal construction (the most common in modern language) is less frequent in those texts.

2. As most texts in our corpus are short (300 words on average), the frequencies were calculated per text type (and not per text).

- (8) Marka-tzen darotazu ez-tuzu-la Qanigon-ekin
 write-IPFV AUX.2SG>3SG<1SG NEG-AUX.2SG>3SG-COMPL Qanigon-SOC
 izatu letra-rik, bainan egin gintuben, [zein-etan
 have letter-PART but make AUX.PST.1PL>3SG which-INDEF.INES
 aizpa Nana Miarritz-en bai-tugu neskato Bodri-rekin]
 sister Nana Miarritz-INES SUB-have.1PL>3SG maid Bodri-SOC
 ‘You write that you haven’t got our letters from Qanigon, but we wrote one, in
 which sister Nana is maid in Biarritz with Bodri.’

(LeDauphin-38, private letter, 1757)

Non-standard variants are most common in private letters, where over 40% of the examples are of this kind (Table 4).

Table 4. Non-standard *zein* relative clauses in different text types

Text type	Tokens of <i>zein</i> RC	Non-standard variants (%)
Adm. doc.	55	20.0
Notaries	37	18.9
Adm. letters	70	7.1
Private letters	65	43.1

Table 5 presents occurrences of each variant (including the non-relative), and the number of sources in which it appears. The variant with *zeinetan* is the most widespread non-standard option.

Table 5. Frequencies of variants of *zein* relative

<i>Zein</i> RC type	Tokens	Texts
Standard	176	88
Non-standard <i>zeinetan</i>	28	22
Non-standard <i>zeina/zeinak</i>	15	8
Non-relative	8	4

The frequency of non-standard relatives seems highest in the eighteenth century, and it decreases in the nineteenth century. This might reflect an increase in literacy and knowledge of Romance languages, which would cause using the standard *zein* relative. Non-standard variants were unsuccessful innovations: they are not attested in the twentieth century.³

3. However, in modern written Basque there is a non-standard relative construction with the invariant particle *non* ‘where’, which seems not unlike the variant with *zeinetan*, but its development and use require further study.

4.4 *Zein* relatives and formulaic language

Zein relatives in our corpus are often found in epistolary formulae, which are defined by Wray (2005: 9) as “a sequence, continuous or discontinuous, of words or other elements, which is, or appears to be, prefabricated: that is, stored and retrieved whole from memory at the time of use, rather than being subject to generation or analysis by the language grammar”.

Following the classification by Wray (2005) and Rutten and Van der Wal (2012), two classes are relevant here: text-constitutive (e.g. letter openings) and intersubjective formulae (e.g. greetings or expressions of health). In a small corpus such as ours it is difficult to decide what counts as a formula. We thus only include two classes frequent in the data: (a) letter openings (“I have received your letter in which ...” and its variations, see Example (9)), and (b) the expression of the sender’s health and hope that the receiver is alright (“we have health, which we hope you also have”, “I want to let you know about my health, which is good” and similar phrases, Example (10)).

- (9) Rezebitu dut atzo zu-k eskribatu karta,
 receive AUX.1SG>3SG yesterday you-ERG write letter
 [zoin-tan erra-ten bai-tuzu jente paubri-ak
 which-INDEF.INES say-IPFV SUB-AUX.2SG>3SG people poor-DEF.ABS.PL
 dire-la zu-ri eska-tzen dute-n-ak
 be.3PL-COMPL you-DAT ask-IPFV AUX.3PL>3SG-SUB-DEF.ABS.PL
 gastu-ak]
 payment-DEF.ABS.PL
 ‘Yesterday I received the letter you wrote, in which you say that it is the poor
 people that asks you for money.’ (administration letter, 1817; Camino (2012))
- (10) Osasun dugu Janko-a-ri esker, [zein-etan dezira-tzen
 health have.1PL>3SG god-DEF-DAT thank which-INDEF.INES wish-IPFV
 bai-kinduke zuri-a hala ba-litz]
 SUB-AUX.POT.1PL>3SG your-DEF.ABS so COND-be.HYP.3SG
 ‘We have health, thank God, which we wish you also have.’
 (LeDauphin-46, private letter, 1757)

Table 6 shows that there is a correlation between the frequency of occurrence in a formula and letter type, on one hand, and type of relative, on the other hand. In private letters relatives appear more frequently in formulae than in administration letters. In *Le Dauphin* collection, this proportion is even higher (39%). Non-standard relatives are also more common as part of a formula than the standard construction.

Table 6. Relative clauses in formulae according to the text type and construction

Context		Tokens	Formulae
Letter type	Administrative	70	12.9%
	Private	65	29.2%
Relative clause type	Standard	102	14.7%
	Non-standard	33	39.4%

5. Discussion

In this paper we focus on texts written in Basque despite the adverse diglossic situation. Their authors usually belonged to upper social strata and were taught how to write in Romance. Thus, literacy in Basque depended on literacy in another language. This fact is patent in spelling: for instance, writers often struggled to render Basque sibilants using Romance systems, which have fewer sibilant phonemes. Here we explore the consequences it had on syntax.

Writing does not usually happen in void and knowing how to write is not limited to the command of the code. It also presupposes knowledge related to different aspects of writing, such as genre knowledge and knowledge related to the social context of writing (Beaufort & Iñesta 2014). For example, writing a letter requires understanding of the conventions ruling this type of texts and of how the given community usually writes them, as well as what is thought to be correct and what not. Genres can in fact be understood as “socially and culturally patterned ways of engaging in activity” (Lillis 2013: 70).

As regards genres analysed here, in terms of Koch and Oesterreicher (2012 [1985]), private letters are closer to the language of immediacy, and formal documents produced by different institutions are examples of language of distance: they constitute a monologue, there is distance between the partners, and topics are fixed. Administrative correspondence falls somewhere in between. Nevertheless, the texts in our corpus were written before the language was standardised, and as a result they exhibit dialectal features and characteristics described for the language of immediacy by Koch & Oesterreicher (2012 [1985]: 454–455): congruence errors, problems with segmentation, low type-token ratio, overall lexical poverty, hesitation phenomena or markers of correction. Moreover, features of illiterate writing (see Montgomery 1995) are common: problems with spelling and word division or missing punctuation marks. This is because those texts were written in Basque by people who were literate in Romance. In a complex sociolinguistic environment, they could not rely on a fully standardised language and had to use a dialectal variety. Nevertheless, they looked for models to construct their writings.

Bilinguals could resort to norms of other languages. For instance, when a clerk prepared a Basque document, he could translate from Spanish or French the typical phrases of administrative language, and, in general, try to copy the style of the text. The same holds for other conventions, for example the use of epistolary formulae.

Writers also treated Basque texts as models. For this to be possible, though, a certain tradition of Basque writing had to exist in the social environment. For instance, when composing a letter, the writer could reuse fragments of Basque letters, provided he had access to any. In general, learning of correspondence-writing conventions usually happened in practice: by reading or listening to letters and then imitation (see van der Wal & Rutten 2013 and references therein). This explains the similarity between texts written by different people.

The *zein* relative construction is thought to have developed in Basque due to language contact (Lafitte 1991 [1944]; Haase 1992; Trask 1998). This might be a change from above spreading in writing from more literate bilinguals (usually belonging to higher social classes) to less proficient ones. Similar explanations have been proposed for changes in relativisation strategies in a number of languages (Romaine 1982; Rissanen 1999; Rutten & van der Wal 2014, among others). In Basque, this change could happen through imitating the Romance style, for instance, through importing epistolary formulae, when writing a text belonging to a genre not-so-entrenched in Basque. We will develop this idea in what follows.

(11)–(14) are typical French eighteenth-to-twentieth-centuries letter openings with different types of RCs (with *qui*, *dont* and *lequel*). Example (14) resembles non-standard Basque relatives: apart from a resumptive pronoun *elle*, the pronoun has the non-standard form (*auquel* with a preposition instead of *laquelle*). Branca-Rosoff (1990) lists incorrect relatives among characteristics of letters of early twentieth-century soldiers.

- (11) Je répond a ta lettre que nous avons reçu il y a qu'elques jours *dont* nous avons été très contents de voir que vous êtes tous en bonne santé
(Bruneton-Governatori & Moreux 2015: 89)
- (12) Je vous aist crist cette Lettre pour vous faire asavoir de mes nousvelle *qui* sontres bonne dieu mersie
(Martineau 2007: 205)
- (13) J'ai reçu hier 20 janvier ta lettre du 17 *dans laquelle* tu me parles du dessus de piano d'Adèle et où se trouve le petit mot de Jeannette
(Vicari 2018: 14)
- (14) J'ai reçu ta longue lettre du 28 *auquel* elle m'a fait grand plaisir
(Branca-Rosoff 1990: 23)

Thus, formulae containing the *zein* RC were used by Basque writers, possibly less experienced, who were in need of prefabricated elements. Imitation was not always perfect, and often it was more like a “broken telephone” transmission, with

misunderstandings and modifications of the construction. One of the reasons for that could be the limited command of Romance languages, and particularly of formal registers.

Bilingual speakers who used primarily spoken registers of Romance languages might have had problems with relatives in these languages, because they are rare in the spoken language and because Basque and Romance relativisation strategies differ. For English, for instance, Biber (1988) found that per 1,000 words there are on average 2.9 relative clauses in conversations, 4.6 in press reports, and 8.6 in official documents. Additionally, the French and Spanish equivalents of the *zein* RC – structures with *lequel* and *el cual*, respectively – pertain to formal written registers (Guiraud 1966; Auger 1995). With limited exposition to those constructions, bilinguals might not have understood their syntax. Speakers were more likely to be exposed to Romance constructions with invariant particles, which also differ from Basque usual relatives, but are simpler to use, and more common in the spoken language. The resulting non-standard Basque relatives are similar to *lequel/el cual* relatives because of the pronoun used, but function like Romance relatives with invariant particles.

Formulaic language played a significant role in these processes. As regards the variant with the inessive *zeinetan*, there is an epistolary formula, in which the inessive argument is relativised (15). Its equivalent is common in other languages (13), and the first step was to translate it into Basque using the standard *zein* construction:

- (15) Recebitu dut *çure guthun-a*, [çoin-etan escriba-tzen
 receive AUX.1SG>3SG your letter-DEF.ABS which-INDEF.INES write-IPFV
 bai-terautaçu nola dezir handi-a ducie-la
 SUB-AUX.2SG>3SG<1SG how wish big-DEF.ABS have.2PL>3SG-COMPL
 guci-ec baque eta arcordu hun bat-en eçar-te-ra bi
 all-DEF.ERG.PL peace and agreement good one-GEN bring-NMLZ-ADL two
 herri hoyen artian]
 community these.GEN between

‘I have received your letter in which you write me that your greatest wish is to bring peace and good agreement between the two communities.’

(Etxart, administrative letter, 1616–1617)

We think that such formulae – or even this *very* formula – were the source of the non-standard *zeinetan* RC. Writers reanalysed the inessive pronoun in a formula (15) as a particle introducing the clause and extended it to situations where another syntactic function is relativised.

This reanalysis could have happened through “recycling” and slight modifications of parts of formulae. This is represented in (16). The first stage is the formula (16a). It is then shortened: from ‘in which you write that X’ we arrive at ‘in which X’. This clause is used in a sentence containing a reference to a letter (16b). Even

though syntactically it cannot function as a RC, semantically it resembles it. We find several such examples in our corpus: the ones we labelled as “non-relative” (see (8) above). In such contexts the pronoun *zeinetan* loses its locative value and becomes a semantically empty relativiser. Afterwards the construction can be used to modify any noun phrase in any circumstances (16c), as in (17).

- (16) a. I have received your letter, [in which you write that X]
 b. You write that you haven't got any letters ... [in which X]
 c. ... [a noun phrase] [which X]

- (17) *salutatzen zaituzte aizpa Nanak, [zeinetan Miarritzen*
greet-IPFV AUX.2PL>3SG sister Nana-ERG which-INDEF.INES Biarritz-INES
bai-ta neskato]
SUB-be.3SG maid

‘sister Nana, who is a maid in Biarritz, sends greetings.’

(LeDauphin-15, private letter, 1757)

Similar scenarios can be proposed for other, less frequent variants. The option with *zeina* can be linked to the health formula, in which the absolutive form of the pronoun is found:

- (18) *Bada, ni-k ere satifa-tzen zaitut guri-a-z, [zeñ-a*
so I-ERG also satisfy-IPFV AUX.1SG>2SG OUR-DEF-INS which-DEF.ABS
bai-tugu haiñitz perfekt-a Jainko-a-ri esker]
SUB-have.1PL>3SG very perfect-DEF.ABS god-DEF-DAT thank

‘And so, I also inform you on our (health) which, thank God, we have perfect.’

(LeDauphin-49, private letter, 1757)

The changes attested in Basque can be considered another example of diaphasic variation in relativisation strategies, common in many European languages, where relative pronouns that are inflected in the standard language are invariant in non-standard varieties.

A particular case of this variation involves locative pronouns as general relativisers. In some languages such relativisers belong to the standard (e.g. *pu* ‘where’ in Modern Greek (Murelli 2011: 184)), but elsewhere they are limited to non-standard varieties. For example, *wo* ‘where’ can relativise all cases in some German varieties (Romaine 1984; Fleischer 2004). In Romance, something similar happens in non-standard Italian with *dove* ‘where’ (Cerruti 2016).

The explanation proposed here for Basque non-standard relative clauses resembles the general scenario in the development of relative particles from locative pronouns proposed by Murelli (2011: 183). According to him, the locative interrogative is first used to relativise locative arguments, but it then becomes an “unspecific connector” linking main and subordinate clauses (which is not a RC syntactically, but there is a shared participant in both clauses). From there it extends to relative

clauses, relativising first inanimate and then also animate obliques, before extending to other syntactic functions. In Basque texts we also find something similar to the unspecific connector (in “non-relative” examples), though we do not have enough information on the order of extension (oblique to grammatical cases). The data discussed here provide an insight into why such changes can happen and how they can proceed in written language. In Basque, the reason appears to be related to language contact, incomplete bilingualism and inexpert writers’ reliance on formulaic language.

Fixed phrases were also important for the development of relative constructions elsewhere, for example in the English *wh*-relative (Romaine 1982; Rissanen 1999). Bergs (2005: 50) argues that *who*, *whom* or *whose* “do not simply enter the linguistic system as grammatical elements in one big swoop, but which originate as part and parcel of complex formulaic expressions which are then slowly (des-)integrated into the grammar.” Non-standard *zein* relatives also originated as a part of fixed phrases, but eventually became a productive relativisation device (the construction occurs also outside formulaic contexts). Moreover, the non-standard *zein* eventually spread to writing of people proficient in Romance.

We could also ask why writers used the *zein* relative construction so often: as compared to Basque religious prose, this construction is more frequent in our corpus, especially in letters and notarial documents. We think that the construction, most probably not employed in informal oral interactions, became a feature associated with written language. Letters are, in general, closer to spoken discourse than many other genres, but they contain elements of more formal registers, too, and reflect people’s awareness of writing conventions (Martineau 2007; van der Wal & Rutten 2016). The use of formulae is a sign of this awareness. More generally, the spread of *zein* relative into the genres other than literary texts can be seen as an example of the creation of Basque written register. Importantly, with non-standard *zein* relatives this process seems to have happened in a “natural” way, in exchanges between writers, and not through influence of printed language. The proof of that is that those non-standard relatives are not found in printed Basque books.

6. Conclusion

The Basque *zein* relative clause has been traditionally considered an uninteresting calque from Romance, but the particularities behind this borrowing were not discussed. Having analysed a corpus of genres understudied in Basque linguistics, such as administrative documents and private letters, we have shown that in order to understand how the construction was introduced into the language it is necessary to take into account the sociolinguistic situation and characteristics of textual genres.

We have reflected on reasons for the emergence and spread of *zein* relative, the mechanisms that underlie its diffusion to different genres and from more literate to less literate writers. Finally, we have explained the creation of non-standard variants of the construction. Similar processes to those observed in Basque were described for other European languages, but our case study provides an insight into how such processes take place in the context of diglossia and lack of standard variety.

This sociolinguistic situation made people look for models in other languages (specifically, imitating Romance constructions), but also in Basque texts they had access to, such as letters. In this way, the relative construction with the pronoun *zein* ‘which’, common in literary texts, spread to administrative or legal language and private letters. In letters, epistolary formulae, partially fixed expressions usually translated from Romance, played a key role in these processes, because a few of them contained the *zein* relative. In a further development, non-standard relatives (e.g. the variant with inessive pronoun) emerged when, similarly to what occurred in other European languages, writers reanalysed the syntax of formulae with the standard *zein* relative. We propose that this innovation happened because of the limited command of Romance languages of some writers, who did not understand the syntax of the standard construction. The new relativisation strategy then became productive and spread to other writers. We argue that *zein* relative, standard and non-standard, eventually became a feature of the emerging Basque written register. The life of the non-standard relative construction, though interesting, was relatively short, as it did not make it into the twentieth century, and the reasons for it appear to be increasing knowledge of Romance languages in the population.

Abbreviations and glossing conventions

ABS	absolutive	INS	instrumental
ADL	adlative	IPFV	imperfective
AUX	auxiliary verb	NEG	negation
COMPL	complementiser	NMLZ	nominalisation
COND	conditional	PART	partitive
DAT	dative	PL	plural
DEF	definite	POT	potential
ERG	ergative	PST	past
GEN	genitive	RM	relational marker
HYP	hypothetical	SG	singular
INDEF	indefinite	SOC	sociative
INES	inessive	SUB	subordinator

In glosses of finite verbs the sign “>” distinguishes ergative and absolutive arguments and “<” distinguishes dative ones.

Funding

The research for the paper was made possible by the grants from the Spanish Ministry of Science and Innovation (FFI2016-76032-P and PGC2018-098995-B-I00). Support given by the research group on historical linguistics (IT1344-19) funded by the Basque Government is also acknowledged.

References

- Amorrortu, Estibaliz. 2003. *Basque sociolinguistics: Language, society, and culture*. Reno: Center for Basque Studies.
- Auger, Julie. 1995. On the history of relative clauses in French and some of its dialects. In Henning Andersen (ed.), *Historical linguistics, 1993: Selected papers from the 11th International Conference on Historical Linguistics*, 19–32. Amsterdam & Philadelphia: John Benjamins. <https://doi.org/10.1075/cilt.124.03aug>
- Azkue, Resurrección María. 1891. *Euskal izkindea*. Bilbao: Astui.
- Azkue, Resurrección María. 1923–1925. *Morfología vasca*. Bilbao: Editorial Vasca.
- Azkue, Resurrección María. 1969 [1905–1906]. *Diccionario vasco-español-francés*. Bilbao: La Gran Enciclopedia Vasca.
- Beaufort, Anne & Anna Iñesta. 2014. Author profiles: Awareness, competence, and skills. In Eva-Maria Jakobs & Daniel Perrin (eds.), *Handbook of writing and text production*, 141–158. Berlin: Mouton De Gruyter. <https://doi.org/10.1515/9783110220674.141>
- Bergs, Alexander. 2005. The role of the individual in language change from the point of view of Social Network Analysis. *Logos and Language* 6. 30–54.
- Biber, Douglas. 1988. *Variation across speech and writing*. Cambridge: Cambridge University Press. <https://doi.org/10.1017/CBO9780511621024>
- Blanche-Benveniste, Claire. 1990. Usages normatifs et non normatifs dans les relatives en français, en espagnol et en portugais. In Johannes Bechert, Giuliano Bernini & Claude Buridant (eds.), *Toward a typology of European languages*, 317–335. Berlin: Mouton de Gruyter. <https://doi.org/10.1515/9783110863178.317>
- Branca-Rosoff, Sonia. 1990. Conventions d'écriture dans la correspondance des soldats. *Mots* 24(1). 21–36. <https://doi.org/10.3406/mots.1990.1535>
- Bruneton-Governatori, Ariane & Bernard Moreux. 2015. Un modèle épistolaire populaire: Les lettres d'émigrés béarnais. In Daniel Fabre (ed.), *Par écrit: Ethnologie des écritures quotidiennes*, 79–103. Paris: Éditions de la Maison des sciences de l'homme.
- Camino, Iñaki. 2012. Bi testu llabur Aezkoa eta Zuberoako. Garralda (1828) & Larraine (1817). *Fontes Linguae Vasconum* 114. 61–72.
- Carter, Karen E. 2011. *Creating Catholics: catechism and primary education in early modern France*. Notre Dame: University of Notre Dame Press.
- Cerruti, Massimo. 2016. Changes from below, changes from above. Relative constructions in contemporary Italian. In Massimo Cerruti, Claudia Crocco & Stefania Marzo (eds.), *Towards a new standard: Theoretical and empirical studies on the restandardization of Italian*, 62–88. Berlin: Mouton de Gruyter.
- Chartier, Roger. 1997. Du livre au lire. *Réseaux* 1. 271–290.
- Chartier, Roger, Dominique Julia & Marie-Madeleine Compère. 1976. *L'éducation en France du XVIe au XVIIIe siècle*. Paris: Société d'Éducation d'Enseignement Supérieur.

- Dávila, Pauli, Ana Eizagirre & Idoia Fernández. 1994. Los procesos de alfabetización y escolarización en Euskal Herria. *Cuadernos de Sección. Educación* 7. 63–99.
- Elosegi, Xabier. 2018. *Senpereko eta Lapurdiko euskara XVIII mendean*. Senpere: Lapurdi 1609.
- Elosegi, Xabier. 2019. *Le Dauphin: 1757ko gutuneriari buruzko osagarriak eta gogoetak*. Bilbao: Euskaltzaindia.
- Ferguson, Charles A. 1959. Diglossia. *Word* 15(2). 325–340.
<https://doi.org/10.1080/00437956.1959.11659702>
- Fishman, Joshua A. 1972. *The sociology of language*. Rowley: Newbury House.
- Fleischer, Jürg. 2004. A typology of relative clauses in German dialects. In Bernd Kortmann (ed.), *Dialectology meets typology*, 211–243. Berlin: Mouton de Gruyter.
- Giacalone Ramat, Anna. 2008. Persistence and renewal in the relative pronoun paradigm: the case of Italian. *Folia Linguistica Historica* 25(1–2). 115–138. <https://doi.org/10.1515/flin.26.1-2.115>
- Grosperin, Bernard. 1984. *Les petites écoles sous l'Ancien Régime*. Rennes: Ouest France.
- Guiraud, Pierre. 1966. Le système du relatif en français populaire. *Langages* 3. 40–48.
<https://doi.org/10.3406/lgqe.1966.2342>
- Haase, Martin. 1992. *Sprachkontakt und Sprachwandel im Baskenland*. Hamburg: Helmut Buske.
- Koch, Peter & Wulf Oesterreicher. 2012 [1985]. Language of immediacy – language of distance: Orality and literacy from the perspective of language theory and linguistic history. In Claudia Lange, Beatrix Weber & Göran Wolf (eds.), *Communicative spaces. Variation, contact, and change*, 441–473. Frankfurt am Main: Peter Lang.
- Krajewska, Dorota. 2017. *Euskararen sintaxi diakronikorantz: Egitura konplexuak / Towards a historical syntax of Basque: Complex constructions*. Vitoria-Gasteiz: UPV/EHU doctoral dissertation. <http://hdl.handle.net/10810/26877>
- Lafitte, Pierre. 1991 [1944]. *Grammaire basque*. Donostia: Elkar.
- Lamikiz, Xabier, Manuel Padilla-Moyano & Charles Videgain. 2015. *Othoi çato etchera*. Le Dauphin itsasontziko euskarazko gutunak (1757) – Correspondance basque du corsaire Le Dauphin (1757). *Special edition of Lapurdum* 2.
- Larramendi, Manuel. 1728. *De la antigüedad y universalidad del bascuense en España*. Salamanca: Eugenio García de Honorato.
- Larramendi, Manuel. 1729. *El imposible vencido*. Salamanca: Antonio Joseph Villargordo Alcaraz.
- Larramendi, Manuel. 1745. *Diccionario trilingüe del castellano, vascuence y latín*. San Sebastián: Bartolomé Riesgo y Montero.
- Lillis, Theresa. 2013. *The sociolinguistics of writing*. Edinburgh: Edinburgh University Press.
- Lyons, Martyn. 2012. *The writing culture of ordinary people in Europe, c. 1860–1920*. Cambridge: Cambridge University Press. <https://doi.org/10.1017/CBO9781139093538>
- Lyons, Martyn. 2014. The power of the scribe: Delegated writing in Modern Europe. *European History Quarterly* 44(2). 244–262. <https://doi.org/10.1177/0265691414526342>
- Madariaga, Juan. 2014. *Sociedad y lengua vasca en los siglos XVII y XVIII*. Bilbao: Euskaltzaindia.
- Martineau, France. 2007. Variation in Canadian French usage from the 18th to the 19th century. *Multilingua* 26(2–3). 203–227. <https://doi.org/10.1515/MULTI.2007.010>
- Montgomery, Michael. 1995. The linguistic value of Ulster emigrant letters. *Ulster Folklife* 41. 26–41.
- Mounole, Céline & Joseba A. Lakarra. 2018. Euskara Arkaikoa. In Joakin Gorrotxategi, Iván Igartua & Joseba A. Lakarra (eds.), *Euskararen historia*, 345–468. Vitoria-Gasteiz: Eusko Jaurlaritza.
- Murelli, Adriano. 2011. *Relative constructions in European non-standard varieties*. Berlin: Walter de Gruyter. <https://doi.org/10.1515/9783110238792>
- Oyharçabal, Bernard. 1987. *Étude descriptive de constructions complexes en basque, propositions relatives, temporelles, conditionnelles et consessives*. Paris: Université de Paris VII doctoral dissertation.

- Oyharçabal, Bernard. 1999. Euskarazko irakaskintzaren historia ororen eskolen ildotik, iraultzaren frantses garaiko eskola liburuxka bat. *Lapurdum* 4. 81–105.
<https://doi.org/10.4000/lapurdum.1545>
- Oyharçabal, Bernard. 2003. Relatives. In José Ignacio Hualde & Jon Ortiz de Urbina (eds.), *A grammar of Basque*, 762–821. Berlin: Walter de Gruyter.
- Padilla-Moyano, Manuel. 2015. Le Dauphin itsasontziko gutuneria (1757) edo euskararen historia behetik. *Special edition of Lapurdum* 2. 45–77.
- Rijk, Rudolf Pieter Gerardus de. 1972. *Studies in Basque syntax: Relative clauses*. Cambridge, Massachusetts: Massachusetts Institute of Technology.
- Rissanen, Matti. 1999. Syntax. In Richard M. Hogg & Roger Lass (eds.), *The Cambridge history of the English language*, vol. 3, 187–331. Cambridge: Cambridge University Press.
- Romaine, Suzanne. 1982. *Socio-historical linguistics*. Cambridge: Cambridge University Press.
<https://doi.org/10.1017/CBO9780511720130>
- Romaine, Suzanne. 1984. Towards a typology of relative-clause formation strategies in Germanic. In Jacek Fisiak (ed.), *Historical syntax*, 437–470. Berlin: Mouton de Gruyter.
<https://doi.org/10.1515/9783110824032.437>
- Rutten, Gijsbert & Marijke van der Wal. 2012. Functions of epistolary formulae in Dutch letters from the seventeenth and eighteenth centuries. *Journal of Historical Pragmatics* 13(2). 173–201. <https://doi.org/10.1075/jhp.13.2.01rut>
- Rutten, Gijsbert & Marijke van der Wal. 2014. Social and constructional diffusion: Relative clauses in seventeenth- and eighteenth-century Dutch. In Ronny Boogaart, Timothy Colleman & Gijsbert Rutten (eds.), *Extending the scope of construction grammar*, 181–205. Berlin, New York: Walter de Gruyter.
- Salaburu, Pello. 2018. El euskera contemporáneo: el largo camino de la unificación literaria. In Joaquín Gorrochategui, Iván Igartua & Joseba A. Lakarra (eds.), *Historia del euskera*, 799–869. Vitoria-Gasteiz: Eusko Jaurlaritza.
- Sarasola, Ibon. 1975. *Euskal literatura numerotan*. San Sebastián: Kriseilu.
- Stark, Elisabeth. 2016. Relative clauses. In Adam Ledgeway & Maiden Martin (eds.), *The Oxford guide to the Romance languages*, 1029–1040. Oxford: Oxford University Press.
<https://doi.org/10.1093/acprof:oso/9780199677108.003.0064>
- Trask, Robert L. 1998. The typological position of Basque: Then and now. *Language Sciences* 20. 313–24. [https://doi.org/10.1016/S0388-0001\(98\)00006-0](https://doi.org/10.1016/S0388-0001(98)00006-0)
- Trebiño, Imanol. 2001. *Administrazio zibileko testu historikoak*. Vitoria-Gasteiz: IVAP.
- Urgell, Blanca. 2018. Lehen Euskara Modernoa. In Joakin Gorrotxategi, Iván Igartua & Joseba A. Lakarra (eds.), *Euskararen historia*, 549–657. Vitoria-Gasteiz: Eusko Jaurlaritza.
- Vicari, Stefano. 2018. Rituels épistolaires dans les lettres des poilus peu et moins lettrés: Une analyse contrastive. *SHS Web of Conferences* 46. 06009.
<https://doi.org/10.1051/shsconf/20184606009>
- Wal, Marijke van der & Gijsbert Rutten. 2013. The practice of letter writing: Skills, models, and Early Modern Dutch manuals. *Language & History* 56(1). 18–32.
<https://doi.org/10.1179/1759753613Z.0000000014>
- Wal, Marijke van der & Gijsbert Rutten. 2016. At the crossroads. Orality and literacy in Early and Late Modern Dutch private letters. In Ann-Catrine Edlund, Timothy G. Ashplant & Anna Kuismin (eds.), *Reading and writing from below: exploring the margins of modernity*, 197–214. Umeå: Umeå University.
- Wray, Alison. 2005. *Formulaic language and the lexicon*. Cambridge: Cambridge University Press.
- Zuazo, Koldo. 2019. *Standard Basque and its dialects*. London-New York: Routledge.
<https://doi.org/10.4324/9780429429729>