

# Linguistic Autobiographies. Towards the Creation of a Multilingual Resource Family

**Silvia Calamai**

Università di Siena, Italy  
silvia.calamai@unisi.it

**Rosalba Nodari**

Università di Siena, Italy  
rosalba.nodari@unisi.it

**Claudia Soria**

CNR-ILC, Italy  
claudia.soria@ilc.cnr.it

**Alessandro Carlucci**

Universitetet i Bergen, Norway  
alessandro.carlucci@uib.no

## Abstract

This paper describes a project aimed at creating a new resource family of multilingual and multimodal resources centered around the concept of a “Linguistics of the self”, i.e. personal reflections on the role of languages in shaping one’s identity. Language portrait silhouettes, drawing bilingualism, and linguistic autobiographies are different types of resources that share this common feature. We describe the resources and the criteria for their metadata annotation, focusing in particular on linguistic autobiographies, where the writer explicitly reflects on the relationship between him/herself and language. These genres are fruitfully used in different educational settings, and research has shown that they help to uncover the social, affective, and psychological dimensions of language learning. The potential of a multilingual and multimodal collection is discussed starting from data collected in Italy and Norway.

## 1 CLARIN Resource Families and Linguistic Diversity

The CLARIN Resource Family (Fišer et al. 2018) is a user-friendly overview per data type of available language resources in the CLARIN infrastructure aimed at the needs of researchers from digital humanities, social sciences, and human language technologies. Resource families are provided according to modality (spoken, multimodal, computer-mediated), genre (historical, academic, literary, newspaper, etc.), different languages, and intended use (reference, L2 learners). These groupings of corpora, lexical resources and other tools are meant to facilitate comparative research: for each resource family, a brief description is provided followed by a list of the resources belonging to the family, together with the most important metadata (name, size, annotation, license, language, description and availability). Thus, resource families provide a curated view of the available CLARIN resources. Over the years, this has proven to be a highly visible initiative appreciated by a broad spectrum of CLARIN users (Leonardič and Fišer 2020) and it therefore deserves to be maintained and enlarged. In this paper, we introduce a new genre that comprises various types of productions (texts and mixed text and drawings) sharing the common feature of containing personal reflections on the role of language in shaping one’s identity. We argue that a resource family devoted to this genre could be useful for both first and second language research and pedagogy, as well as for a better understanding of the linguistic landscape in European schools and universities, as it is shown in the following section.

---

This work is licensed under a Creative Commons Attribution 4.0 International License. License details: <http://creativecommons.org/licenses/by/4.0/>

Silvia Calamai, Rosalba Nodari, Claudia Soria and Alessandro Carlucci 2023. Linguistic Autobiographies. Towards the Creation of a Multilingual Resource Family. *Selected papers from the CLARIN Annual Conference 2022*. Ed. by Tomaž Erjavec and Maria Eskevich. Linköping Electronic Conference Proceedings 198, pp. 23–32. DOI: <https://doi.org/10.3384/ecp198003>

## 2 Self-Reflective Practices for Shaping Europe’s Multilingual Landscape

Pedagogical research has shed light on the importance of reflective practices for both teachers and students. As far as the former are concerned, reflective practices can help in reshaping their knowledge and their teaching practices (Farrell 2022). Students, too, can be encouraged to engage in self-reflection to help them focusing on the emotional and cognitive components of their attitudes toward the subjects they are studying. In language education there are several age-appropriate tasks that are typically employed in assisting students to reflect on their multilingual selves in order to develop meta-sociolinguistic competence. Among the most popular methods we can list: i) language portrait silhouettes; ii) drawing bilingualism; iii) linguistic autobiographies. Silhouettes and drawing bilingualism are discussed in §2.1, while a separate paragraph is devoted to linguistic autobiographies (§2.2).

### 2.1 Language Portrait Silhouette and Drawing Bilingualism

Language portrait silhouette (LPS) and drawing bilingualism (DB) are two multimodal self-reflective tasks that are suitable for early childhood and primary education because they do not require particular abilities in writing. However, LPS and DB can also be employed in other contexts such as secondary schools and universities, especially in multilingual classes where different levels of linguistic competence can be found. In LPS tasks, respondents are given a body silhouette (see fig. 1) and are asked to paint it, choosing different colors and body part for each language they want to mention (Busch 2010). They can be adopted also at a policy level: in the Autonomous multilingual Province of Bozen\Bolzano in South Tyrol, LPS is an accredited part of the European Portfolio of languages. Respondents are asked to fill the silhouette with different colors, using larger spots for the most frequently used languages, and to discuss their LPS with schoolmates (Provincia Autonoma di Bolzano 2004).



Figure 1. A template for LPS

Similarly, in the drawing bilingualism task (Favaro 2013), children can freely draw themselves without having a silhouette to fill in. Children have to imagine themselves as plurilingual individuals and choose which representation of themselves can better capture their plurilingual mind. Figure 2 shows a specimen collected in an Italian secondary school: according to the image, French is the language of the ‘head’ because it has been studied at school, while Italian is the language of the ‘heart’; German is the language of the ‘hand’ since the student has to struggle with its grammar and finally English is the language of the ‘legs’, because it opens up doors in space and time.

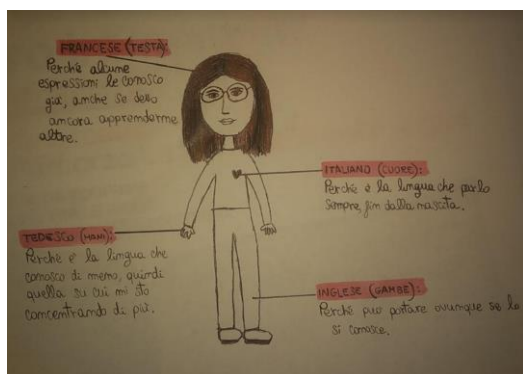


Figure 2. A drawing bilingualism example from the University of Siena corpus

These two tasks are commonly used across Europe (cf. for the Italian setting Manconi 2019; Biagioli 2021) as they do not require specific training; LPS can also easily be downloaded from different websites (see for example <https://cluss.unistrasi.it/public/articoli/157/Silhouette%20linguistica.pdf>). In these two tasks, the linguistic repertoire is truly interpreted as embodied cognition. With LPS and DB, researchers can look at linguistic repertoires from a multimodal perspective (Kress & van Leeuwen 2006): they can analyse the relation between language, colours and cultural tropes, the affiliation and emotional component of language learning, as well as the multimodal representation of language using different symbols. For teachers, LPS and DB could be used as an ice-breaker activity during the first days of school.

## 2.2 Linguistic Autobiographies

Unlike the two tasks explained in 2.1, linguistic autobiographies constitute a non-fictional genre where the writer explicitly reflects on the relationship between him/herself and language. Different labels have been used for describing this genre other than linguistic autobiographies, such as language memoirs (Pavlenko 2001, 2007; Kramsch 2004) or language stories (Gohard-Radenkovic, Rachedi 2009). In this self-reflective writing practice, language becomes the overarching organising principle for re-tracing salient moments in the writer's life. The idea behind this genre is that the acquisition and the interaction of different languages can be seen as the acquisition of selfhood (Ramsdell 2004). Linguistic autobiographies can be considered as both a research and a pedagogical tool. They are used by professors and teachers in secondary schools and university classrooms to help students in developing their metalinguistic and metapragmatic abilities; within superdiverse multilingual classrooms, linguistic autobiographies allow students to narrate their multilingual selves and they help make their languages more visible. Linguistic autobiographies can also help linguists in gaining access to language ideologies and attitudes towards language varieties: in particular, these narratives can help understand how ideologies about languages can have an impact on the linguistic behaviour of speakers.

Linguistic autobiographies are highly versatile in that they can easily be collected without requiring specific skills or academic knowledge. Several templates are available to facilitate the production of linguistic autobiographies, which offer some suggestions to think about languages and self-reflect on key points of the writers' own life (cf. Canobbio 2006; D'Agostino 2007; Luppi, Thüne 2022)<sup>1</sup>.

For example, one possible template requires mentioning:

- 1) Family members and personal data (place of birth, eventual relocations, etc.);
- 2) Family linguistic background: L1 of the grandparents, L1 of the parents;
- 3) Family linguistic situation: parents and grandparents' linguistic preferences (they speak which language to whom and when); which languages are used for ordinary communication between family members (with children, with the rest of the family); family choices in linguistic education (which

<sup>1</sup> The Council of Europe provides its own template for compiling language biographies of users' plurilingual profiles. The scheme can be found here: <https://rm.coe.int/CoERMPublicCommonSearchServices/DisplayDCTMContent?documentId=09000016804932c5>)

language is taught to children); which language is spoken at home; which other varieties are used at home, and for what purposes (communicative, expressive, affective needs, identity stances, etc.);

4) Family and school attitudes and behaviours: repression of non-standard varieties; are any specific varieties preferred to other varieties? Are there any disfavoured languages? Are there any forbidden languages used in secret between friends or family members?

5) Meeting with linguistic diversity (holiday trips to different regions, community of practices, peer groups, school environment, extended family etc.). Formal and informal language learning (foreign language schools, friends from abroad etc.); are different languages used with different groups of people? Are non-standard varieties used for performing specific identities? Etc.

6) Personal evaluation of language learning in and out of school; ability to perceive different linguistic varieties, social evaluation of accents, stigma and stereotypes towards accents, varieties, languages, etc.

Here is an instance of a linguistic autobiography produced by a second-generation Italian student of Chinese origin, followed by its translation into English.

*Io ho iniziato a studiare l'italiano quando andavo all'asilo e l'ho imparato insieme al cinese.*

*Quando andavo all'asilo soprattutto io parlavo italiano perché c'erano solo 2 o 3 compagni cinesi. A casa io parlo cinese con i miei genitori e alcune volte italiano o inglese con mia sorella, i miei genitori parlano sempre il dialetto e vuole anch' io parlare il dialetto a casa e mi dicono non devi mai dimenticare il dialetto perché questo è il nostro radice. Io vorrei migliorare il mio italiano perché questa lingua l'ho iniziato a studiare a l'asilo e per un altro motivo: io abito in Italia e userò più italiano di cinese. Io vorrei migliorare francese e tedesco più d'italiano perché queste due lingue forse le userò per lavorare<sup>2</sup>.*

*I started studying Italian when I was in kindergarten and I learned it together with Chinese.*

*In kindergarten I used to speak Italian because there were only 2 or 3 Chinese classmates. At home I speak Chinese with my parents and sometimes Italian or English with my sister, my parents always speak the dialect and I want to speak the dialect too at home and they tell me you must never forget the dialect because this is our root. I would like to improve my Italian because I started studying this language in kindergarten and for another reason: I live in Italy and I will use more Italian than Chinese. I would like to improve French and German more than Italian because I will perhaps use these two languages for work.*

The template (and the terminology used) can be adapted depending on the age, educational background and other characteristics of those involved. In any case, linguistic autobiographies are deeply personal texts that outline the writers' own thoughts and feelings, giving them the possibility for spontaneous expression.

### **2.3 The Societal Potential of Self-Reflective Practices in Education**

All these tasks are fruitfully used in different educational settings, and research has shown that these tools help to uncover the social, affective, and psychological dimensions of language learning (Franceschini, Miecnikowski 2004; Groppaldi 2010; Cavagnoli 2020; Salvadori, Blondeau, Polimeni 2020). Language biographies are also an important element of the European Language Portfolio, as they encourage learners to write information on linguistic and cultural experiences gained in and outside formal educational contexts (Council of Europe 2001). They permit students to develop an awareness of cultural and linguistic diversity, and to learn about the social value of languages. In superdiverse settings, linguistic autobiographies help students in understanding mechanisms of stereotyping and linguistic discrimination and are considered an empowering tool. Teachers can also gain access to the students' linguistic learning process, and they can discover students' learning practices and reasons for studying languages, as well as their needs and expectations. For policy makers and stakeholders, self-reflective practices can help in understanding students' motivation for language learning, thus allow-

---

<sup>2</sup> The linguistic autobiography reported in the paper is directly quoted from the original text, without any editing from the authors.

ing them to develop specific school curricula addressing students' communicative needs. For those who are interested in the sociology of languages, these data can also provide unique information about informal language-learning opportunities and the different values that speakers attach to different types of multilingualism.

### 3 Towards a CLARIN Resource Family for Self-Reflective Practices

Language portrait silhouettes, drawing bilingualism, and linguistic autobiographies constitute different genres of great scholarly significance and with high potential of impacting education. The genres above can all be considered self-reflective practices, with language(s) being the overarching characterizing principle. For this reason, we believe that a multilingual and multimodal resource family that collects LPS, DB and linguistic autobiographies produced by L1 and L2 speakers of different languages in different countries is an important contribution to the scholarly community gathered around CLARIN. This new resource family will thus deal with peculiar genres with specific formal and content characteristics, and it is also open to hosting other genres and text types, such as oral interviews and personal narratives. From our point of view, it is desirable that all textual and non-textual products sharing the trait of personal metalinguistic reflection can be collected, and that their collection is carried out in ways that allow scholars to easily compare them and practitioners to use them immediately. The new resource family, called "Linguistics of the Self", will thus offer the possibility for scholars to access all these kinds of data, and to contribute with their own data. In this section we report about the principles and criteria that we will follow in building the resources and making them available as a CLARIN resource family. We will provide examples from the corpus of linguistic autobiographies, which so far represents the largest component of the family.

#### 3.1 Adding Linguistic Autobiographies to CLARIN

The University of Siena corpus of linguistic autobiographies will be the core of this new multilingual and multimodal resource family. At present, this resource consists of about 300 linguistic autobiographies collected during university courses in linguistics (educational linguistics and sociolinguistics)<sup>3</sup>, about 50 autobiographies written by secondary school students, and about 40 produced by secondary school teachers. An example is available in Figure 3. In addition to the core, the resource will also include autobiographies collected in Norway. A small pilot collection has been carried out at the University of Bergen, in Norway, and has produced texts in different languages, including English and Spanish (see Fig. 4). The collected texts will be digitized, and data processing will be carried out for anonymization, automatic linguistic analysis with the Profiling-UD tool (see Brunato et al. 2020) and metadata description, on the basis of a shared procedure. A strong requirement that constrains how the corpus is going to be structured is that it should be easy to update, so that the corpus can be integrated with new linguistic autobiographies. Indeed, linguistic autobiographies, as well as LPS and DB, are commonly produced and collected by others, in addition to the original creators of the resource. Possible potential contributors are schoolteachers who collect this type of data according to the procedure that we suggest, but also other university professors. Linguistic autobiographies are currently being collected as part of teaching activities at several Italian universities (University of Pavia, Macerata, Roma Tre, among others), especially in educational linguistics and sociolinguistics courses. Even if this resource is not particularly large in number, compared to other CLARIN resources, one of its main strengths is its transversality, in that it can be used in any higher education degree. This will allow the resource to grow steadily each year.

In this respect, we have identified the structure of the ParlaMint project (Erjavec et al. 2022), where each file of the corpus represents a single transcript session, as the most appropriate for our needs.

---

<sup>3</sup> At the University of Siena, approximately 60 autobiographies per academic year are collected. In particular, the collection is an integral part of the course in Sociolinguistics: autobiographies are collected at the beginning and the end of the course and function as a tool both to introduce some key notions from sociolinguistics and to allow individual reflection by students at the end of the course, once analytical skills and terminology have been acquired.



This will allow the corpus to grow constantly, with new autobiographies being periodically added to the original collection.

SEMPRE RIMASTA A VIVERE IN QUELLA ZONA  
 SPOSTATA SOLO RESIDENZA TRA DUE COMUNI  
 LIMINIAPI -  
 - PADRE NATO NELLO STESSO COMUNE  
 - MADRE NATA IN ALTRO COMUNE, MA A SOLI 15 KM.  
 ANCHE LOO ~~REMO~~ VISSUTO NELLA STESSA VALLATA  
 IL CASENTINO (ALTO CASENTINO)  
 - IO NONNE VENIVANO DALLA STESSA AREA GEOGRAFICA  
 MA IN GIOVENTÙ SI ERANO TRASFERITI PER  
 LAVORO A FIRENZE O PROVINCIA  
 - HO VISSUTO CON UNA NONNA CHE PARLAVA "OSCANO"  
~~OSCANO~~ CON UNA CADENZA INCARCATI DI  
 DIALETTO NEMPOLESE / COME ANA /  
 \* LEL USAVA PAROLE DIVERSE DAGLI ALTRI COMPONENTI  
 DELLA FAMIGLIA (E NE ANDAVA ORGOGLIOSA).  
 - HO STUDIATO E HO SONO LAUREATA IN LINGUE STRANIERE  
 (INGLESE - SPAGNOLO) HO SOGGIORNATO ALL'ESTERO  
 - QUANDO ~~ARRIVAI~~ AD ABBEZZO A SCUOLA GLI AMICI MI  
 CHIEDEVANO SE VENIVO DA FIRENZE "PARLI FERENTINO  
 ?? A ME NON SEMBRAVA AFFATTO !!  
 - QUANDO OSPITAVO AMICHE STRANIERE A CASA  
 MIA NONNA PARLAVA IN ITALIANO, MA ALZANDO IL  
 TONO DICEVA "COME SE FOSSEVO STATE SORDE" !!

Figure 3. A hand-written linguistic autobiography from the University of Siena corpus

**Autobiografía lingüística**

En mis 20 años de vida he aprendido más o menos los siguientes idiomas: el inglés casi como si mi existencia dependiera de ello, el alemán por obligación, el italiano por gusto y el español, mi lengua materna, quizás nunca lo termine de aprender.

Desde el primer año de colegio empecé a escuchar el inglés. En mi casa nadie hablaba un idioma diferente al español así que aprendí la pronunciación inventada por mis papás tratando de leer el "Workbook", más adelante mis profesores de bachillerato sufrieron para corregir esa pronunciación a la que estaba acostumbrada. Mis papás se niegan completamente a que viva en Colombia e hicieron todo lo posible para que yo aprendiera inglés, por eso siempre estuve en colegios bilingües donde enseñaban inglés británico e insistían en que escucháramos música, viéramos películas y leyéramos todo en inglés. Dado que los libros que me gustaban se publicaban primero en inglés, aprovechaba la excusa de practicar para que me los compraran o para que me dejaran pasar más tiempo viendo televisión, y llegó un punto en el que mi entorno estaba tan plagado del idioma que se me hacía raro cuando alguien me contaba que no lo hablaba.

Al entrar a la universidad no tuve que cursar las clases obligatorias de inglés porque ya tenía un buen nivel, para abrir la posibilidad de irme a otro país empecé a estudiar alemán. Me ilusionaba la idea de leer los textos teóricos de mi carrera en su idioma original y entender canciones, pero la emoción se fue tan rápido como llegó, las clases pasaron a la virtualidad por la pandemia y me costaba mucho concentrarme o siquiera entusiasarme por ver la ventana de zoom presentar un libro de ejercicios mientras la voz del profesor iba y venía. Tan pronto como se acabó el curso dejé de practicar y olvidé casi completamente todo lo aprendido, cuando se dieron cuenta mis papás se enojaron diciendo que era desperdiciar una oportunidad muy grande y que debería seguir estudiando alemán a pesar de que no me sintiera verdaderamente motivada ni quisiera hacerlo.

Aprendí italiano por una canción que me gustaba mucho y más tarde descubrí que la letra está en napolitano. Al inicio entraba a las clases con miedo de que se repitiera la experiencia del alemán, de hecho ni siquiera le conté a mis papás que me inscribí al curso. Sin embargo fue totalmente diferente a lo que esperaba y mis dudas se transformaron en entusiasmo y pasión por el idioma, gracias a la clase de italiano conocí personas maravillosas con las que me encanta hablar y practicar a pesar de que hace tiempo concluimos el curso. Además, la profesora me hizo entender que un idioma no es la traducción del otro, y nos pedía que pensáramos directamente en italiano y nos expresáramos con las herramientas que teníamos aunque al principio fuera difícil. Así fue como sentí que conecté más con el idioma y me interesé por la historia y el entorno que lo hicieron ser como es, incluso empecé a interesarme esos aspectos de mi propio idioma.

El español ha sido una cuestión compleja porque he vivido en tantas ciudades diferentes de Colombia que mi acento es irreconocible, aunque preferiría que se notara el acento de la ciudad donde nací inconscientemente adopté las formas de hablar del sitio donde estoy, eso le encanta a mi mamá porque dice que en mi tierra la gente es muy corroncha y no saben hablar. Ella odia todo lo que no sea español académico, al punto de que considera a algunos géneros musicales, como el vallenato, inferiores. Después de varias peleas con ella por mi interés en la tradición cantada del español, pude disfrutar de esa historia tanto amo y que hace parte de mí. Ahora, junto a mi deseo de viajar y vivir en diferentes países, está también el deseo de atesorar y llevar conmigo la belleza de mi idioma.

Figure 4. A linguistic autobiography collected at the University of Bergen

This new CLARIN resource is designed to be comparable with other types of resources, such as LPS and DB, that share some of the features of linguistic autobiographies. Additionally, linguistic autobiographies could share some features with already existing or future resources, such as focus groups on language attitudes, qualitative interviews and oral narratives in which the interviewee describes his/her relationship with his/her mother tongue(s), language acquisition, etc.

This approach derives from the fact of having devised our resource as a member of a resource family from the outset, and it has important consequences for the choice of metadata (see §3.2). Furthermore, it is important to us that the new resource can easily be compared with similar resources that are

already part of CLARIN collections. For example, by searching the corpus according to speakers' L1, it will be possible to compare the biographies of speakers with the same L1 across different countries, educational settings, etc. This possibility may be used for various purposes, e.g. for a comparative analysis of the effects of language contact in different settings and with different L2s. Therefore, we foresee that rich metadata will be added to the corpus of linguistic autobiographies: all the valuable sociodemographic data, the language used and all the languages and varieties mentioned in the autobiographies will have to be made explicit. This will greatly enhance the usability and comparability of the resource (see below).

### **3.2 Criteria for Metadata Selection for a Multilingual and Multimodal Resource Family on the “Linguistics of the Self”**

The choice of metadata is a crucial issue in building resource families (Leonardič and Fišer 2020), and it is therefore of paramount importance. When a resource family is created from already available resources, it is often the case that the curated resources have different depths and breadths of metadata description, which in turn has consequences for their final usability. In our case, where the individual resources are going to be created and described having in mind their role as members of a wider resource family, it is of utmost importance that all the different genres are described in such a way that their collection under a resource family is straightforward and allows for their maximum comparability. Here we introduce some of the criteria that will guide us in the selection and adoption of metadata for the resources composing the family.

We believe that a metadata description that is oriented to making a resource fit in a resource family must satisfy two interacting main requirements: a) exhaustiveness and b) comparability.

Exhaustive metadata description is important not only for the sake of describing any given individual resource accurately, but also in order to maximise its traceability and the possibility for it to be discovered and included in future collections. All aspects belonging to the “linguistics of the self” will need to be highlighted accordingly: the self-descriptive aspect, the modality of the task (spontaneous drawing, written autobiography, oral interview, etc.), all the languages that are mentioned and the language in which a given text is produced, whether the language is an L1 or L2 for the author, in addition to more customary socio-demographic factors such as sex, age and provenance. Metadata description will also take into account other important aspects, namely school grade (resources collected in primary or secondary schools, or universities) and metalinguistic competence (autobiographies collected before and after linguistic courses).

Making this information explicit will make it possible, for instance, for a scholar interested in studying the attitudes of students towards English as a second language to select all autobiographies written in different languages and mentioning English as a second language.

A TEI Header will be developed to encode this information. Each autobiography will then be annotated according to the TEI header scheme. The TEI Header scheme will be made available to download, in order to facilitate the annotation of similar resources already owned by other scholars.

In order to ensure the highest possible degree of comparability with other resources, either already present in the CLARIN ecosystem or to be added in the future, metadata description for the new resources will have to take into account the metadata sets used for describing resources that are partially overlapping for content and/or genre with linguistic autobiographies. From a first analysis of the VLO we have identified oral interviews, general autobiographies and personal narratives as the most similar genres already represented in CLARIN collections. The metadata used for describing these resources will be studied and analyzed to identify which elements of their description can be included in metadata description for linguistic autobiographies and linguistic silhouettes.

Since these are very peculiar genres not currently available in CLARIN, none of the available profiles for describing the resource is entirely suitable. The VLO repository, for instance, already offers a selection of linguistic autobiographies collected in two (non-exclusively) Italian-speaking settings, namely Language Biographies from South Tyrol and from Basel. However, these two corpora consist of audio interviews, together with their transcriptions. Written linguistic autobiographies, on the other hand, show some peculiar features such as a) written modality vs. mainly oral modality, and b) their strong emphasis on the linguistic component: language is the key around which the narrative is built

and articulated, and the recollection of one's life follows a linguistic path, while, in general, oral narratives focus on the main events in the lives of speakers.

### 3.3 Working Agenda

Available CMDI profiles will be studied and, if necessary, an ad hoc profile will be created to adequately describe the new resources. In addition to the metadata definition, legal and ethical issues will be addressed for personal data protection, in order to clarify i) the legal basis for future fieldwork collection, especially in the case of children/minors; ii) the procedure for obtaining consent for texts already collected – also before the introduction of GDPR; iii) anonymization strategies, especially in cases where explicit consent is absent.

In addition to the resource, support material for the collection of new resources will then be made available: three scripts (in English, Italian, and Norwegian) to guide the researcher/teacher in the elicitation of the texts (linguistic autobiographies, LPS, DB); guidelines for anonymizing the texts; and a template for their representation in TEI format. The script for the collection of autobiographies will only contain minimal information, such as the one mentioned in 2.2. For LPS, the silhouette will be provided, together with general instructions in three languages. The proposed family will thus provide a consistent and shared framework for the collection and curation of texts and multimodal products that are remarkably peculiar and therefore highly subject to the risks incurred by any data-scarce collection, such as inconsistency of structure, format, and level of analysis. The common template for metadata description of the different resources belonging to the family and the protocol for elicitation of the interviews will prove useful for ensuring coherence and comparability across different initiatives and will greatly help researchers and teachers by providing them with an established format.

Alongside this investigation, we will implement specific activities to promote linguistic autobiographies as an educational tool, also in order to discover any uses already in place but unknown to us. We are confident that in some of the countries involved in the CLARIN network, linguistic autobiographies are already used in school and university settings. In this respect, ILC-CNR and the University of Siena have now the advantage of being involved in the CIRCE Erasmus+ project (2023-25) aimed at studying accent discrimination in education. The school environment is a hotspot for investigating this issue: students are exposed to different accents and form and reinforce their attitudes and beliefs towards them also on the basis of peer pressure. Teachers are also confronted daily with regional and non-native accents of the national language, and may unconsciously succumb to prejudice and negative evaluations of non-standard varieties. In the project, a specific task will be devoted to training school teachers in using linguistic autobiographies in their classes. These autobiographies will then be FAIRified and will be added to the CLARIN resource family. This will permit to collect at least 200 multilingual autobiographies in different school settings and in different countries (in a digital format).

We thus wish to help uncover any already existing material and encourage the production of new material. A multilingual collection of such written material will indeed offer an invaluable picture from several perspectives. Firstly, it can be used as teaching material, from school classes of any grade to university courses, in order to raise awareness of heritage languages, accentism and glottophobia. Secondly, it can help teachers to better understand the most used and known languages in their classrooms. And thirdly, this collection represents a useful tool to verify – among pupils, students and teachers – the pervasiveness of the concept of linguistic error and deviation in describing linguistic repertoires.

Comparable corpora of linguistic autobiographies will also provide valuable quantitative and qualitative data to researchers interested in a variety of topics – such as language attitudes, language and migration, multilingualism and language contact. Finally, this new resource can help policymakers in designing linguistic policies that are more consistent with the different linguistic landscapes existing in different European schools and universities.



## Acknowledgments

For Siena University and CNR authors, the project described in the text is partially co-funded by the European Union (CIRCE project (Erasmus+ AGREEMENT NUMBER: 2022-1-IT02-KA220-SCH-000087602).

## References

- Biagioli, R. 2021. Promuovere un'educazione linguistica interculturale. *Educazione interculturale*, 19(2): 33-45.
- Brunato D., Cimino A., Dell'Orletta F., Montemagni S., Venturi G. 2020. Profiling-UD: a Tool for Linguistic Profiling of Texts. In *Proceedings of 12th Edition of International Conference on Language Resources and Evaluation (LREC 2020)*, 11-16 May, 2020, Marseille, France.
- Busch, B. 2010. School language profiles. Valorizing linguistic resources in heteroglossic situations in South Africa. *Language and education* 24(4): 283-294.
- Canobbio, S. 2006. Dialecto dei giovani e politiche linguistiche delle famiglie, appunti dal Piemonte. In: Marcato, G.: *Giovani, lingue e dialetti, Atti del Convegno*, Sappada - Plodn, 29 June - 3 July 2005. Unipress, Padova: 239-244.
- Cavagnoli, S. 2020. Diventare insegnanti: l'importanza della riflessione di futuri/e docenti sul proprio percorso di apprendimento linguistico. *Italiano LinguaDue* 12(2): 146-160.
- Council of Europe 2001. *Common European Framework of Reference for Languages: learning, teaching, assessment*. Cambridge University Press, Cambridge.
- D'Agostino, M. 2007. *Sociolinguistica dell'Italia contemporanea*. Il Mulino, Bologna.
- Erjavec, T., Ogrodniczuk, M., Osenova, P. et al. 2022. The ParlaMint corpora of parliamentary proceedings. *Lang Resources & Evaluation*. <https://doi.org/10.1007/s10579-021-09574-0>
- Farrell, T. S. 2022. *Reflective practice in language teaching*. Cambridge University Press, Cambridge.
- Favaro, G. 2013. Il bilinguismo disegnato. *Italiano LinguaDue*, 5(1): 114-127.
- Fišer, D., Lenardič, J., and T. Erjavec. 2018. CLARIN's Key Resource Families. In *Proceedings of the Eleventh International Conference on Language Resources and Evaluation (LREC 2018)*: 1320–1325.
- Franceschini, R., and Mieczkowski, J. (eds) 2004. *Leben mit mehreren Sprachen. Vivre avec plusieurs langues. Sprachbiographien. Biographies langagières*. Lang, Bern.
- Gohard-Radenkovic, A., and Rachedi, L. 2009. *Récits de vie, récits de langues et mobilités*. L'Harmattan. Paris.
- Groppaldi, A. 2010. L'autobiografia linguistica: strumento per una moderna didattica dell'italiano L2-LS. *Italiano LinguaDue*, 2(1): 89-103.
- Kramsch, C. 2004. The multilingual experience: insights from language memoirs. *TRANSIT*, 1(1). <http://dx.doi.org/10.5070/T711009691>
- Kress, G., & van Leeuwen, T. 2006. *Reading images: The grammar of visual design*. Routledge, London.
- Leonardič, J. and D. Fišer. 2020. Extending the CLARIN Resource and Tool Families. In: Navarretta, C. & Eskevich, M.: *Proceedings of CLARIN Annual Conference 2020*, 05-07 October 2020, Virtual Edition: 1–5.
- Luppi, R, and Thüne, E.-M. (eds) 2022. *Biografie linguistiche. Esempi di linguistica applicata*, Centro di Studi Linguistico-Culturali (CeSLiC), Bologna.
- Manconi, T. 2019. La “silhouette des langues”: dessiner et représenter les identités plurielles. *Annali online della Didattica e della Formazione Docente*, 11(17): 107-123.
- Pavlenko, A. 2001. Language learning memoirs as gendered genre. *Applied Linguistics*, 22(2): 213-240.
- Pavlenko, A. 2007. Autobiographic narratives as data in applied linguistics. *Applied Linguistics* 28(2): 163–188.
- Provincia autonoma di Bolzano 2004. *Portfolio europeo delle lingue, per alunne e alunni dai 9 agli 11 anni*. Direzione Istruzione e Formazione tedesca, Bolzano.
- Ramsdell, L. 2004. Language and Identity Politics: The Linguistic Autobiographies of Latinos in the United States. *Journal of Modern Literature* 28(1): 166-176.

Salvadori, E.; Blondeau, N.; Polimeni, G. (eds.) 2020. Lingue maestre. autobiografia linguistica e autoformazione dei docenti di L1 e L2. *Italiano LinguaDue*, 12(2).