



Citation: Gorla, E., & Gasparini, F. (2025). Towards an Oral Corpus for Heritage Piedmontese, *Oral Archives Journal*, 1: 35-55. doi: 10.36253/oar-3337

Received: June 13, 2024

Accepted: November 15, 2024

Published: February 20, 2025

Data Availability Statement: All relevant data are within the paper and its Supporting Information files.

Competing Interests: The Author(s) declare(s) no conflict of interest.

ORCID

EG: 0000-0003-0713-8961

FG: 0000-0002-9196-8695

Ai soli fini concorsuali, l'attribuzione dei paragrafi è la seguente: Eugenio Gorla e Fabio Gasparini per il § 1, Eugenio Gorla per i §§ 2, 3, 4; Fabio Gasparini per i §§ 5, 6, 7.

© 2025 Author(s). This is an open access, peer-reviewed article published by Firenze University Press and USiena PRESS (<https://www.fupress.com>) and distributed, except where otherwise noted, under the terms of the CC BY 4.0 License for content and CC0 1.0 Universal for metadata.

Oral Data Production

Towards an Oral Corpus for Heritage Piedmontese

EUGENIO GORLA^{1*}, FABIO GASPARINI²

¹ *Università di Torino, Italy*

² *Freie Universität Berlin, Germany*

E-mail: eugenio.gorla@unito.it; fabio.gasparini@fu-berlin.de

*Corresponding author

Abstract. This paper presents the creation of an oral corpus as part of the ongoing PILAR project (Piedmontese Language in Argentina, 2019 – present), focused on the linguistic and ethnographic documentation of Piedmontese (Italo-Romance) as a heritage language in Argentina. The project gathers linguistic autobiographies and video recordings of grassroots initiatives that either actively use Piedmontese or reference cultural elements from Piedmont, such as music, folk songs, and traditional cuisine. Piedmontese has been spoken in Argentina since the late 19th century, when the government encouraged European migration to support agricultural development in the central provinces of Córdoba, Santa Fe, and Entre Ríos. High numbers of Piedmontese-speaking migrants and the isolation of these communities from urban centres allowed the language to persist longer than other Italian dialects in Argentina. As language shift eventually occurred, a revival movement emerged in Argentina, inspired by similar efforts in the Piedmont region. This paper details the corpus, its contents, and its significance in preserving Piedmontese as a living cultural and linguistic heritage in Argentina.

Keywords: oral data production, heritage language, Piedmontese.

1. INTRODUCTION

Various definitions have been given of heritage languages (HLs) (Rothman 2009; Benmamoun et al. 2013; Nagy 2014; Polinsky 2018; Aalberse et al. 2019; among many others), focussing on a number of typical features that define these language varieties and the communities in which they are spoken. On the social side, they are spoken by minority groups within a composite society who have a different ethnic origin from the dominant group and, most typically, have a migratory background. On the linguistic side, HLs are characterised by specific acquisitional features: they are learned and passed on in

the family environment, but their acquisition is often interrupted once the speakers, to the latest when they start going to school, interact more frequently in the dominant language, so that heritage speakers tend to display noticeable innovations in their use of the language compared to homeland speakers. Moreover, since the use of the HL tends to become less frequent over the lifespan of an individual, HLs are also exposed to language attrition, so that their speakers tend to lose parts of their L1 grammar (see Schmid 2005; Sorace 2011; among many others). Finally, language contact also plays an important part: lexicon and grammatical constructions from the dominant language are frequently used by heritage speakers through code-mixing (Auer 2014; 2022) potentially causing long-term grammatical changes in the HL in the long run¹.

Oral archives containing data from HLs are particularly useful for linguistic research, because they provide evidence on phenomena that are hardly observable in speakers who belong to other linguistic ecologies; furthermore, their realisation is urgent because they preserve evidence of small-scale sociolinguistic situations that are often unstable over time, due to language shift towards the majority language. Nevertheless, archives of HLs are usually not included in the most known repositories in language documentation such as ELAR or The Language Archive. This is mainly due to the fact that the rationale for inclusion in language documentation repositories is the overall level of endangerment of the language: since HLs are fundamentally linguistic varieties of languages that are often not endangered in the homeland, their documentation tends to fall outside of the scope of language documentation in a narrow sense. One of the few exceptions is Nagy's (2020) dataset, which has been partly made available online.

In this paper, we present the methodological choices adopted for collecting spoken data and building a language documentation corpus for the variety of Piedmontese (Italo-Romance) spoken as a HL in Argentina, in the provinces of Córdoba and Santa Fe. Throughout the paper we will refer to this variety as 'heritage Piedmontese' (HP), while the name 'Piedmontese' will be used for the homeland variety. In Section 1, we provide a short description of the sociolinguistic situation of Piedmontese in Italy, while Section 2 is dedicated to a socio-historical overview of the history of the HP community. We then move, in Section 3, to the description of the fieldwork methods adopted in the ongoing PILAR project, aimed at the documentation of HP, and the final structure of the dataset. Section 4 focuses into greater detail on the post-fieldwork treatment of the materials and sketches out envisaged possibilities for permanent archiving of the dataset. Final remarks are presented in Section 5.

2. THE STATUS AND DOCUMENTATION OF PIEDMONTESE IN ITALY

Piedmontese (ISO: pms) is an Italo-Romance language² spoken in the Italian North-Western region *Piemonte* ('Piedmont' in English). According to the Ethnologue database, its

¹ Another contact dynamics that is typical of HLs is koineisation between various dialects of the same language being spoken in the same migrant setting. In the sociolinguistic readership, the products of these dynamics have been sometimes referred to as 'migrant koines' (Kerswill 2006). Since the present paper is more concerned with the methodology used for data collection and organisation, we will not discuss this aspect in detail, but for an evaluation of koineisation processes in heritage Piedmontese, see Cerruti et al. (submitted).

² In the Italian linguistic and dialectological tradition, it is commonplace to refer to Italo-Romance varieties, including Piedmontese, as *dialetti*, 'dialects'. While the term points to their lack of an official status within the Italian legislation

estimated speakers range from 10 thousand to one million; Regis (2012) assesses its speakers to approximately 700 thousand. Most crucially, though, virtually all speakers of Piedmontese are also speakers of Italian, which represents the sociolinguistically dominant language. As well described in key works in Italian sociolinguistics (see e.g. De Mauro 1963; Berruto 2012; among many others), Piedmontese started undergoing language shift towards Italian after the Second World War, especially in urban environments, due to social reasons such as the introduction of compulsory education in Italian throughout the national territory, and the diffusion of Italian-speaking mass media.

In terms of endangerment, Piedmontese is considered a threatened language in the Ethnologue database, and 'definitely endangered' according to UNESCO's Atlas of the World's Languages in Danger (Moseley 2010); it is not included among the officially recognised linguistic minorities of Italy, and therefore it does not benefit from public policies aimed at its revitalisation. However, as emerges from the overview presented in Duberti and Miola (2022), the lack of safeguard at the national level has been partially counterbalanced by the emergence of various grassroots initiatives since as early as the beginning of the 20th century, with aims and purposes that changed over time. While at the beginning, local intellectuals organised to promote literary production in Piedmontese, more recently the efforts have been focusing on trying to prevent language shift by creating contexts for language use and, in some cases, by organising language courses at various levels. Partial support to these initiatives also came from local public institutions such as the Regional government and private cultural institutions like the *Centro Studi Piemontesi* (Centre for Piedmontese Studies). At present, Piedmontese is also taught at the University of Turin, as an optional subject in MA programmes in Linguistics and Italian literature (Duberti and Miola 2022).

Piedmontese has undergone various attempts of normative standardisation based on the linguistic features of the variety used in Torino, the capital city of Piedmont. Grammars and dictionaries have been produced, with different purposes, at least since the 18th century; for reasons of space we limit ourselves to mention the descriptive grammar by Tosco et al. (2023). Various orthographies have been elaborated over time, but without gaining universal consensus even among language activists (see Regis 2012; Regis and Rivoira 2019; for an overview). Currently, most of the written formal productions follow the norms elaborated in Brero and Bertodatti's (1988) normative grammar, but alternative proposals also exist (see e.g. Villata 2009). However, studies on more spontaneous uses (see e.g. Gorla 2012), such as commercial writing, brand names, and so on also reveal the presence of improvised orthographies often based on Italian spelling.

3. PIEDMONTESE AS A HERITAGE LANGUAGE IN ARGENTINA

Piedmontese migration to Argentina reached its peak in what historian Nascimbene (1987) refers to as the 'North-Western phase' in Italian migration to this country, which spans

and their functional subordination to the national language, it is sometimes looked down upon, under the assumption that it misrepresents their being autonomous grammatical systems, and could convey social stigmatisation. We will not go further into this terminological dispute. In this paper we will use the term Italo-Romance language since it is more used in the international readership.

from the last decades of the 19th Century to the outbreak of the First World War in Italy, in 1915. During these decades, a vast majority of the migrants relocating from Italy to Argentina came from Piedmont: according to Italian records, between 1879 and 1890 most of the Italian migrants came from Piedmont (22% of the total), followed by Lombardy (19%) and Veneto (12%) (Devoto 2006, 106; see also Bagna 2011).

In this period, The Argentine government explicitly invited the arrival of migrant workforce from Europe through the *Ley Avellaneda* (“Avellaneda Law”, from the name of the President by which it was enforced), which introduced a regulation for immigration, in order to improve the agricultural exploitation of the grasslands in the provinces of Córdoba and Santa Fe. Migration and agricultural colonisation are deeply interconnected in this context (Djenderedjan 2008), as migrants were also the founders of new settlements where they represented a majority. This had major consequences also on the linguistic history of these communities: while in urban migrations (e.g. in the city of Buenos Aires) the ‘argentinisation’ of Italian migrants had been quick and necessary to their integration, this was not the case in the rural settlements of Córdoba and Santa Fe (Crolla 2015). Due to their relative isolation, Piedmontese migrants here had the occasion to retain their linguistic repertoires for a longer period of time, and thus, Piedmontese underwent language shift at a much slower rate compared to other situations involving Italo-Romance varieties; at the same time, it developed the typical features that characterise heritage languages (see Section 0) both on the social side and on the linguistic side, hence the label ‘heritage Piedmontese’ that is used in this paper.

The picture drawn so far allows us to identify major differences with respect to other communities of Italian origin in the world. Most of the existing studies have been focusing on migration flows that took place after the Second World War, which show major differences from the situation we are dealing with. In general accounts such as Turchetta (2005) and Vedovelli (2011), the typical situation of Italian communities is best described by the 3-generations shift model (Fishman 1966), where the community with a migratory background, in the span of three generations, loses the heritage language and is fully absorbed by the dominant linguistic repertoire of the new country of residence. Moreover, Italo-Romance varieties are believed to be abandoned at a very early stage in the migrant community, as in some cases they may hinder interaction even inside the community itself (Bettoni and Gibbons 1988).

The case of Piedmontese in Argentina stands apart from these situations, as it represents one of the few cases in which an Italo-Romance vernacular has been kept as the main language of the community, or at least the language with which the community identifies itself. A similar case is represented by Talian, a variety of Venetian dialect that has been retained in some rural areas of Rio Grande Do Sul, in Brazil (Brambatti Guzzo 2023 for a recent contribution). The case of Talian represents, to our knowledge, the sole parallel to the situation of Piedmontese in Argentina involving an Italo-Romance vernacular. However, while the Talian has been the object of recent initiatives of linguistic documentation, and a corpus of the language is being built, the same cannot be said for Piedmontese in Argentina prior to the launch of the project described here. Recent works on this community have been focusing on the reception of regional identities in rural Argentina (Crolla 2015) or have used an emic approach, which is central also in our methodology, in the analysis of personal and family narratives of the community. Following this methodology, Giolitto (2010) draws an extensive account of the community, reconstructing through oral narratives the process of language

shift towards Spanish, hinting at the emergence of a process of linguistic revival that is further being developed in the present day.

Associationalism started to spread in Argentina in the 1970s and reached its climax during the last decades of the 20th Century. Associations that were founded in this period often have the name *Familia Piemontesa* (Piedmontese Family), and operate on a strictly local basis. Their main goal is to promote cultural and recreational activities aimed at the celebration of Piedmontese identity within the community. HP associationism spread at the same pace as the practice of town-twinning between Piedmontese and Argentine cities, also due to the activity of the Italian association *Piemontesi nel mondo* (Piedmontese worldwide), which sought to strengthen Piedmontese identity by (re-)creating relationships between Piedmontese descendants and their homeland (see also Giolitto 2010). At present the vast majority of these associations are part of a single federation named *FAPA Federación de Asociaciones Piemontesas de la Argentina* (Federation of Piedmontese Associations in Argentina), whose central administration at the national level coordinates the actions undertaken by each local association.

Crucially, Italian language plays a relatively small part in this process: on the one hand language revival has reinforced relationships not only with the region Piedmont, but with Italy in general, which includes a greater exposure to Italian overall; on the other hand, though, the regional basis of this revival even in more recent times prevented more systematic contact between heritage Piedmontese and Italian.

To conclude, we may identify two main gaps in the description of this community. The first one is represented by the absence of a sizable documentation of heritage Piedmontese: written testimonies³ are scattered across private archives both in Italy and in Argentina, but have never been digitalised; most notably, no attempt has been made, to our knowledge, to document spoken heritage Piedmontese – or at least, no attempts have been made to organise data coming from previous collections, and particularly the one that informed Giolitto's (2010) work, in the form of a digital and openly accessible language documentation corpus. This is in our view the first step towards the linguistic description of heritage Piedmontese, and represents one of the long-term goals of the PILAR project, described in Section 3.

The second open point that needs to be addressed concerns the sociolinguistic dynamics in which heritage Piedmontese is involved. While the narratives that were collected by Giolitto (2010) seem to point to a “linear” pattern of language shift towards the dominant language, little attention has been given to the counterweight to this tendency that is represented by linguistic and cultural revival of HP, as noticed by Giolitto himself. As emerges from preliminary analyses (Goria 2015; 2023), most recent uses of Piedmontese are in fact to be located in a climate of renewed interest towards Piedmontese cultural heritage from the Argentinian side, and also from the Italian side, in terms of a rediscovery of the emigration from Piedmont. This led to various initiatives, ranging from town-twinning projects between Piedmont and Argentina to international summits, and in general the strengthening of transnational relations, both at a private and a more institutional level. It can thus be argued that due to the increase of literacy and education within a Spanish-speaking society, the heritage

³ Heritage Piedmontese developed a written tradition only to a limited extent; exceptions include Piedmontese-Spanish local journals that began being published in Argentina in the 1970s, at the start of the linguistic revival, and private correspondences.

language is now only spoken on a limited number of occasions, related to the ongoing revival of Piedmontese language and culture, where Piedmontese associations sponsor programmes of language maintenance and revitalisation. With a few hundred Piedmontese descendants nowadays actively involved in associationism, Piedmontese is identified as an important resource to manifest local identity and claim community membership; see Gasparini and Gorla (in prep.).

Therefore, we contend that a more systematic documentation and description of the cultural practices that are related to the use of Piedmontese in Argentina, may sensibly contribute to our understanding of the sociolinguistic dynamics that characterise this variety and this community.

4. THE PILAR CORPUS OF HERITAGE PIEDMONTESE

The project *PILAR – Piedmontese Language in Argentina* (<https://sites.google.com/unito.it/pilar/pilar>) seeks to follow the two research perspectives mentioned in Section 2: (i) collecting and organising first-hand language documentation materials on heritage Piedmontese, and (ii) offering an updated perspective on macro-level social processes of language shift and language revival. This required, as will be illustrated, the adoption of a mixed methodology that combines common techniques in language documentation, with ethnographic observation, which led to the collection of heterogeneous materials in terms of formats (audio vs. video), languages of interaction (HP vs. Spanish), types of observation (interview vs. participant observation), number of subjects involved. For this reason, the materials have been divided and classified based on the setting in which the recordings took place, which resulted into two sets: (i) sociolinguistic interviews and group conversations conducted with HP as the target language, to be primarily used for the linguistic analysis, and (ii) cultural practices performed during public gatherings, relevant for the ethnographic documentation of semi-spontaneous multilingual interaction and the self-representation of the community. For this reason, this section is structured as follows: in 3.1 we will present a global overview of the methods adopted during the two fieldwork sessions that have been carried out in 2019 and 2022; in sections 3.2 and 3.3, we will then provide a detailed account of the amount and type of materials that were collected, respectively for language documentation (interviews and group conversations) and for ethnographic documentation (cultural events).

4.1. Fieldwork methods

Given the twofold nature of the research questions discussed in Section 2.2, the fieldwork methodology adopted for this study had to benefit from a combination of the methods that are typically used in neighbouring sub-disciplines of linguistics, and in particular language documentation (Austin 2006), ethnography (Blommaert and Dong 2015) and sociolinguistic research on language variation and change (Tagliamonte 2006). As argued for other research works on Italo-Romance varieties such as Mereu (2022), on the one hand we followed traditional approaches to the documentation of endangered languages, and gave priority to the gathering of a corpus of oral recordings of HP. On the other hand, as a complement to the

corpus, we also collected rich sociolinguistic data, concerning personal biographies, language attitudes and reported linguistic practices.

Two fieldwork sessions of one month each were carried out in 2019 and in 2022. The first one had as its main goal to identify the social network formed by the speakers of heritage Piedmontese (see further) and locate the individuals and communities available for further investigation. Also, the 2019 fieldwork was primarily aimed at collecting materials for narrow linguistic analysis, and thus exclusively focused on audio recordings. The 2022 fieldwork session was carried out mainly to include in the project multimodal ethnographic documentation of cultural events related to Piedmontese language, and for this reason, sessions were both audio-recorded and filmed. Audio recordings were carried out with a Zoom H4n pro stereo recorder in wav format, with a 24bit/48KhZ sampling rate; in 2022, when possible, an external lavalier microphone was used with single speakers in order to reduce ambient noise. For video recordings, we used a professional Sony ILCE-7RM3A camera in .mp4 format and 4k resolution.

Informants and locations were chosen based on the presence of associations of Piedmontese descendants. FAPA (see Section 2) was thus formally involved in the second stage of the project: the researchers were clear about their scientific goals since the first contacts with the organisation, presenting themselves as linguists who were carrying out a study on heritage Piedmontese, and stating their interest in how associations were working to promote language revitalisation and language use. To justify the systematic use of video recordings in each location, we proposed to the community the realisation of a documentary film with a selection of these materials, to which the community agreed (see Section 4).

A first contact was made in order to identify the associations that were more active and which had among their members some who were able to speak Piedmontese. Once on the field, the same members offered to provide contacts with other Piedmontese speakers in the area, which qualifies the methodology adopted here as ‘snowball sampling’ (Buchstaller and Khattab 2013), or ‘friend-of-a-friend’ methodology (Milroy and Milroy 1987; Tagliamonte 2006). This technique is considered particularly effective in cases where the community is small and the participants are difficult to find outside of the social network of which they are part; for a similar case see Mereu (2019). In this situation, such methodology was particularly needed, especially because of the relatively small number of Piedmontese speakers and the great distance between the various cities and villages. By contrast, other forms of sampling would have been less effective due to the fact that etic categories such as age, gender, social class, or even an attested Piedmontese ancestry would have failed to identify an adequate number of actual Piedmontese speakers, even within the network of Piedmontese associations.

The primary criterion for selecting an interviewee was thus based mainly on in-group designation as an enthusiast member and good representative of the HP community. This simplified our search for adequate consultants, but inevitably shifted the object of our analysis from Piedmontese descendants in a merely genealogical sense (i.e. any individual of Piedmontese ancestry living in the area), to a category of ‘language revivers’ who, besides having Piedmontese ancestry, are actively committed to the promotion and diffusion of Piedmontese as a heritage language in Argentina, or simply take part in the activities of the local association. As is often the case, this choice blurs our view on the vitality of the language, as we deliberately choose to focus exclusively on linguistic practices within one social network. In this data,

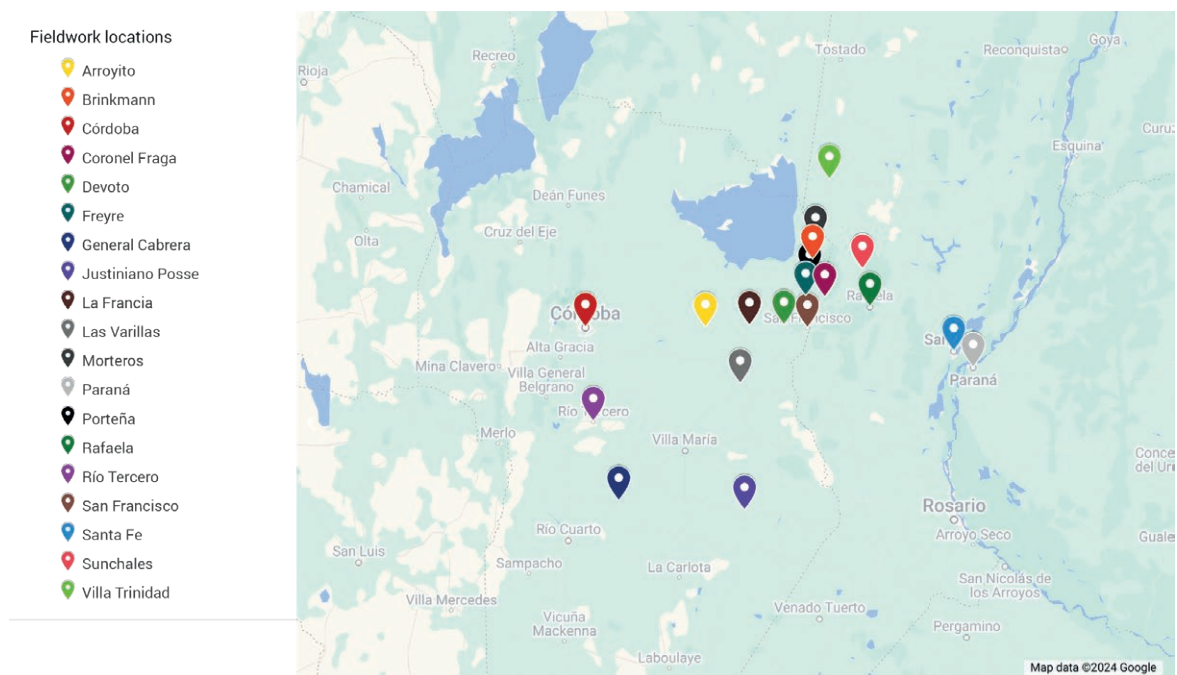


Figure 1. Sites of data collection.

linguistic knowledge of the heritage language is enacted performatively, as a tool to build and express group membership rather than employed for daily communication needs (Gasparini and Gorla in prep.) Nonetheless, language is not the only element through which Piedmontese identity is claimed by the community. It was often the case that non-proficient speakers were introduced to the researchers as relevant members of the community, or volunteered in order to share their experience and point of view. They were often included in the interview sessions.

In both fieldwork sessions, the research was carried out in rural areas in the provinces of Córdoba and Santa Fe. As can be seen in the map in Figure 1, the research only included four large cities, namely Buenos Aires (not included in the map), Córdoba, Santa Fe and Paraná. The locations considered during the first inquiry are: Arroyito, Córdoba, Coronel Fraga, Devoto, Freyre, La Francia, Morteros, Paraná, Porteña, Rafaela, San Francisco, Sunchales, and Villa Trinidad. During the second fieldwork session, recordings were carried out in Buenos Aires, Brinkmann, General Cabrera, Justiniano Posse, Las Varillas, Morteros, Rafaela, Río Tercero, San Francisco, and Santa Fe.

A total amount of 130 individuals participated as interviewees in the data collection, either in face-to-face interviews, or, especially in the 2019 session, in group conversations with the researchers. For this reason, not all the informants contributed in the same measure to the data collection. Moreover, even though the researchers explicitly mentioned the need to record individuals who were able to engage in a conversation in heritage Piedmontese, some interviews have been carried out in Spanish. This occurred when the researchers felt the need to accommodate the preference of single informants who declared to not be able of speaking Piedmontese or believed to have an insufficient level of proficiency (see Section 2.2). Moreover, no a priori sampling was made based on etic categories that are typical in

sociolinguistic research, such as age and social class. Gender has been considered a potentially relevant social feature that may have an effect on linguistic practices, but even in this case no sampling has been made, due to the fact that the availability of informants of each gender was unavoidably dependent on the number of men and women who participate in Piedmontese cultural activities in each location, and a deeper insight into community practices would have been needed. At the same time, since the interview involved the narration of family histories and autobiographical narratives, we were able to collect rich information concerning the time in which the families migrated from Italy, and the degree of proximity of the individuals with the first generation of Italian migrants.

As a major point of difference with most studies on heritage languages, we chose not to rely programmatically on a distinction between first-generation speakers and those who belong to the second and subsequent generations. A considerable amount of readership (see e.g. Benmamoun et al. 2013; Polinsky 2018; Polinsky and Scontras 2020) is strongly based on the idea of a radical distinction between generation 1 speakers, who are typically adults who learned the heritage language in the homeland, and generation 2 speakers, who learned the heritage language within the household, but in most of the other domains, and from the school age onwards, have been mostly using the dominant language in society. In the case of heritage Piedmontese, though, the most intense migration wave dates back to the late 19th and early 20th century, and most members of the community in present day are fully integrated in the Argentine society and are separated from the first generation by various degrees of relationship, up to the fourth and in some cases fifth generation. The few first-generation speakers of Piedmontese that were interviewed were, in any case, individuals who arrived in Argentina during the 40s and 50s and in most cases attended school in Argentina: in terms of linguistic practices are thus hardly comparable to the first generation of ‘early migrants’. Based on autobiographical narratives, we prefer to distinguish between simultaneous bilinguals, who learned both Spanish and Piedmontese in the same environment, and sequential bilinguals, who first learned the heritage language, and subsequently learned Spanish at school. These two categories are represented in Figure 2, which focuses on the type of bilingualism, associated to the time of the arrival of the families in Argentina; it must be clarified though

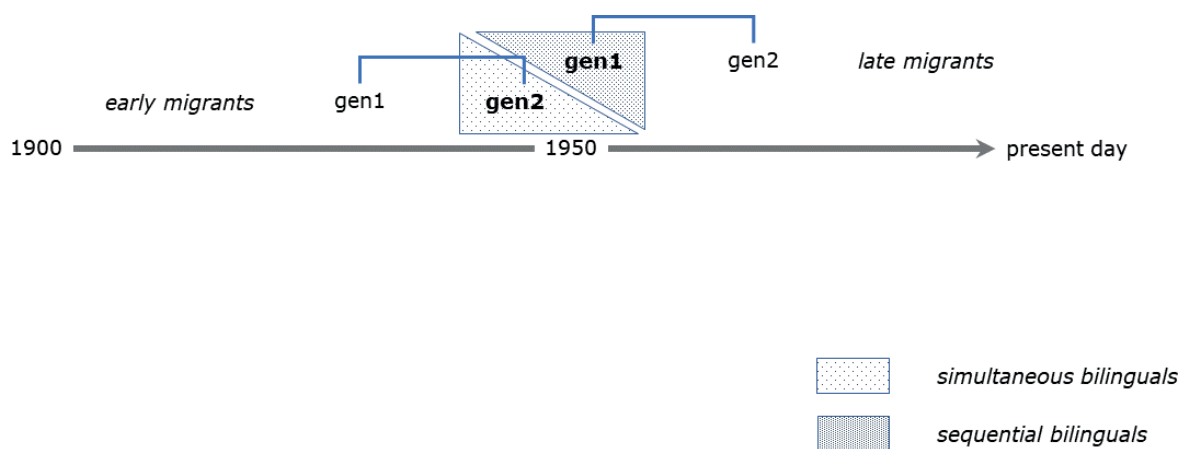


Figure 2. Bilingualism among different generations of Piedmontese migrants in Argentina.

Table 1. Summary of the collected materials in the two fieldwork sessions, divided per type of activity, main language, and number of sessions.

Year	Type	Working language	N Sessions	Total length
2019	Interviews	pms	16	08:03:09
		pms-spa	7	06:01:40
TOTAL				14:04:00
2022	Song	pms-spa	6	01:54:26
	Community, song	pms-spa	2	02:42:57
	Community	pms-spa	4	01:21:37
	Community, Cuisine	pms-spa	1	00:16:29
	Conversation	pms-spa	5	03:09:45
	Cuisine	pms-spa	1	02:10:41
	Photograph	pms-spa	10	00:56:34
	Photograph, Conversation	pms-spa	1	00:50:26
	Theatre	pms-spa	1	01:00:47
TOTAL				14:23:42
2022	Interview	spa	22	07:59:54
	Interview	spa-pms	6	03:27:38
	Interview	pms	23	06:49:45
TOTAL				18:17:17

that the category of ‘late migrants’ indicated in the figure is by no means comparable in size to the ‘early migrants’ whose descendants, as said, represent the vast majority of the community. The label ‘late migrants’ was introduced *ex post* in order to make explicit the qualitative difference between first-generation speakers who are observable in the present, and belong to the ‘late’ wave, and early first-generation speakers, who are the ancestors of the speakers observed in this study. Furthermore, a preliminary analysis of the fieldwork materials and participant observation revealed that the use of HP, and more generally, identification within the HP community, are situated practices mediated by specific cultural activities, which in some cases also include semi-guided teaching of the language. As such, language use is not so much dependent on ‘linear’ intergenerational transmission as on individual biographies and ideological orientations.

To conclude, the final dataset consists of approximately 46 hours of audio and video recordings that were classified in the following way. We considered ‘interviews’ those interactions where the researcher actively participates leading the conversation through a series of recurring questions, situationally adapted to the context and the speaker. A summary of the recorded sessions is given in Table 1, and a more detailed table is provided in the Appendix of this paper.

4.2. Interviews

Interviews were conducted by one of the two researchers present on the field, using the heritage language in order to stimulate answers in this language, which made it possible to

record several hours of semi-spontaneous interactions in HP (see Blommaert and Dong 2015). Some speakers who declared to be unable to speak heritage Piedmontese were also interviewed in Spanish, as they volunteered to participate in the interviews in order to be able to tell their family histories or about their personal involvement in the local association. For the reasons outlined in Section 2, Italian was never used for data collection.

We chose to adopt as a model the sociolinguistic interview (Labov 1984; Eckert 2000; Tagliamonte 2006; among others). This technique consists in creating an environment for the researcher to engage in conversation with their informant(s) about a set of topics that are relevant for the ongoing inquiry. Questions are not fixed and follow, as much as possible, the global flow of the conversation; the interviewer often adapts to the type of the contribution that the interviewee is willing to give, without forcing questions that are not relevant for that specific interaction. Below follows a list of topics that were covered:

(1) Script used for interviews

Age of the speaker

Family history

- Who in the family came from Italy?
- In which year?
- Where did they come from?

Personal relationship with Piedmontese

- Where did you learn Piedmontese?
- Have you ever been to Italy?
- With whom did you use Piedmontese in the past?
- With whom do you use Piedmontese in the present?
- How and when did you get involved in activities related to Piedmontese?
- What activities do you do related to Piedmontese culture?

Attitudes towards Piedmontese?

- Why are you committed to Piedmontese culture?
- Do you think that there has been a renewed interest in Piedmontese in the last decades? Why?
- Is Piedmontese spoken in Argentina different from that spoken in Italy?
- Is it good or bad that Argentine Piedmontese is different from homeland Piedmontese?

When possible, informants were interviewed singularly or in pairs, in order to elicit a comparable amount of speech for each participant. These were labelled as ‘interviews’ proper. In some cases, though, the researchers had little control over the interactional setting, and, due to time constraints, ‘group interviews’ had to be introduced as a second subtype of the genre interview. This second type of interaction shares with full-fledged interviews the semi-spontaneous character of the interaction, determined by the presence of an external researcher who openly poses questions to the group; at the same time, greater freedom is left to the participants in terms of the amount and quality of information that is given.

As for linguistic choices, interviews globally show an “intended monolingual” (Clyne 2003; Dal Negro 2013) behaviour: speakers who were able to speak Piedmontese firmly adopted this language during the whole activity, while in other cases Spanish was negotiated as the language of interaction. This reflects a typical paradox in the documentation of minority languages, where, in spite of the fact that the language is spoken in a highly plurilingual

setting, informants – and activists even more so – tend to present themselves as ‘ideal’ speakers of the language, and therefore to adopt a typically monolingual style where contributions from the other languages of the repertoire are kept to a minimum. At the same time, use of the other language can never be completely avoided and both conversational code-switching and pragmatically neutral code-mixing (Auer 1999) can be observed. Consider example (2)

- (2) Y ESTO ES COMO APARE- belessì COMO l’ha APARECí (0.3) e:hm (0.3) ël FERROCARRIL (0.4) tut ël mond a vnisí an sa (.) noi l’ha fasse gròs Morteros (0.3) como l’era UNA PUNTA DE LINEA (0.2) pèrchè FINALIZava belessì (0.2) no? ENTONCES (0.7) pi gent a vnisia PARA VER (.) SER a ramba dël FERRO-CARRIL (0.2) pèrchè ël FERROCARRIL a l’è col che a l’ha portate (.) ël PROGRESO

And this is how it appeared here how it appeared (0.3) ehm (0.3) the railway (0.4) everybody would come here (.) and we, it became big, Morteros (0.3) as it was the end of the line (0.2) because it ended here (0.2) no? So (0.7) more people would come in order to see (.) be close to the railway (0.2) because the railway is what brought (.) the progress.

At a global level, Piedmontese appears to be selected as the language of the interaction, as can be seen also from the self-repair at the beginning of the quote. From an emic perspective, the speaker is thus speaking in Piedmontese. At the same time, the type of speech produced in this activity corresponds to a bilingual mode (Grosjean 2013) where both Spanish and Piedmontese are activated. The speaker in fact resorts to various insertions from Spanish, with different extent, ranging from single morphemes (e.g. APAREC-ì “appeared”) to entire phrases (e.g. UNA PUNTA DE LINEA “an end of the line”).

In group interviews, the researchers maintained the same behaviour as in single ones, and Piedmontese was always offered as the default language of the interview. The observed behaviours were, however, more diverse: in some cases, the presence of proficient speakers of Piedmontese encouraged those who had lesser competence to use Piedmontese anyway, in order to accommodate the language of the interviewer. In other cases, though, Spanish was negotiated as the language of the activity by the majority of the group, and thus the whole activity was carried out in this language.

4.3. *Ethnographic documentation*

Ethnographic documentation was conducted with a handheld camera with the aid of a tripod and a gimbal support, based on the situation’s requirements, and was aimed at collecting information on the types of activities that were carried out when HP is used in the community. This resulted in documenting various activities organised by the Piedmontese associations. A selection of our collected visual material has been subsequently used to produce a short documentary (Gorla and Gasparini 2023, see Section 4).

The ethnographic documentation involved group activities performed during celebrations and gatherings of various sorts, such as culinary events (Figure 3), choir singing, traditional music (Figure 4) and theatre exhibitions. It was not possible to record any lesson of Piedmontese. Observation of these events was participant, in that the researchers were actively involved in the activities that were being carried out, or in some cases, were acknowledged spectators and intended addressees of the activity, especially in the case of music exhibitions. The data collected in this way complement and contribute to linguistic research in that



Figure 3. Preparation of *bagna caoda*.



Figure 4. Music performers during a convivial gathering.

they show ‘what is going on’ when the heritage language is used: they help contextualise the observed linguistic practices and provide access to the attitudes and beliefs of the speakers. Analysis of the activities carried out by Piedmontese associations was useful, in particular, to interpret the observed practices in terms of an ongoing process of language revival.

The main culinary event that is documented in the corpus is the preparation of *bagna caoda* (literally “hot sauce”): this is a traditional dipping sauce from lower Piedmont, prepared with oil, chopped anchovies and garlic. It is eaten with raw and cooked vegetables. The HP community prepares it as a traditional local dish during the *Semana Santa* (Holy Week) that precedes Catholic Easter. The picture portrays members of the *Familia Piemontesa* of *Rafaela* during the preparation of the sauce: this operation takes a full day and requires the participation of several people. Interaction during the process takes place mostly in Spanish, but in various occasions during the day Piedmontese and Italian songs are voiced spontaneously by the people at work. The sauce is then canned and sold as a fundraising activity for the association. This practice is consistently recognised within the community as a tangible sign of the Piedmontese presence: pots of *bagna caoda* are sold with the logo of the association, which is a modified version of the Piedmontese regional flag, and indirectly fulfil the mission of promoting and preserving Piedmontese culture in Argentina. It must also be noted that if on the one hand *bagna caoda* is treated and presented as a symbol of Piedmontese identity in the area, on the other hand people who are not part of Piedmontese associations consume this recipe merely as traditional Easter food of the area, without a specific interest in its origin. In fact, in other locations it was possible to see the same product in shops whose aesthetics and design were completely unrelated to Piedmont or Italy.

Traditional music plays an important role in the identity discourse of the HP community. Singing is a very widespread activity, as is proven by the presence of an amateur choir in almost every association that was visited. It is often the first HP-related activity that many members of the community undertake and some informants report to having started using Piedmontese again after beginning choir practice (see Gorla 2023). The repertoire consists of songs in Italian and Piedmontese, often composed by songwriters in the first half of the 20th century and later becoming part of the popular tradition. Some songs are more recent and were probably adopted during cultural exchanges between Piedmont and Argentina, or were composed specifically for similar occasions. It is also noteworthy that some choir singers in various locations are unable to speak Piedmontese outside of this context. The lyrics sheets include Spanish translations and notes for proper pronunciation, and additionally, choir directors often assist the singers in correctly pronouncing the lyrics.

A different type of performance that was observed involves the practice of playing folk songs, both traditional and authored, in Italian and Piedmontese by solo musicians or groups. Many practise the most known traditional instruments of Piedmontese tradition (accordion, mouth organ) together with instruments of other traditions (e.g. the classical guitar) and instruments of more recent diffusion, such as the electric keyboard. Notably, the folk instruments employed in the Piedmontese tradition, such as the *torototela* (a kind of monochord built out of a desiccated pig bladder or a bucket, chord and a stick or pole), woodwinds (such as the hurdy-gurdy), and traditional pifferos, flutes and bagpipes (see Raschieri 2019), are basically unknown to the community, probably meaning that memory of such instruments was likely not transmitted by the earlier migrants. In the events documented in the corpus, such performances are spontaneously organised by associations and groups committed to the Piedmontese revival during informal gatherings. These performances alternate with improvised narrations of anecdotes concerning the history of Piedmontese communities and humorous stories, both in Spanish and Piedmontese. As in choir performances, the focus remains strongly on the celebration of Piedmontese identity, and the intended audience is the community itself. On these occasions, however, a greater influence of local, or non-Piedmontese, music is noticed. This reveals a higher level of integration between local Argentine cultural elements and those brought by immigration than what emerges, for example, in explicit statements collected among the community. Additionally, in both types of performances, the noticeable presence of Italian music and language contrasts with the absence of Italian in other observed linguistic practices. This could be interpreted as a sign that these musical events have likely become more common in a more recent stage in the history of HP communities, benefiting from stronger contacts with Piedmont and Italy during a period when Italian was already a major language of communication in Italy.

5. TRANSCRIPTION AND ANNOTATION

The corpus collected with the methodology outlined in Section 3 is being transcribed and annotated with the ELAN software (Sloetjes and Wittenburg 2008). The proposed template for transcription and annotation has the structure exemplified in Figure 5.

The first line, indicated as (i) in Figure 5, corresponds to the orthographic transcription of heritage Piedmontese. Given the plurilingual nature of this data, two separate systems have

The screenshot shows the ELAN software interface with a transcription of a sentence in Piedmontese. The sentence is "mi quand andasian ant la (scòla),". The interface displays a timeline at the top with time markers from 00:09:26.800 to 00:09:28.200. Below the timeline, the transcription is shown with five levels of annotation:

Annotation Level	Text
(i) Attilio Be [464]	mi quand andasian ant la (scòla),
(ii) atilio-tok [2704]	mi quand andasian ant la scòla
(iii) atilio-m [3295]	mi quand andasi -an ant la scòl -a
(iv) atilio-g [3295]	1SG when go:IMP 3PL in the:F.SG school SG
(v) atilio-l [2760]	piem piem piem spa piem piem piem piem

Figure 5. Example of the annotation template.

been introduced for Spanish and Piedmontese. Spanish is always transcribed following the orthographic norm for written Spanish, but with an ad hoc treatment of sociolinguistically marked forms that characterise Argentinian Spanish. Typical phonetic features of this variety (e.g. phonetic reduction of /s/ in syllable codas, pre-tonic vowel lengthening in some parts of the Cordoba region) have been normalised following the orthography of standard Spanish. We chose however to retain lexical or morphological features, such as the use of *vos* and *ustedes* for 2sg and 2pl informal personal pronouns (*tu* and *vosotros/as* in Standard Spanish) and local verb inflection (e.g. Argentinian *tenés* vs. Standard *tienes*). The transcription system for heritage Piedmontese is based on the so-called literary orthography, which was introduced in the 1930s and further elaborated by Brero and Bertodatti (1988) in order to represent, fundamentally, the dialect of Turin. For a detailed description we refer to Tosco et al. (2023), who use the same writing system in their descriptive grammar of Piedmontese.

The transcription is then (ii) tokenised, so that each word corresponds to a single annotation. Subsequently, with ELAN's interlinearisation function the text is semi-automatically⁴ (iii) segmented into morphemes and (iv) glossed according to the Leipzig glossing rules (Bickel et al. 2008). Finally, each morpheme also (v) receives a language tag: the labels introduced are *spa* (= Spanish), *pie* (= Piedmontese), *ita* (= Italian), and *pie-spa* for homophonous forms.

6. ARCHIVING AND TRANSFER TO THE COMMUNITY: FUTURE PERSPECTIVES

Fieldwork data is collected with great effort, both in terms of time and money, by the researchers; such material can be of greater relevance to the community involved, since it can represent their culture, legacy and identity (Carroll et al. 2020). Digital archives can be a suitable form of restitution, in case the community has full access to computers and the internet (Kung 2020). The lack of open archives of Piedmontese built according to the academic standards, and more in general the absence of a corpus of homeland Piedmontese, makes the

⁴ We define this function as semi-automatic because the automatic operation of segmenting and glossing each morpheme is based on the manual construction of a Lexicon. For the same reason, since the work is still in progress, we are not able to share a list of the labels used to annotate the files.

creation of a digital archive for HP all the more tempting and urgent. For these reasons, after transcription and annotation of the recordings, data will be shared on an open access repository. The XML-based structure of eaf ELAN files seems particularly adequate to this task, as the same format is adopted for documentation materials of endangered languages.

In order to comply with ethical standards for research of this kind, data collection was planned in cooperation with the FAPA association at the national level (see Section 3.1): besides helping the researchers identify the most adequate contexts for data collection, its contribution was also to inform the local communities of the ongoing research and especially of the fact that the researchers needed to record and film interviews and particular activities that were held in these associations. Moreover, in each location the researchers were asked to publicly present the aims of their research. Informants recruited in this project then consented to share their voice and image both for research purposes and for the production of a documentary film; consent was acquired, when possible, by asking to sign a form, or in other cases by asking in the video recording permission to record and use the interview for the mentioned purposes.

It is common practice for researchers working on endangered languages to share .eaf files together with the recordings made available in their online corpus. Data are usually deposited on larger archives dealing with language documentation in general: some examples are DOBES (<https://dobes.mpi.nl/>) and the Endangered Languages Archive (ELAR: <https://www.elararchive.org/>). Further examples of the potentials of online corpora making use of ELAN transcribed files for deeper linguistic inquiries are CORPAFROAS (Mettouchi et al. 2010) and MultiCAST (Haig and Schnell 2023).

Besides, the corpus will enable researchers to investigate the dynamics of language contact between Spanish and Piedmontese that are unique to this scenario. Thanks to systematic annotation of the language, it will be possible to obtain a qualitative and quantitative account of how the two languages interact in grammar and discourse. With ELAN becoming a standard tool for linguistic research, experimental approaches aimed at taking advantage of the great potentiality of ELAN have been extensively developed. Such strategies can be successfully applied to our case study in the near future. As an example, ELAN could be useful for an investigation following variationist sociolinguistics approaches (Nagy and Meyerhoff 2015), or for the analysis of gesture (Azar et al. 2020) in relation to multilingualism and language contact. In particular, the corpus will be extremely adequate for the description of a still debated phenomenon in language contact literature, such as code mixing between structurally similar languages.

However, community needs do not always coincide with those of academia, and a digital archive may not be a really attractive option to lay people: they may not have the will to deal with ELAN and transcriptions, and may prefer a more immediate, direct access to the data. The making of a documentary out of the material gathered on the field seemed like an engaging way to provide the community with tangible results of their efforts and kindness as participants in the investigation and hosts during our fieldwork: thanks to the collaboration with a professional filmmaker and editor, Silvia Pesce, we were able to produce a short documentary (Gorla and Gasparini 2023) which, eventually, was screened different times among the local communities during a short trip back in the field in April 2024.

7. CONCLUSIONS AND FURTHER DEVELOPMENTS

The main aim of the PILAR project was to collect and organise language documentation materials for heritage Piedmontese. In order to build a comprehensive corpus, various types of data were collected during fieldwork: namely, linguistic, sociolinguistic and ethnographic. Dealing with a very specific language ecology, on a small-scale sociolinguistic situation and with a tight-knit social network required the adoption of typical methods of ethnographic fieldwork, which include a relatively long period of observation, active participation of the researchers in the observed cultural practices, and creation of stable relationships with the community. Use of video recordings was fundamental to this approach, as it allowed to collect rich information also on cultural activities indirectly related to the use of heritage Piedmontese, and to extend the documentation also to such activities, as is typical in “documentary linguistics” (Riessler and Wilbur 2017). From a strictly linguistic perspective, the scope of the project was more general, as it had to deal with the great level of endangerment of Piedmontese, both in its homeland and heritage varieties, and with the absence of available corpora; for this reason, more traditional research tools, such as sociolinguistic interviews in the heritage language, to ensure that a sufficient amount of data was collected.

In terms of fieldwork methodology, our ethnographic recording of cultural practices during the second stint of fieldwork was inevitably influenced by the way we interacted with the community on the field (Kilani 1995): even if most of the social events we recorded were planned independently from our presence, many participants saw in our investigation the perfect occasion to showcase individual activities as well as the association’s engagement in preserving Piedmontese traditions. Also, given the great distance between the locations and the relatively short amount of time, our fieldwork was not immersive, but mostly interview-based: such limitations did not allow us to undergo that process which De Sardan (1995) defines “saturation”, meaning the assimilation of social behaviours without the mediation of formal education or direct research inquiry, which can provide further insights about the researched group, as also advocated by Aikhenvald (2007). But at the same time, personal bonds with some members of the community were built during both the time on the field and through constant digital communication in the months following the two fieldwork sessions, thereby helping in mitigating this shortcoming. Another major drawback was represented by the little time that was spent in each location. Since the research design was based on a systematic collaboration with the FAPA federation, our schedule had to be agreed upon and organised along with local associations prior to our arrival on the field, also based on each association’s requirements and needs. Therefore, we could spend only a few days with each community along the way, and we had only few occasions to observe real daily language practices within the household, outside of controlled recording sessions. However, as the data collected so far reveal, HP is hardly used outside of the context of Piedmontese associationism. We therefore argue that immersive fieldwork, without the support of local associations, would have led to poorer results for a documentation of HP.

The main research outcomes expected from this fieldwork experience reflect the twofold need for an easily accessible documentation and for rich metadata that satisfyingly describe the scrutinised situation. On the one hand, the publication of the PILAR data as a corpus of spoken heritage Piedmontese will be able to fill the existing gap in the documentation of Piedmontese

as an endangered language, both in the homeland and in Argentina. On the other hand, the collection of ethnographic material and linguistic autobiographies will make it possible to carry out in-depth studies on various aspects of language contact between Spanish and HP as well as on language revival as a sociological phenomenon that characterises the HP community.

Documentary videos played an important role in data collection as well as in the subsequent stages. As said, using videos enabled to stimulate the participation of the community in fieldwork activities and, while on certain occasions it prevented less obtrusive forms of observation, it enabled in a short time to collect documentation also on various cultural practices related to the use of heritage Piedmontese, that would have been impossible to observe in a fully spontaneous setting. Besides, the availability of professionally recorded sound and video enabled the researchers to produce a documentary film (Gorla and Gasparini 2023) in order to transfer part of the research outcomes to the community itself.

To conclude, the PILAR oral archive will constitute the first accessible dataset of spoken heritage Piedmontese, and therefore it will not only enable studies on multiple aspects of this contact situation, but it will also preserve a structured documentation corpus for this situation, as the effects of language shift towards Spanish become more pronounced in the future.

ACKNOWLEDGEMENT

The 2019 fieldwork session was funded by personal research funds of some fellow linguists at the University of Torino; for this reason we are deeply grateful to Massimo Cerruti, Davide Ricca and Riccardo Regis for their financial – and obviously scientific – support. The 2022 fieldwork was funded by the Early Career Grant of the Societas Linguistica Europaea, to which we are certainly obliged. We also acknowledge the invaluable contributions of the local communities who generously shared their time, knowledge, and hospitality, making this study feasible and enriching.

REFERENCES

- Aalberse, Suzanne, Ad Backus, and Pieter Muysken. 2019. *Heritage Languages: A Language Contact Approach*. Studies in Bilingualism, Vol. 58. Amsterdam: John Benjamins. <https://doi.org/10.1075/sibil.58>.
- Aikhenvald, Alexandra Y. 2007. “Linguistic Fieldwork: Setting the Scene.” *STUF – Language Typology and Universals*, 60(1): 3-11. <https://doi.org/10.1524/stuf.2007.60.1.3>.
- Auer, Peter. 1999. “From Codeswitching via Language Mixing to Fused Lects: Toward a Dynamic Typology of Bilingual Speech.” *International Journal of Bilingualism*, 3(4): 309-32. <https://doi.org/10.1177/13670069990030040101>.
- Austin, Peter K. 2006. “Data and Language Documentation.” In *Essentials of Language Documentation*, edited by Jost Gippert, Nikolaus P. Himmelmann, and Ulrike Mosel, 87-112. Boston: de Gruyter Mouton. <https://doi.org/10.1515/9783110197730.87>.
- Azar, Zeynep, Ad Backus, and Asli Özyürek. 2020. “Language Contact Does Not Drive Gesture Transfer: Heritage Speakers Maintain Language Specific Gesture Patterns in Each Language.” *Bilingualism: Language and Cognition*, 23(2): 414-28. <https://doi.org/10.1017/S136672891900018X>.
- Bagna, Carla. 2011. “America Latina.” In *Storia linguistica dell’emigrazione italiana nel mondo*, edited by Massimo Vedovelli, 305–58. Rome: Carocci.

- Benmamoun, Elabbas, Silvina Montrul, and Maria Polinsky. 2013. "Heritage Languages and Their Speakers: Opportunities and Challenges for Linguistics." *Theoretical Linguistics*, 39(3-4). <https://doi.org/10.1515/tl-2013-0009>.
- Berruto, Gaetano. 2005. "Dialect/Standard Convergence, Mixing, and Models of Language Contact: The Case of Italy." In *Dialect Change*, edited by Peter Auer, Frans Hinskens, and Paul Kerswill, 1st ed., 81-95. Cambridge: Cambridge University Press. <https://doi.org/10.1017/CBO9780511486623.005>.
- Berruto, Gaetano. 2017. *Sociolinguistica dell'italiano contemporaneo*. Nuova edizione, 2a edizione. Manuali universitari Linguistica 131. Rome: Carocci.
- Bettoni, Camilla, and John Gibbons. 1988. "Linguistic Purism and Language Shift: A Guise-Voice Study of the Italian Community in Sydney." *International Journal of the Sociology of Language*, 1988(72). <https://doi.org/10.1515/ijsl.1988.72.15>.
- Bickel, Balthasar, Bernard Comrie, and Martin Haspelmath. 2008. "Leipzig Glossing Rules: Conventions for Interlinear Morpheme-by-Morpheme Glosses." Online manuscript. Leipzig: Max Planck Institute for Evolutionary Anthropology. <https://www.eva.mpg.de/lingua/resources/glossing-rules.php>.
- Blommaert, Jan, and Jie Dong. 2010. *Ethnographic Fieldwork: A Beginner's Guide*. Bristol; Buffalo: Multilingual Matters.
- Brero, Camillo, and Remo Bertodatti. 1988. *Grammatica della lingua piemontese: Parola, vita, letteratura*. Turin: Edizione "Piemont/Europa."
- Buchstaller, Isabelle, and Ghada Khattab. 2014. "Population Samples." In *Research Methods in Linguistics*, edited by Robert J. Podesva and Devyani Sharma, 1st ed., 74-95. Cambridge: Cambridge University Press. <https://doi.org/10.1017/CBO9781139013734.006>.
- Carroll, Stephanie Russo, Desi Rodriguez-Lonebear, and Andrew Martinez. 2019. "Indigenous Data Governance: Strategies from United States Native Nations." *Data Science Journal*, 18(1): 31. <https://doi.org/10.5334/dsj-2019-031>.
- Clyne, Michael. 2003. *Dynamics of Language Contact: English and Immigrant Languages*. Cambridge: Cambridge University Press. <https://doi.org/10.1017/CBO9780511606526>.
- Dal Negro, Silvia. 2013. "Dealing with Bilingual Corpora: Parts of Speech Distribution and Bilingual Patterns." *Revue Française de Linguistique Appliquée*, 18(2): 15-28.
- De Mauro, Tullio. 1963. *Storia linguistica dell'Italia unita*. Biblioteca storica Laterza. Rome: Laterza.
- Devoto, Fernando. 2006. *Storia degli Italiani in Argentina*. Rome: Donzelli.
- Djenderedjian, Julio César. 2008. "La Colonización Agrícola En Argentina, 1850-1900: Problemas y Desafíos de Un Complejo Proceso de Cambio Productivo En Santa Fe y Entre Ríos." *América Latina En La Historia Económica*, 15(2): 127.
- Eckert, Penelope. 2000. *Linguistic Variation as Social Practice: The Linguistic Construction of Identity in Belten High*. Language in Society, 27. Malden: Blackwell Publishers.
- Fishman, Joshua. 1966. "Language Maintenance and Language Shift: The American Immigrant Case within a General Theoretical Perspective." *Sociologus*, 16(1): 19-39.
- Garcia, Guilherme D., and Natália Brambatti Guzzo. 2023. "A Corpus-Based Approach to Map Target Vowel Asymmetry in Brazilian Veneto Metaphony." *Italian Journal of Linguistics*, 35(1): 115-38. <https://doi.org/10.26346/1120-2726-205>.
- Gasparini, Fabio and Eugenio Gorla. In preparation. *Heritage languages and acts of identity: the case of Piedmontese migrants in Argentina*. Manuscript.
- Giolitto, Marco. 2000. "Pratiche linguistiche e rappresentazioni della comunità piemontese d'Argentina." *Éducation et Sociétés Plurilingues*, 9: 13-19.
- Giolitto, Marco. 2010. *La Communauté Piemontaise d'Argentine: Evolution, Fonction et Image Du Piemontais Dans La Pampa Gringa Argentine*. München: Martin Meidenbauer Verlagsbuchhandlung.

- Gorla, Eugenio. 2012. "Il dialetto nella comunicazione commerciale: Il caso torinese". *RID: Rivista italiana di dialettologia*, 36: 129-49.
- Gorla, Eugenio. 2015. "Il piemontese di Argentina. Considerazioni generali e analisi di un caso". *Rivista Italiana di Dialettologia, Lingue Dialetti e Società*, 39: 127-58.
- Gorla, Eugenio. 2023. "Il piemontese in Argentina. Aspetti linguistici ed etnografici". In *Confini nelle lingue e tra le lingue*, edited by Daniela Mereu and Silvia Dal Negro, 219-35. Milan: Officinaventuno.
- Gorla, Eugenio, and Fabio Gasparini. 2023 [Film]. *Pilar. Piedmontese Language in Argentina*. Italy.
- Grosjean, François. 2012. "Bilingual and Monolingual Language Modes." In *The Encyclopedia of Applied Linguistics*, edited by Carol A. Chapelle, 1st ed. Hoboken: Wiley. <https://doi.org/10.1002/9781405198431.wbeal0090>.
- Haig, Jeffrey, and Stefan Schnell. 2023. "Multi-CAST: Multilingual Corpus of Annotated Spoken Texts." Bamberg: University of Bamberg. <https://multicast.aspra.uni-bamberg.de>.
- Kilani, Mondher. 2000. *L'invention de l'autre: essais sur le discours anthropologique*. Repr. Anthropologie. Lausanne: Payot.
- Kung, Susan Smythe. 2020. "Data Archiving, Access, and Repatriation." In *The International Encyclopedia of Linguistic Anthropology*, edited by James Stanlaw, 1st ed., 1-4. Hoboken: Wiley. <https://doi.org/10.1002/9781118786093.iela0430>.
- Labov, William. 1984. "Field Methods of the Project on Linguistic Change and Variation." In *Language in Use*, edited by John Baugh and Joel Sherzer, 28-52. NJ: Prentice-Hall, Englewood Cliffs.
- Mereu, Daniela. 2019. *Il sardo parlato a Cagliari: Una ricerca sociofonetica*. Materiali Linguistici, 80. Milan: Franco Angeli.
- Mereu, Daniela. 2022. "Documentazione linguistica e studio della variazione sociolinguistica: Il caso delle varietà dialettali in via di estinzione." In *Per una pianificazione del plurilinguismo in Sardegna*, edited by Daniela Marzo, Simone Pisano, and Maurizio Viridis, 127-45. Cagliari: Condaghes.
- Mettouchi, Amina, and Christian Chanard. 2010. "From Fieldwork to Annotated Corpora: The CorpAfroAs Project." *Faits de Langues*, 35-36(2): 255-65. <https://doi.org/10.1163/19589514-035-036-02-900000011>.
- Milroy, James, and Lesley Milroy. 1997. "Network Structure and Linguistic Change." In *Sociolinguistics*, edited by Nikolas Coupland and Adam Jaworski, 199-211. London: Macmillan Education UK. https://doi.org/10.1007/978-1-349-25582-5_17.
- Miola, Emanuele, and Nicola Duberti. 2022. "Sulla testualità degli elaborati scritti del laboratorio di piemontese dell'Università di Torino." *Bollettino dell'Atlante Linguistico Italiano*, 46: 161-80.
- Moseley, Christopher, ed. 2007. *Encyclopedia of the World's Endangered Languages*. London: Routledge.
- Nagy, Naomi. 2020. "HLVC Transcriptions and Recordings." *Borealis*.
- Nagy, Naomi, and Miriam Meyerhoff. 2015. "Extending ELAN into Variationist Sociolinguistics." *Linguistics Vanguard*, 1(1): 271-81. <https://doi.org/10.1515/lingvan-2015-0012>.
- Nascimbene, Mario C. 1987. "Storia della collettività italiana in Argentina (1835-1965)." In *La popolazione di origine italiana in Argentina*, edited by Francis Korn, 209-504. Turin: Fondazione Agnelli.
- Olivier de Sardan, Jean-Pierre. 1995. "La Politique du terrain: Sur la production des données en anthropologie." *Enquête*, 1(October), 71-109. <https://doi.org/10.4000/enquete.263>.
- Polinsky, Maria. 2018. *Heritage Languages and Their Speakers*. Cambridge: Cambridge University Press.
- Polinsky, Maria, and Gregory Scontras. 2020. "Understanding Heritage Languages." *Bilingualism: Language and Cognition*, 23(1): 4-20. <https://doi.org/10.1017/S1366728919000245>.
- Raschieri, Guido. 2019. "The Museo del Paesaggio Sonoro (Riva presso Chieri, Turin)". *Etnografie sonore/Sound Ethnographies*, 2(1): 155-169.

- Regis, Riccardo. 2011. "Koinè dialettale, dialetto di Koinè, Processi di Koinizzazione." *Rivista Italiana di Dialettologia, Lingue Dialetti e Società*, 35: 7-36.
- Regis, Riccardo, and Matteo Rivoira. 2019. "L'anello che non tiene': Ai margini di un sistema ortografico." *Lengas*, 86(November). <https://doi.org/10.4000/lengas.3318>.
- Rießler, Michael, and Joshua Wilbur. 2017. "Documenting Endangered Oral Histories of the Arctic: A Proposed Symbiosis for Language Documentation and Oral History Research, Illustrated by Saami and Komi Examples." In *Oral History Meets Linguistics*, edited by Erich Kasten, Katja Martina Roller, and Joshua Karl Wilbur, 31-64. SEC Publications Exhibitions and Symposia. Fürstenberg: Kulturstiftung Sibirien.
- Rothman, Jason. 2009. "Understanding the Nature and Outcomes of Early Bilingualism: Romance Languages as Heritage Languages." *International Journal of Bilingualism*, 13(2): 155-63. <https://doi.org/10.1177/1367006909339814>.
- Schmid, Monika. 2011. *Language attrition*. Cambridge: Cambridge University Press.
- Sloetjes, Han, and Peter Wittenburg. 2008. "Annotation by Category: ELAN and ISO DCR." In *Proceedings of the Sixth International Conference on Language Resources and Evaluation (LREC'08)*, edited by Nicoletta Calzolari, Khalid Choukri, Bente Maegaard, Joseph Mariani, Jan Odijk, Stelios Piperidis, and Daniel Tapias. Marrakech: European Language Resources Association (ELRA).
- Sorace, Antonella. 2011. "Pinning down the concept of "interface" in bilingualism." *Linguistic Approaches to Bilingualism*, 1(1): 1-33. <https://doi.org/10.1075/lab.1.1.01sor>.
- Tagliamonte, Sali A. 2006. *Analysing Sociolinguistic Variation*. 1st ed. Cambridge: Cambridge University Press. <https://doi.org/10.1017/CBO9780511801624>.
- Tosco, Mauro, Emanuele Miola, and Nicola Duberti. 2023. *A Grammar of Piedmontese: A Minority Language of Northwest Italy*. Grammars and Sketches of the World's Languages. Leiden: Brill.
- Turchetta, Barbara. 2005. *Il mondo in italiano: varietà e usi internazionali della lingua*. 1. ed. Manuali Laterza, 220. Rome: Laterza.
- Vedovelli, Massimo. 2011. *Storia linguistica dell'emigrazione italiana nel mondo*. 1a ed. Studi Superiori, 641. Rome: Carocci.
- Villata, Bruno. 2009. *La lingua piemontese: fonologia, morfologia, sintassi, formazione delle parole*. Turin: Savej.