ETHNORÊMA

Lingue, popoli e culture

Rivista annuale dell'associazione Ethnorêma

ANNO XX - N. 20 (2024)

www.ethnorema.it

Ethnorêma, dal greco *ethnos* 'popolo, etnia' e *rhêma* 'ciò che è detto, parola, espressione', ma anche 'cosa, oggetto, evento'. Nella linguistica pragmatica *rema* sta ad indicare la parte di una frase che aggiunge ulteriore informazione a quello che è stato già comunicato (il *tema*).

Ethnorêma, from the Greek words *ethnos* 'people, ethnicity' and *rhêma* 'what is said, word, expression', but also 'thing, object, event'. In linguistics, *rheme* indicates the part of a sentence that adds further information about an entity or a situation that has already been mentioned (the *theme*).

Ethnorêma è la rivista dell'omonima associazione. L'associazione senza scopo di lucro Ethnorêma intende promuovere attività di studio e ricerca nel campo linguistico, letterario, etnografico, antropologico, storico e in tutti quei settori che hanno a che fare, in qualche modo, con le lingue e le culture del mondo.

Ethnorêma is the journal of the association of the same name. The Italian non-profit association Ethnorêma works to promote study and research activities in the fields of linguistics, literary enquiry, ethnography, anthropology, history and in all those sectors which have to do, in some way, with the languages and cultures of the world.

Valutazione ANVUR: Classe A (Area 14/B2)

La rivista è disponibile gratuitamente in rete, scaricabile dal sito *www.ethnorema.it.*

The journal can be viewed and downloaded free of charge at *www.ethnorema.it.*

ISSN 1826-8803

Registrazione n. 1/05 del 28/10/2005 presso il Tribunale di Tortona

ETHNORÊMA

Lingue, popoli e culture

Rivista annuale dell'associazione Ethnorêma

ANNO XX - N. 20 (2024)

www.ethnorema.it

INDICE / CONTENTS

Editoriale

Moreno Vergari (Direttore di *Ethnorêma*)

Questo numero speciale di *Ethnorêma* vuole celebrare due piccoli, ma per noi importanti traguardi: **la decima edizione di una serie di seminari sulla Documentazione Linguistica** e il **ventesimo numero della nostra rivista**. Per farlo abbiamo pensato di unire in qualche modo questi due eventi. Questo numero di *Ethnorêma* è infatti dedicato a vari aspetti della documentazione linguistica e viene presentato ufficialmente in occasione della Scuola Estiva di Documentazione Linguistica che si tiene quest'anno presso la Libera Università di Bolzano dal 15 al 19 luglio (https://ldsummerschool2024.wixsite.com/ldss2024).

Ethnorêma nasce nel 2003 come un'**associazione di volontariato** (oggi ODV iscritta al RUNTS) che vuole occuparsi di lingue e culture del mondo con un approccio multidisciplinare, oltre ad avere anche una particolare attenzione a programmi di cooperazione allo sviluppo. Per queste ed altre iniziative e progetti rimandiamo alle pagine dedicate sul sito di Ethnorema.

Il **primo numero della rivista dell'associazione viene pubblicato nel 2005**. Da subito si è fatta una scelta precisa: la rivista doveva essere scaricabile gratuitamente per poter garantirne la libera e massima diffusione. Vediamo con piacere che questa è la scelta seguita da sempre più pubblicazioni scientifiche, ma vent'anni fa non era così.
Da allora circa 100 diversi autori hanno contributo a far crescere la nostra rivista, facendola conoscere e a farle trovare un suo spazio nel mondo accademico e non solo.

L'idea di una serie di **seminari dedicati alla documentazione linguistica** nasce dopo lo spostamento della sede legale dell'Associazione, e quindi della rivista, dal Piemonte a Bolzano e da un primo contatto con l'allora Centro di Competenza Lingue della Libera Università di Bolzano (LUB). Dal 2010, anno della prima edizione, la collaborazione

tra Ethnorêma e la LUB ha permesso la realizzazione di altri sei seminari nella sede della LUB. È proprio durante la settima edizione che nasce la collaborazione con docenti della Università della Tuscia di Viterbo e dell'Università di Napoli L'Orientale. Si decide quindi per un'alternanza annuale delle sedi tra Bolzano, Viterbo, dove si è tenuta l'ottava edizione nel 2022, e Napoli, sede nel 2023 della nona edizione.
Dal 2022 il ciclo di seminari assume il nome ufficiale di *Scuola Estiva di Documentazione Linguistica* (in inglese *Language Documentation Summer School*) .
Nel 2023 la collaborazione tra i quattro enti coinvolti viene sancita con un Protocollo d'intesa che, all'articolo 1, ne illustra gli obiettivi:

> ***Articolo 1 – Obiettivi della collaborazione***
> *Le Parti, nell'ambito delle rispettive funzioni, competenze e ruoli istituzionali, si impegnano a collaborare per la realizzazione di un'azione congiunta finalizzata a*
> *a. approfondire la conoscenza e l'innovazione relative alle tecniche di raccolta dati sul campo, analisi e catalogazione di dati linguistici provenienti da lingue minori, lingue minacciate, dialetti e gerghi;*
> *b. approfondire la conoscenza e l'innovazione relative alla rivitalizzazione linguistica di lingue minacciate;*
> *c. svolgere una formazione didattica, eventualmente connessa con stage e tirocini, sugli strumenti di raccolta, analisi e archiviazione di dati linguistici;*
> *d. condividere, confrontare, diffondere tali conoscenze in occasione di convegni, seminari e scuole estive rivolte a pubblico italiano e internazionale;*
> *e. stabilire e sviluppare una proficua collaborazione con lo scopo di promuovere tra loro la collaborazione scientifica e lo scambio per attività di formazione.*

Nel corso delle nove precedenti edizioni, una trentina di docenti si sono alternati nell'insegnamento delle varie tematiche collegate alla documentazione linguistica.
I vari autori di questo numero speciale, che ringraziamo per il loro prezioso contributo, sono stati tutti docenti in una o più edizioni della Summer School.
Sul sito di Ethnorêma, nella sezione dedicata alla Language Documentation, è possibile consultare la lista completa degli interventi (https://www.ethnorema.it/language-documentation/). Caratteristica di tutte le edizioni è stata anche la presenza di sessioni di attività pratica, dedicate all'utilizzo di software utili alla raccolta di dati, la loro analisi e archiviazione, come Elan, Praat, FLEx, WeSay, SayMore, DiaTech, ecc.
In questo numero presentiamo anche una lista di libri, riviste e siti internet, che riteniamo essere particolarmente utile, anche se certamente non esaustiva, per chi volesse conoscere in maniera più approfondita tutto quanto ha a che fare con la documentazione linguistica e la ricerca sul campo. Abbiamo incluso i link, ove disponibili, a siti web di riferimento e a documenti scaricabili.

Editorial

Moreno Vergari (Director of *Ethnorêma*)

This special issue of Ethnorêma aims to celebrate two small but important milestones for us: **the tenth edition of our seminar series on Language Documentation and the twentieth issue of our journal.** To do so, we thought we would combine these two events. Thus, this issue of Ethnorêma is dedicated to various aspects of language documentation and is being officially launched at the Language Documentation Summer School held this year at the Free University of Bozen/Bolzano 15 to 19 July.
See https://ldsummerschool2024.wixsite.com/ldss2024.

Ethnorêma was founded in 2003 as a **voluntary association** (now registered in the National Single Register of the Third Sector as a Voluntary Organization) that aims to deal with the world's languages and cultures using a multidisciplinary approach, as well as having a special focus on development of cooperation programs. We refer to the dedicated pages on the Ethnorema website for these and other initiatives and projects.

The **first issue of the Association's journal was published in 2005**. From the beginning, a clear choice was made: the journal was to be freely downloadable to ensure its free and maximal circulation. We see with pleasure that this is now the choice followed by more and more scientific publications, but 20 years ago this was not the case.
During this time about 100 different authors have contributed to the growth of our journal, making it known and helping it find its place in the academic world and beyond.

The idea of a series of **seminars dedicated to language documentation** arose after the Association's, and therefore the journal's, registered office moved from Piedmont to Bolzano and from an initial contact with the then Language Competence Center of the Free University of Bozen-Bolzano (FUB). Since 2010, the first year of the series, collaboration between Ethnorêma and the FUB has led to six more at the FUB headquarters. It was during the seventh cycle that collaboration with the University of Tuscia in Viterbo and the University of Naples L'Orientale was born. It was then decided to hold an annual alternation of venues between Bolzano, Viterbo, where the eighth edition was held in 2022, and Naples, the venue in 2023 for the ninth edition.

Since 2022, the seminar series has taken on the official name *Language Documentation Summer School* (LDSS).

In 2023, the collaboration between the four agencies involved was enshrined in a Memorandum of Understanding, Article 1 of which outlines the objectives:

> ***Article 1 - Objectives of Collaboration***
> *The Parties, within the scope of their respective functions, competencies, and institutional roles, undertake to collaborate on the implementation of a joint action aimed to:*
> *a. deepen knowledge and innovation related to the techniques of field data collection, analysis, and cataloguing of linguistic data from minor languages, endangered languages, dialects, and jargons;*
> *b. deepen knowledge and innovation related to linguistic revitalization of threatened languages;*
> *c. carry out educational training, possibly connected with internships and apprenticeships, on the tools for collecting, analyzing, and archiving language data;*
> *d. share, compare, and disseminate such knowledge at conferences, seminars, and summer schools aimed at Italian and international audiences;*
> *e. establish and develop fruitful cooperation with the aim of promoting scientific collaboration and exchange among themselves for training activities.*

Throughout the nine previous editions, some 30 faculty members have taken turns teaching the various topics related to language documentation.
The various authors of this special issue, whom we thank for their valuable contributions, were all presenters at one or more editions of the Summer School.
On the Ethnorêma website, in the section dedicated to Language Documentation, you can see the complete list of speakers (https://www.ethnorema.it/en/language-documentation/). A characteristic of all editions has also been practical activity sessions dedicated to software useful for data collection, analysis, and archiving, such as Elan, Praat, FLEx, WeSay, SayMore, DiaTech, ecc.
In this issue, we also present a list of books, journals, and websites, that we believe to be particularly useful, though certainly not exhaustive, for those who would like to learn more about everything to do with language documentation and linguistic fieldwork. We have included links, where available, to relevant websites and downloadable documents.

Models and methods in linguistic fieldwork: a case study in reflexive meta-documentation

Peter K. Austin
pa2@soas.ac.uk
SOAS, University of London

ABSTRACT

Linguistic fieldwork has seen a progression of frameworks of research since the 1960s, from an ethical model ('fieldwork on'), to advocacy ('fieldwork for'), to collaboration ('fieldwork with'), to empowerment ('fieldwork by') (Cameron et al. 1992, Grinevald 2003, 2007, Grinevald & Bert 2011). Each of these impacts on relationships to the people researchers work with and the methods they use in the field. In addition, meta-documentation (documentation of the research itself) has emerged as an area of concern to fieldworkers and in need of elaboration and practice (Austin 2013). In this paper, I give a personal account of the models and methods I have used over 50 years of engagement with the Diyari (Dieri) Aboriginal community in South Australia (Austin 2014) as a form of reflexive meta-documentation in linguistic field-based research.

Keywords: *language description, language documentation, language revitalisation, meta-documentation, Australian Aboriginal languages, Diyari, South Australia*

ISO 639-3 code: dif

DOI: 10.23814/ethn.20.24.aus

1. Introduction[1]

Linguistic fieldwork is a method of engagement with speakers and communities that is employed by linguists and others to explore language structures and use (Bowern 2008, Meakins et al. 2018). It is useful to distinguish three types of approach to linguistic fieldwork: description, documentation, and revitalisation. We understand description as the study of language as a structural system separated from its actual use by speakers and from the social-political-cultural-economic situation in which it is used. This requires abstraction and the search for general structural principles (phonology, morphology, syntax, semantics, pragmatics, sociolinguistics). It demands idealisation, and typically involves 'cleaning up' of recordings of actual use when they are cited as instances of some descriptive phenomenon (Austin and Grenoble 2007, Austin 2010).

Data collection for description may often involve elicitation through surveys, or interviews or experiments with individuals or groups of speakers/signers. Frequently, study of a language or variety that the researcher does not themselves speak or sign is

[1] Earlier versions of the materials incorporated in this paper were presented at the Australian National University, SOAS University of London, and Mahidol University. I am grateful to audiences at these presentations for comments and feedback, and to Lise Dobrin, Julia Sallabank, and David Nathan for discussions of the general theoretical approaches taken. I received very helpful detailed feedback on an earlier draft from Lise Dobrin, David Nash, David Nathan, Jan Scott, and an anonymous reviewer. My thanks also to Greg Wilson for collaboration 2010-2013. My greatest debt is to the Diyari community, especially the Merrick, Murray, Kemp, and Warren families who have taught me about their language and culture, and welcomed me so warmly into their lives since 1974.

undertaken via translation or asking for speaker/signer judgements. Commonly, the records of interview or survey are not of interest in themselves, but are seen as a way to accumulate 'the data' for analysis. Description is typically carried out for several reasons:

- to present language structures for others to understand;
- to identify common features and differences across languages (typology);
- to reconstruct language histories and/or contact;
- to investigate how the human mind works (psycho-linguistics, neurophysiology);
- to understand the principles behind how humans interact and express personal, social, and cultural relationships.

The analysis that results from description is often highly structured and written in an abstract metalanguage (which may or may not be formalised). The audience for description is typically other researchers, and it is distributed in articles or books (grammars, dictionaries, maps, graphs, narratives, text collections).

Language documentation, by contrast, is, according to Himmelmann (2006: v):[2]

> concerned with the methods, tools, and theoretical underpinnings for compiling a representative and lasting multipurpose record of a natural language or one of its varieties.

According to Himmelmann (2006: 15), it differs from description by virtue of its focus on collecting and analysing primary data (instances of linguistic performances), accountability (analytical statements are supported by transparent access to the primary data), long-term storage and preservation of primary data (archiving), work in interdisciplinary teams, and cooperation with and direct involvement of the speech/signer community. The outcome of language documentation is frequently seen as an annotated and translated corpus of archived representative materials on a language or a variety (see also Woodbury 2011).

Language revitalisation refers to efforts to improve linguistic vitality by taking action to extend the domains of use of a language and/or to increase the number of speakers/signers (often in the context of reversing language shift), both adults and children (Austin and Sallabank 2018). Speech/signer community members may often be more interested in revitalisation than description or documentation, though increasingly many are also using documentary and descriptive approaches to support their work on language and culture learning and recovery.[3] There are several models of revitalisation in use in different situations (language nest, master-apprentice, immersion, language awareness) and communities may associate revitalisation with formal language learning in a school context.

2. Meta-documentation

Austin 2013 (building on Nathan 2010: 196) proposes the term 'meta-documentation' to refer to documentation of language documentation, description, and revitalistion research projects. This includes the nature of the materials collected and analysed, as well as project goals, history, stakeholders, biographies, attitudes and politics, methods,

---

[2] See also Austin (2010), Austin and Grenoble (2007).

[3] Communities are also being trained in applying descriptive techniques to legacy materials, as in workshops of the *National Breath of Life* in the US (https://mc.miamioh.edu/nbol/, accessed 2024-05-28) and *Paper and Talk* in Australia (https://www.livinglanguages.org.au/paper-and-talk, accessed 2024-05-27).

tools, relationships, agreements, and outcomes of a given project. These aspects of research are rarely explicitly described by linguistics researchers (other than at the outset in a project proposal, e.g. in a grant application, however this typically does not receive ongoing representation as the research is actually carried out). Woodbury (2011: 161) uses a narrower term 'project design' to cover 'the participants, their purposes, and the various stakeholders in the activity or program of activity or project'.

We suggest that meta-documentation is important for:

- developing good ways of presenting and using language research;
- future preservation of the outcomes of current projects, assisting sustainability by ensuring continuity of projects, people, and products;
- helping future researchers learn from successes and failed experiments; and
- documenting intellectual property contributions and career trajectories.

We develop below (Section 5) a case study of meta-documentation of our work with the Diyari (Dieri)[4] Aboriginal community of South Australia. We hope that in future other researchers will publish similar accounts of their description, documentation, and revitalisation projects (see also Grinevald 2003, 2005, 2007 for her work with the Rama community in Nicaragua).

3. Research frameworks

Cameron et al. (1992), Grinevald (2003, 2007) and Grinevald and Bert (2011) identify four frameworks within which research with speakers/signers has been carried out over the past sixty years:

1. *Ethical research* – research **on** a language or speakers/signers. This is defined by Cameron et al. (1992: 14-15) as:

   > a wholly proper concern to minimize damage and offset inconvenience to the researched, and to acknowledge their contributions. … But the underlying model is one of 'research on' social subjects. Human subjects deserve special ethical consideration, but they no more set the researcher's agenda than the bottle of sulphuric acid sets the chemist's agenda.

2. *Advocacy research* – research **for** speakers/signers. Cameron et al. (1992: 15) say this is:[5]

   > characterized by a commitment on the part of the researcher not just to do research on subjects but research on and for subjects. Such a commitment formalizes what is actually a rather common development in field situations, where a researcher is asked to use her skills or her authority as an 'expert' to defend subjects' interests, getting involved in their campaigns for healthcare or education, cultural autonomy or political and land rights, and speaking on their behalf.

[4] I spell the language name as Diyari (see Austin 1981, 2013, 2021 for discussion). Other spellings in the literature are Dieri, Diari, Dieyerie and variants thereof. The community preference for the name of the group is Dieri and hence the political body is named The Dieri Aboriginal Corporation TDAC (https://dieri.org.au/, accessed 2024-05-09).

[5] For Native Title in Australia (see 5.2) advocacy research "is conducted in order to provide evidence to support – or not – rather than speak on behalf of the claimants" (Jan Scott, p.c.).

3. *Collaborative research* – research **with** speakers/signers. Cameron et al. (1992: 22) describe this as:

   > the use of interactive or dialogic research methods, as opposed to the distancing or objectifying strategies positivists use. Community members participate as agents working together with researchers.

4. *Empowering research* – research **by** speakers/signers. According to Cameron et al. (1992: 24):[6]

   > In this model: (a) 'people are not objects and should not be treated as objects.' (b) 'Community members have their own agendas and research should try to address them' (c) 'If knowledge is worth having, it is worth sharing'.

Note that the last framework may involve research training and full participation of speakers/signers as equals in the research design, process, and outcomes. It has been seen by many as the favoured model for documentation and revitalisation work over the past 15 years or so (see Czaykowska-Higgins 2009, Glenn 2009, Leonard and Haynes 2010, Rice 2010, Sapien 2018, Yamada 2007).

4. Reflexive meta-documentation

I use the term 'reflexive meta-documentation' to refer to research which analyses and interprets researchers' experiences in their work, and elaborates and contextualises its goals, history, relationships, and outcomes (see Section 2). This requires reflexive and critical consideration of the broad context of historical, socio-cultural, political, and personal issues (within the community and more widely) at the time. It has been practised for the last 25 years within anthropology (including in the form of what Van Maanen (1988) called "confessional tales"), but has not been seen in mainstream linguistics. This is because linguists generally adopt a positivistic empiricist approach that excludes these kinds of contextual and experiential issues from considerations of language structures and uses. Notable exceptions are Dixon (1983), the papers by Grinevald describing her research in Nicaragua on Rama (Grinevald 2003, 2005, 2007, Grinevald and Pivot 2013), and the collection on legacy materials edited by Dobrin and Schwartz (2021).[7]

In the following sections, I describe and analyse my experiences working with the Diyari Aboriginal community, from 1975 to the present day, categorising this work into several phases according to the research framework and methods adopted at each stage. The goal of this writing is to elucidate the intellectual history of my research, make explicit its contextual circumstances and dynamics, as well as meta-document the various kinds of materials collected and produced during this 50 year period. It is to be hoped that other researchers will engage in similar reflection on their research, whether or not they adopt this particular method as a means to represent it.

---

[6] Here 'knowledge sharing' refers to researchers sharing their knowledge with those they are working with.

[7] Other autobiographies by linguists who have worked in Australia, such as Oates (2003), Glass (2018) and Swartz (2020), present their work in a narrative, rather than reflexive, way.

5. Phases of research with the Diyari Aboriginal community

The Diyari (or Dieri) are an Australian Indigenous First Nations group whose traditional lands are in the far north of South Australia, east of Kati Thanda-Lake Eyre (in Diyari, *kati thandra*) and centred along the lower reaches of Cooper Creek (*kudnarri* in Diyari). For an overview history of the language since the community was missionised by German Lutherans in 1867 see Austin (2014).

5.1 Phase I – descriptive 'research on'

The first period of my research extends from 1974 to 1990, with fieldwork in Maree, South Australia, in 1974 (for my BA Linguistics Honours at the Australian National University ANU) and 1975-1977 (PhD at ANU, submitted 1978).[8] It is important to recall that this was less than seven years after the landmark 27th May 1967 Australian Referendum[9] that removed references in the Australian Constitution that discriminated against Aboriginal and Torres Strait Islander (ATSI) people, and enabled the Australian Parliament to make laws for First Australians, as well as counting them as part of the population. It was also just two years after the erection of the Aboriginal Tent Embassy to protest for ATSI land rights,[10] and a year after the election of the Whitlam Labor federal government which instituted reforms to Indigenous rights, such as creation of the Aboriginal Land Fund for the purchase of private property, and the introduction of bilingual education in the Northern Territory (Devlin et al. (eds.) 2017). ATSI people generally held subordinate social positions, being discriminated against in a wide range of ways, and having dependent economic roles, often relying on government funds for support. Many people, including the Diyari, had moved into towns (like Maree, Port Augusta, and Broken Hill) from station properties, following the 1966 Gurindji strike in Wave Hill in the Northern Territory that established equal pay for ATSI workers.

I was introduced to the community of Diyari speakers in Maree by Luise Hercus, a researcher at ANU who was working on the neighbouring Arabana-Wangkangurru and Kuyani languages. Through observing Hercus' interactions and methods as a kind of apprenticeship,[11] I adopted what could be called an 'ethical research' approach that relied on elicitation and recording of narrative texts, as well as music performances (for more on Hercus' work and general approach see Nathan 2016, and other chapters in Austin et al. 2016). This also involved learning to speak Diyari well enough to be able to engage in qualitative dialogic interactions, with sessions recorded on tape and in fieldnotes. At the time there were about 20 multilingual speakers who had learned Diyari as children, and the language was in daily use in some families (especially among Frieda Merrick and her daughters and grandchildren). The outcome of this research was a descriptive grammar (Austin 1981, revised in a practical orthography in 2013), and academic papers about literacy, language classification, and history.[12] Austin et al.

[8] In 1978 I took up a one-year position at the University of Western Australia and began fieldwork in the Gascoyne region in the mid-north of the state. This research on Ganyara, Mandharda and Gardu languages, which could be classified as primarily 'descriptive research on', continued until 1995.

[9] See https://www.naa.gov.au/students-and-teachers/student-research-portal/learning-resource-themes/government-and-democracy/constitution-and-referendums/1967-referendum, accessed 2024-05-29.

[10] https://en.wikipedia.org/wiki/Aboriginal_Tent_Embassy, accessed 2024-05-29.

[11] I had taken a course on 'Field Methods' at ANU in 1973, but this involved one-to-one 50 minute weekly interviews over 10 weeks in a classroom with an 'informant', asking for Hua (a Papuan language) translations of English sentences on the topic of complex sentence constructions. It did not address general issues of fieldwork outside the classroom, and did not prepare me for interactions with ATSI people.

[12] For a full list see https://peterkaustin.com/publications/published-books-and-articles/, accessed 2025-05-15.

(1988) is a biography of Ben Murray, one of my main teachers; it is notable for presenting his life history partly through a collection of narrative excerpts in Diyari and Wangkangurru with English translations. Further Diyari texts and translations were published as co-authored with Ben Murray (Murray and Austin 1981).

5.2 Phase II – applied 'research for'

Beginning in around 1990, a number of significant socio-political and economic developments took place within Australia, and among the Diyari in particular. The 1992 the Australian High Court Mabo decision overturned the application of *terra nullius*, opening the possibility for ATSI groups to demonstrate a close and continuous relationships to their traditional lands, and thence to argue for title to unalienated Crown land through a land claim process (Native Title Act 1993). In 1997 a group of Diyari lodged such a land claim; a consent determination was handed down on 1st May 2012 (SCD2012/001)[13] with award of non-exclusive native title to 47,000 square kilometres east and north-east of Lake Eyre (see Austin 2014 Figure 4). A second consent determination was awarded on 26th February 2014, adding to this native title (SCD2014/003),[14] and a third one on 28th September 2017 for the eastern shore of Lake Eyre (SCD2017/001).[15] Meanwhile, in 2001 The Dieri Aboriginal Corporation (TDAC) was formed, with 600 members in New South Wales and South Australia (increasing to 1,500 by 2024); members typically are identified via family links that can be traced back to apical ancestors associated with the 19th century Lutheran mission. TDAC has negotiated various agreements concerning mining rights with a number of corporations, including Beach Energy and Santos Ltd (their role in funding language activities is mentioned below), and has several business activities in Port Augusta and Maree. Together these have resulted in TDAC being in a strong financial position to undertake social and cultural activities within the community, especially among the largest populations located in Maree, Port Augusta, and Broken Hill.

Although I had kept in touch with individual Diyari people through other researchers such was Luise Hercus and Philip Jones following the end of my fieldwork in 1977, it was not until 2010 that I had the opportunity to revisit the Diyari community.[16] This was in collaboration with Greg Wilson, who was working as a teacher-linguist on language issues for the South Australian Department of Education. Beginning in 2008, Wilson undertook numerous field trips to Port Augusta and Whyalla in collaboration with the Dieri Resources Development Group established by TDAC, and identified a range of people with knowledge of the language, most commonly words and simple sentences. All of my teachers had passed away by then, and there were only a handfull of people alive who had learnt Diyari as children, such as Rene Warren, born in 1930, and her late sister, Winnie Naylon (the grandchildren of Frieda Merrick, mentioned in 5.1). As Austin (2014, section 5) points out:

---

[13] http://www.nntt.gov.au/searchRegApps/NativeTitleClaims/Pages/Determination_details.aspx?NNTT_Fileno=SCD2012/001, accessed 2024-05-29

[14] For details of the determination see http://www.nntt.gov.au/searchRegApps/NativeTitleRegisters/Pages/NNTR_details.aspx?NNTT_Fileno=SCD2014/003, accessed 2024-05-15.

[15] https://nntt.gov.au/searchRegApps/NativeTitleRegisters/Pages/NNTR_details.aspx?NNTT_Fileno=SCD2017/001, accessed 2024-05-29

[16] In 2001 I left Australia to take up employment in Germany, the Netherlands, and the UK.

> Wilson (p.c.) reports that collection of materials for the Dieri Yawarra project required a great deal of time and that many of the people identified by TDAC as speakers struggled to remember words and sentences in Diyari after years of disuse and lack of practice. It appeared then that all that remained were a number of 'semi-speakers' or 'rememberers' of the language (Grinevald and Bert 2011), but no-one with conversational fluency or ability to record even short narratives.

Wilson elicited solely via translation from English and recorded around 2,000 items, using a revision of Austin (1981) to elucidate their grammatical structure. This led to the 2011 booklet *Dieri Yawarra: Dieri Language*, which was presented as "a handbook for community and school revitalisation and second language learning", and an accompanying CD that included cartoons and audio recordings.[17] A pilot language learning programme was initiated at Willsden Primary School in Port Augusta involving speakers and trained Diyari teacher's aides.

Wilson also began work on a larger language learning textbook entitled *Ngayana Dieri Yawarra Yathayilha: We are all speaking Dieri now* under the auspices of the South Australian Department of Education and with the support of TDAC and Beach Petroleum.[18] Recordings of all the sentence examples in the textbook were made, along with translation of several English children's songs,[19] and *Folsom Prison Blues* by Johnny Cash, a favourite of Diyari elder Rene Warren.[20] In terms of overall design, the structure of the learning materials and textbook was decided by Wilson and follows a grammar-translation model that introduces learning goals via modules focussing on particular grammatical structures, such as:

Nhawurda ngakarni mar̲a.	This is my hand.
Ngayana Dieri wima wangkayilha.	Let's sing a Dieri song.
Ngakarni para mar̲u marla.	My hair is really black.

I served as an unpaid consultant on the textbook, creating a vocabulary list, and checking the transcription, translation, and grammatical analysis of all components. Unfortunately, the textbook was not completed, and remains in 2013 draft form only.

In 2011, with encouragement and support from Wilson and myself, TDAC successfully applied for a grant from the Indigenous Languages Support (ILS) scheme funded by the Federal Government Department of Regional Australia, Local Government, Arts and Sport. The grant ran from July 2012 to September 2013 and included four community language and culture workshops led by Wilson and myself (February and April 2013 in Adelaide, and March and August 2013 in Port Augusta).[21]

[17] Sadly the CD-ROM no longer functions on current computers and the materials on it are locked in an inaccessible format.

[18] Work on both textbooks was funded through Commonwealth Shared Responsibility Agreement (SRA) grants involving The Dieri Aboriginal Corporation (TDAC), the Department of Water, Heritage and the Arts (DEWHA), Families, Housing, Community Services, Indigenous Affairs (FaHCSIA), the South Australian Department of Education & Children's Services (DECS), and Beach Petroleum.

[19] Such as *Old McDonald's Farm* (https://dieriyawarra.wordpress.com/2013/03/01/diyari-wima/, accessed 2024-05-15)

[20] For the Johnny Cash song see https://dieriyawarra.wordpress.com/2013/03/19/folsom-prisonanhi/, https://dieriyawarra.wordpress.com/2013/04/23/folsom-prison-mandru/, https://dieriyawarra.wordpress.com/2013/08/04/folsom-prison-parkulu/, and https://dieriyawarra.wordpress.com/2013/08/05/folsom-prison-mandru-mandru/, all accessed 2024-05-15.

[21] TDAC provided travel funds in support of the ILS grant in order to enable participants from Broken Hill and Port Augusta to participate.

These workshops brought together 60 Diyari participants on each occasion, ranging in age from five to eighty. The curriculum covered basic vocabulary and grammatical structure (such as pronouns, imperative verb forms) and was essentially teacher-led by Wilson and myself (in the spirit of ‘research for’). Through my conversations in Diyari with Rene Warren it became apparent that she was a fluent, if rusty, speaker, while her son Reg Warren could understand everything that his mother and I said, but was not as productively skilled as his mother. They served as pronunciation models in the ILS workshops.

To support the ILS activities I created the *Ngayana Dieri Yawarra Yathayilha* blog which presented posts reporting on the workshops and illustrating and explaining aspects of vocabulary, grammar, and simple conversations.[22] The reasons for choosing the blog format were that it enables timely reports on activities, bite-sized language lessons, links between related posts, and the inclusion of images and media. I continued posting after the ILS project concluded, and by early May 2024 the blog had 116 posts, and has had 48,100 page views since January 2013. It typically attracts 50-100 views per week from an international audience. The blog was publicised on social media, especially *Facebook*. In 2023 a podcast called *Diyari Yawarra* was begun,[23] presenting some of the blog posts in audio form in response to requests from community members to be able to hear spoken Diyari, as well as read it.

A favourite activity at the 2013 workshops was singing, and on the participants’ initiative Wilson and I created a Diyari translation of a country music song written and performed by then TDAC Chairperson and award-winning performer Chris Dodd.[24] Diyari lands are arid with an average of just 270mm (10.5 inches) of rain annually,[25] and the song celebrates a special phenomenon which is very occasional heavy rainfall in the Channel Country of Queensland (around 800km north-east of Killalpaninna) and the subsequent flooding of Cooper Creek that runs through Diyari traditional lands. This brings with it abundant water, fish, and bird and animal life, while promoting subsequent luxuriant plant growth:

ngapa-ngapa pirna ngar̲iyi	Lots of water is coming down
ngarrimatha wakar̲ayi	A flood is coming
thalar̲a pirna kurdayi	Lots of rain is falling
ngayanarni mithanhi	In our country
daku pirna thana	There are big sandhills
matya ngayana pankiyilha	So we are happy now
ngapa pirna ngakayi	Lots of water is flowing
parru pirna pakarna	And big fish (are coming) too

As well as celebrating links to country and a significant recurring historical event, the vocabulary and grammatical structure of songs such as this can serve as a resource for language learning.

One challenge I faced at the time with continuing and expanding this work was that I was fully-employed at a university on the other side of the world, and that the outcomes of the work, including externally-refereed publications on topics like meta-

[22] See https://dieriyawarra.wordpress.com/, accessed 2024-05-15.
[23] See https://open.spotify.com/show/6PnZ6YkkxfF8c0PvL5uS8c, accessed 2024-05-15.
[24] See https://www.youtube.com/watch?v=5hFVjQFJQzM, accessed 2024-05-15.
[25] https://www.worldweatheronline.com/v2/weather-averages.aspx?q=dlk, accessed 2024-06-06

documentation and intellectual history, were deprecated by my Head of Department and considered as "not Linguistics", especially in the context of the periodic UK Research Assessment Exercise. I retired from academia in December 2018, which left me free to be more mobile and to carry out research and write in ways that were outside 'academic norms' (including blog posts, podcasts, and interactions on social media).

5.3 Phase III – 'research with'

The third phase begins in 2022 with TDAC funding and organising a research trip to the Lutheran mission site at Killalpaninna near Cooper Creek, focussing on documentation of ethno-botanical knowledge and practice, as well as traditional cooking methods. The research team comprised four generations of Diyari (including members of the Warren family), myself as linguist, an anthropologist, an archaeologist, a plant specialist, a community development specialist, and a videographer. The Diyari group included two teenagers (great-grandchildren of Rene Warren) as an experiment in taking members of the younger generation to traditional country and encouraging them to learn about their heritage language and culture. For this project, the goals and methods were set in collaboration with the Diyari participants, and centred around video-recorded interviews, including several by Michelle Warren with her grandmother Rene Warren in the form of experientially-embedded conversations about mutually-shared occasions of identifying and using plants. One of these is *yawa* a small tuber, often called 'wild onion, bush onion' (*Cyperus bulbosus*) that was a subsistence staple on Diyari country until groups moved into the towns of Maree and Port Augusta in the 1960s.[26] This project was thus collaborative and engaged the Diyari members as active co-participants. One activity was particularly co-operative. Taryn Debney told the group that her archaeological research in Diyari country had identified numerous sites with instances of clay balls (*tyaputyapu*) being used in ground ovens as heat retainers. The ethnographic and historical record, including oral histories, had shown no evidence of their use since the mid-19th century. The two teenage participants experimented with the clay soil at Killalpaninna and were able to recreate clay ball heat retainers and to cook damper bread using them.[27] I updated the Diyari blog and podcasts to present these activities and provided vocabulary and example sentences drawn from them, thus potentially supporting language learning by users.

5.4 Phase IV – 'research by'

The final phase covers 2023-2024 and includes activities wholly organised and carried out by members of the Diyari community, with my role being a supporting consultant. In 2023 the sub-community in Broken Hill (NSW) expressed interest in holding a family-oriented weekend (12-13th November) of *Yarning About Language* focussed on language and culture activities, in collaboration with knowledge-holders in Port Augusta.[28] I was in Adelaide as part of my research ('on' and 'for') funded by the Leverhulme Trust on missionary J. G. Reuther's massive Diyari-German dictionary (see Austin 2023), and was invited by the families to participate. Approaches by the

---

[26] For more details and photographs see https://dieriyawarra.wordpress.com/2022/12/20/word-of-the-day-yawa/, accessed 2024-05-15.

[27] For further details and photographs see https://dieriyawarra.wordpress.com/2022/12/21/tyaputyapu/, accessed 2024-05-15.

[28] For further details and pictures see https://dieriyawarra.wordpress.com/2023/11/15/yarning-about-language-dieri-families-workshop/, accessed 2024-05-12.

interested Diyari participants to TDAC for funding were refused as part of wider political issues and disagreements within the community,[29] so the participating family members and I paid for all the costs of travel, accommodation, food, and room hire. The two days were attended by 60 participants, including Diyari local knowledge holders, artists, musicians, and interested persons aged from 5 to 93 years. Taryn Debney also attended and gave a short presentation on archaeology and the research trip discussed in 5.3 above, including presenting a compilation video.

The workshop curriculum was decided and facilitated by Michelle Warren, on the topics of greetings, body parts, children's songs, and an interactive game which she and I invented called lingo-bingo.[30] Michelle had received some basic linguistic training a few years before via the Research Network for Linguistic Diversity (now Living Languages),[31] and had attended a workshop in Alice Springs several years prior on the topic of language revitalisation through the master-apprentice model.[32] Michelle had adapted this approach to the Diyari situation with her grandmother Rene Warren as a Diyari master (and to a lesser extent myself) and the participants as apprentices. All learning was interactive and engaged every participant, building on existing knowledge, and exploring new contexts. So, for example, in discussing the term *thina* 'foot', some of the children knew the expression *thina puta parlu* 'without any shoes' as it was regularly used by their father before they left the house (*thina* 'foot', *puta* 'shoe, loan from English boot', and *parlu* 'naked'). Thus, existing knowledge by community members was incorporated into the sessions. Michelle emphasised self-expression without concern for literacy and 'correct spelling' in order to encourage spoken language use. Lingo-bingo was particularly popular. This is a card matching game where participants joined teams that each created 20 playing cards with pictures on one side and their Diyari names on the other (e.g. *nganthi* 'meat', *ngapa* 'water'). Teams selected 10 of their cards to play with and placed them picture-side-up in front of them. Michelle then randomly called out Diyari names, and teams with a match turned them over to show the Diyari word; the winner was the team that turned over all their cards and called "lingo-bingo". The goal of the game is to create an enjoyable context for vocabulary listening and recognition skills, as well as passive literacy learning.

As a follow up to the workshop, Michelle Warren had two sets of lingo-bingo cards and instructions professionally drawn (by her daughter, a talented artist) and printed, with one set to be available in Port Augusta and one in Broken Hill. Also, during November and December 2023 I wrote a series of posts on *Facebook* with a picture of a local event or cartoon and a description of it in Diyari as a way to demonstrate that the language can be used to talk about anything that people experience in their daily lives. An example is a picture of a snake catcher in Broken Hill with the caption *karnali wanku pardakayi yakuthanhi wirripalha* 'The (Aboriginal) man is picking up a snake to put in a bag' (the vocabulary and grammatical analysis is given in a comment under the Facebook post). These posts attracted positive responses from Diyari Facebook users,

[29] Indeed, I received a letter on 6th November 2023 from the TDAC Chairperson stating that the Board had made a "recent decision to work with SA Museum & Mobile Language Team to get Dieri Language out to the whole of the Dieri Community … Due to the TDAC Boards decision to not proceed with this, on behalf of the TDAC Board of Directors I ask that the Dieri Language Workshop in Broken Hill, does not proceed".

[30] Compare an earlier version developed in the ILS workshops that relied on literacy, rather than emphasising oral listening and speaking (https://dieriyawarra.wordpress.com/2013/09/03/ngayana-pirkirna-warayi-lingo-bingo/, accessed 2024-05-28).

[31] See https://www.livinglanguages.org.au/training, accessed 2024-05-29.

[32] See https://en.wikipedia.org/wiki/Master-Apprentice_Language_Learning_Program, accessed 2024-05-15.

especially Diyari people in Broken Hill, but did not lead to further public language use, primarily because of worries about literacy and "getting the spelling wrong". In a written context it is difficult to address these concerns in a meaningful way.

The final activity to be reported in this phase is a pilot translation project carried out in Port Augusta on 16th April 2024. This was organised by Michelle Warren, who invited me to join her and Diyari language experts her grandmother Rene Warren and father Reg Warren. The goal was to collaborate with a local primary school teacher (with permission from the parents) to translate some young Diyari children's writing from English to Diyari as a means to assist them with their literacy skills, self-esteem, sense of identity, and, potentially, Diyari language learning. The stories were based on photographs of the children doing various everyday activities for which they wrote a description, for example, "this is me on the sports oval". After discussions among the team, we decided on an appropriate Diyari expression (in this case *nhaniya nganhi kanthanhi tharkayi* 'this is me standing on the grass') and Michelle then recorded Rene, Reg and herself saying it. This translation process, led by the Diyari speakers, revealed interesting issues about demonstratives in Diyari which I had not been consciously aware of previously. Thus, Diyari has a feminine versus non-feminine contrast in proximal and distal demonstratives that is important when translating "This is me" as correct usage depends on the gender of the speaker (in this case the children are female so *nhaniya* is the correct translation for 'this' rather than *nhawuya* 'this (non-feminine)').[33] Future work on children's writing is planned which will include creation of printed and multimedia outputs, and applying for funding to scale up the project to include more Diyari children at the school.

6. Conclusions

I have been engaged in language and culture work with the Diyari community of South Australia for some 50 years and this has progressed from description to documentation to applied revitalisation, and from 'research on' to 'research for' to 'research with' and 'research by'. These developments have been rewarding personally and professionally, and have resulted in various outcomes (including academic and non-academic products), and insights into language structure and use, both within the community and in the wider academic world. The recent language support work (Phases III and IV) has relied on the solid documentary and descriptive base established in the earlier phases, including the fact that I had learnt to speak Diyari with some degree of fluency. This revitalisation work has involved substantial personal, academic, social, and political challenges (including intra-community disagreements) that have involved me taking a range of roles over time from outside academic researcher determining models and methods, to specialist consultant on community-led projects.

As Amery (2009) and Wilkins (1992) argue, it is easy for researchers to fall into simplistic "solutions" that do not work, either linguistically or socio-politically – the workshops and learning materials discussed in Phase II seem to me to fall under this description. It is important to be ready to listen and try to understand what people mean

[33] For further discussion see https://dieriyawarra.wordpress.com/2024/04/20/how-to-say-this-and-that-in-diyari/. Subsequent blog posts discuss the forms and meaning of plural demonstratives (https://dieriyawarra.wordpress.com/2024/04/21/how-to-say-these-and-those-in-diyari/), inflectional forms (https://dieriyawarra.wordpress.com/2024/04/21/advanced-more-about-this-and-that-and-these-and-those-in-diyari/), and use with non-singular pronouns (https://dieriyawarra.wordpress.com/2024/04/24/how-to-say-this-is-us-in-diyari/), all accessed 2024-05-15.

by what they say, and this relies on close, open and long-term personal relationships, as well as effective participant observation (Dobrin and Schwartz 2016). It is essential to set one's own political assumptions aside and to seek possible solutions and sustainable outcomes through open and equitable discussions in a realistic context, while recognising that conflict and disputation is an inevitable part. It is important to develop concrete outcomes, while being careful not to overpromise and thereby raise expectations that cannot be fulfilled and whose non-achievement will lead to disappointment. Sometimes the best outcomes are the processes of shared learning and oral language use, while apparently desirable printed works like dictionaries and textbooks are little more than talismans. In addition, academic requirements such as publication of journal articles and books can be in conflict with other forms of writing, such as social media posts and blogs, even though the latter garner more users and can communicate more effectively with community members.

Sometimes, it is essential for outsiders to simply say and do nothing, and be patient. Timing, personalities, and the willingness to be flexible and change roles and relationships are all important variables in the success of research of any type. Finally, in presenting this case study, I hope I have made a strong case for reflexive meta-documentation of research projects to assist with understanding and interpreting their goals, trajectories, outcomes, and longer-term impacts.

References

AMERY, Rob. (2009). “Phoenix or relic? Documentation of languages with revitalization in mind”. *Language Documentation and Conservation* 3(2): 138-148.

AUSTIN, Peter. (1981) *A Grammar of Diyari, South Australia*. Cambridge: Cambridge University Press.

AUSTIN, Peter K. (2010) “Current issues in language documentation”. *Language Documentation and Description* 7: 12-33.

AUSTIN, Peter K. (2013) “Language documentation and meta-documentation”. In JONES, Mari, and Sarah OGILVIE (eds.) *Keeping Languages Alive: Documentation, Pedagogy and Revitalization*, 3-15. Cambridge: Cambridge University Press.

AUSTIN, Peter K. (2014) “And still they speak Diyari: the life history of an endangered language”. *Ethnorema* 10: 1-17.

AUSTIN, Peter K. (2020) “Language documentation and revitalisation”. In OLKO, Justyna and Julia SALLABANK (eds.) *Revitalizing endangered languages: a practical guide*, 199-219. Cambridge: Cambridge University Press

AUSTIN, Peter K. (2021) *A Grammar of Diyari, South Australia*. 2nd edition, version 2.10. Manuscript, SOAS, University of London.

AUSTIN, Peter K. (2023) “Making 2,180 pages more useful: the Diyari dictionary of Rev. J. G. Reuther”. In DEHERMI Eda, and Christopher MOSELEY (eds.) *Endangered Languages in the 21st Century*, 241-257. London: Routledge.

AUSTIN, Peter K. and Lenore A. GRENOBLE. (2007) “Current trends in language documentation”. *Language Documentation and Description* 4: 12-25.

AUSTIN, Peter, Luise A. HERCUS and Phillip JONES. (1988) “Ben Murray Parlku-Nguyu-Tangkayiwarna”. *Aboriginal History* 12: 115-188.

AUSTIN, Peter, Harold KOCH, and Jane SIMPSON (eds.) (2016) *Language, land & song: Studies in honour of Luise Hercus*, 57-69. London: EL Publishing.

AUSTIN, Peter K. and Julia SALLABANK (2018) “Language documentation and language revitalisation: Some methodological considerations”. In HINTON, Leanne, Leena HUSS and Gerald ROCHE (eds.) *Handbook of Language Revitalisation*, 207-215. London: Routledge.

BOWERN, Claire. (2008) *Linguistic Fieldwork: A Practical Guide*. Basingstoke: Plagrave Macmillan.

CAMERON, Deborah, Elizabeth FRAZER, Penelope HARVEY, Ben RAMPTON, and Kay RICHARDSON. (1992) *Researching Language: Issues of Power and Method*. London: Routledge.

CZAYKOWSKA-HIGGINS, Ewa. (2009) “Research Models, Community Engagement, and Linguistic Fieldwork: Reflections on Working within Canadian Indigenous Communities” *Language Documentation and Conservation* 3: 15-50.

DEVLIN, Brian Clive, Samantha DISBRAY, and Nancy Regine Friedman DEVLIN. (2017) History of Bilingual Education in the Northern Territory: People, Programs and Policies. Singapore: Springer.

DIXON, R. M. W. (1983) *Searching for Aboriginal languages: Memoirs of a field worker*. St. Lucia, Qld.: University of Queensland Press. (available for download at https://espace.library.uq.edu.au/view/UQ:203856)

DOBRIN, Lise and Saul SCHWARTZ. (2016) "Collaboration or Participant Observation? Rethinking Models of 'Linguistic Social Work'". *Language Documentation and Conservation* 10: 253-277.

DOBRIN, Lise and Saul SCHWARTZ. (2021) "The social lives of linguistic legacy materials". *Language Documentation and Description* 21. special issue.

FLORES FARFÁN, José Antonio, and Justyna OLKO. (2020) "Types of Communities and Speakers in Language Revitalization." In OLKO, Justyna, and Julia SALLABANK (eds.) *Revitalizing endangered languages: a practical guide*, 85-103. Cambridge: Cambridge University Press.

GLASS, Amee. 2018. *Taking the word to the desert people*. Kangaroo Ground, Vic.: Wycliffe Bible Translators Australia.

GLENN, Akiemi. (2009) "Five Dimensions of Collaboration: Toward a Critical Theory of Coordination and Interoperability in Language Documentation" *Language Documentation and Conservation* 3: 149–160

GRENOBLE, Lenore A. (2013) "Language revitalization". In BAYLEY, Robert, Richard CAMERON and Ceil LUCAS (eds.) *The Oxford Handbook of Sociolinguistics*, 792-811. Oxford: Oxford University Press.

GRINEVALD, Colette. (2003) "Speakers and documentation of endangered languages". *Language Documentation and Description* 1: 52-73.

GRINEVALD, Colette. (2005) "Why Rama and not Rama Cay creole?", *Language Documentation and Description* 3: 196-224.

GRINEVALD, Colette. (2007) "Linguistic fieldwork among speakers of endangered languages". In MIYAOKA Osahito, Osamu SAKIYAMA, and Michael E KRAUSS (eds.) *The Vanishing Languages Of The Pacific Rim*, 35-76. Oxford: Oxford University Press.

GRINEVALD, Colette and Michel BERT. (2011) "Speakers and communities". In AUSTIN Peter K., and Julia SALLABANK (eds.) *The Cambridge Handbook of Endangered Languages*, 45-65. Cambridge: Cambridge University Press.

GRINEVALD, Colette and Bénédicte PIVOT. (2013) "On the revitalization of a 'treasure language': The Rama language project of Nicaragua", In JONES, Mari, and Sarah OGILVIE (eds.) *Keeping Languages Alive: Documentation, Pedagogy and Revitalization*, 181-197. Cambridge: Cambridge University Press.

HIMMELMANN, Nikolaus P. (1998) "Documentary and descriptive linguistics". *Linguistics* 36: 161-195

HIMMELMANN, Nikolaus P. (2006) "Language documentation: What is it and what is it good for?". In GIPPERT Jost, Nikolaus P. HIMMELMANN, and Ulrike MOSEL (eds.) *Essentials of Language Documentation*, 1-30. Berlin: Mouton de Gruyter.

HIMMELMAN, Nikolaus P. (2012) "Linguistic Data Types and the Interface between Language Documentation and Description". *Language Documentation and Conservation* 6: 187-207.

LEAVY, Patricia. (2020) *Method Meets Art: Arts-Based Research Practice*. New York: Guilford Press.

LEONARD, Wesley Y. and Erin HAYNES. (2010) "Making 'collaboration' collaborative: An examination of perspectives that frame linguistic field research." *Language Documentation and Conservation* 4: 268-293.

MEAKINS, Felicity, Jennifer GREEN and Myfany TURPIN. (2018). *Understanding Linguistic Fieldwork*. London: Routledge.

MURRAY, Ben and Peter AUSTIN. (1981) Afghans and Aborigines: Diyari texts. *Aboriginal History* 5: 71-79.

NATHAN, David. (2010) "Language documentation and archiving: from disk space to MySpace". *Language Documentation and Description* 7: 172-208.

NATHAN, David (2016) "'Don't tell them we're coming!': learning to document languages with Luise Hercus". In AUSTIN, Peter, Harold KOCH, and Jane SIMPSON (eds.) *Language, land & song: Studies in honour of Luise Hercus*, 57-69. London: EL Publishing.

OATES, Lynnette. (2003) *Against the wind: Wycliffe Bible Translators Australia in action*. Lavington, N.S.W.: G. Van Brummelen.

OLKO, Justyna and Julia SALLABANK (eds.) (2020) *Revitalizing endangered languages: a practical guide*. Cambridge: Cambridge University Press.

RICE, Keren. (2010). The linguist's responsibilities to the community of speakers. In GRENOBLE Lenore A. and N. Louanna FURBEE (eds.) *Language documentation: Practice and values*, 25-36. Amsterdam: John Benjamins.

SAPIÉN, Racquel María. (2018) "Design and Implementation of Collaborative Language Documentation Projects". In REHG, Kenneth L. and Lyle Campbell (eds.) *The Oxford Handbook of Endangered Languages*, 203-224. Oxford: Oxford University Press.

SWARTZ, Steve. (2020) *Broken Pot: The Making of the Warlpiri Bible*. Santa Monica: Initiate Media.

VAN MAANEN, John. (2011) *Tales of the Field: On Writing Ethnography*. Second Edition. Chicago: University of Chicago Press.

WILKINS, David P. (1992). "Linguistic research under aboriginal control: A personal account of fieldwork in Central Australia". *Australian Journal of Linguistics* 12: 171-200.

WOODBURY, Anthony C. (2003) "Defining documentary linguistics". *Language Documentation and Description* 1: 35-51.

WOODBURY, Anthony C. (2011) "Language documentation". In AUSTIN, Peter K., and Julia SALLABANK (eds.) *The Cambridge Handbook of Endangered Languages*, 159-186. Cambridge: Cambridge University Press

YAMADA, Racquel-María. (2007). "Collaborative linguistic fieldwork: Practical application of the empowerment model". *Language Documentation and Conservation* 1(2): 257-282.

Documentazione linguistica e ultimi parlanti: l'esperienza con i dialetti walser

Silvia Dal Negro
Libera Università di Bolzano
Silvia.DalNegro@unibz.it

ABSTRACT

This paper discusses the various challenges encountered in documenting endangered languages spoken in extremely small language communities by speakers who are not always fluent in those languages. On the one hand, some specificities of these languages will be considered, namely their predominantly oral nature and the inherent multilingualism of these communities. On the other hand, some considerations on the nature of the linguistic datum in relation to different elicitation techniques will be proposed.

The following reflections are based on a long-lasting fieldwork experience in Walser communities in Italy.

Keywords: *endangered languages, Walser communities, data elicitation*

DOI: 10.23814/ethn.20.24.dal

Introduzione

Il crescente interesse, anche in contesto italiano, per l'ambito di ricerca che fa capo alla documentazione linguistica richiede, credo, qualche riflessione di carattere più generale per evitare che tale etichetta, se applicata a qualsiasi indagine che includa una fase di raccolta di dati su lingue diverse dalla propria, si svuoti completamente di senso[1]. Di fatto, ragionare di documentazione linguistica implica anche una riflessione sulla natura del dato linguistico, sulle modalità di elicitazione dello stesso, e sulla sua rappresentatività in relazione alla comunità linguistica di riferimento. In questo contributo si metterà al centro dell'attenzione un tipo specifico di ricerca, e cioè la documentazione di lingue prossime all'estinzione per le quali gli ambiti d'uso per così dire naturali risultano estremamente ridotti, così come il numero di parlanti fluenti, a fronte di una maggioranza di membri della comunità che dispone di competenze linguistiche fortemente erose o che non ne dispone affatto. Come è noto, infatti, sono proprio questi i contesti che, più di altri, richiederebbero un lavoro di documentazione che sia quanto più possibile ricco e vario: se non necessariamente finalizzato alla rivitalizzazione della lingua, almeno a restituirne una testimonianza realistica.

Gli esempi dai quali trarrò spunto derivano dall'esperienza di ricerca che si estende su quasi tre decenni in area walser piemontese e valdostana, nell'ambito di progetti di ricerca e di valorizzazione della lingua condotti o in prima persona o come parte di gruppi più ampi. Come vedremo, questi progetti rientrano solo in parte in quella che

1 Il presente testo riprende in larga parte quanto presentato in occasione del convegno "Documenting languages, Documenting cultures" (Napoli, Università Federico II, 5-6 ottobre 2023). Ringrazio, per gli stimoli molto utili ricevuti, l'organizzatrice del convegno, Margherita Di Salvo, le persone intervenute in quell'occasione, e i revisori anonimi di questo saggio.

potremmo definire documentazione linguistica in senso stretto e per questo daranno la possibilità di riflettere sul lavoro di ricerca sul campo e sulle sue finalità.

Come contesto d'indagine, quello delle piccole comunità alpine del Piemonte e della Valle d'Aosta parlanti dialetti tedeschi di tipo alemannico, cosiddetti walser[2], rientra a pieno titolo nel novero delle minoranze linguistiche a fortissimo rischio di erosione. Fra l'altro, sebbene localizzati nel cuore dell'Europa, i dialetti parlati in queste comunità sono stati, fino a pochi decenni fa, in larga parte carenti di accurati lavori di documentazione e di descrizione linguistica. Frutto di migrazioni medievali da nord a sud delle Alpi e sopravvissute per secoli in un fragile equilibrio fra isolamento e contatti plurimi (con il territorio circostante, transfrontalieri e transalpini dovuti a migrazione stagionale o permanente), queste comunità hanno conosciuto, a partire dalla seconda metà del Ventesimo secolo, un rapido decadimento in termini sia demografici sia, soprattutto, linguistici. Oggi le parlate walser italiane possono essere definite, utilizzando i parametri UNESCO relativi alla vitalità delle lingue minacciate (UNESCO 2003), *severly* o *critically endangered* (Dal Negro 2011). Paradossalmente, questo processo di *language shift* ha coinciso con un risveglio di consapevolezza etnico-linguistica sostenuto, fra l'altro, dalla *Internationale Vereinigung für Walsertum* 'Associazione Internazionale Walser' con sede a Briga, in Svizzera, che dal 1962 riunisce tutte le associazioni locali in rappresentanza delle comunità walser (incluse diverse già estinte linguisticamente) di Svizzera, Italia, Austria e Liechtenstein. Oltre a ciò vanno segnalate le diverse forme di tutela regionale e nazionale (per una panoramica recente cfr. Fiorentini 2022) e i numerosi interventi di politica linguistica atti ad arginare l'inesorabile declino di parlanti, facendo leva soprattutto su un mutato atteggiamento linguistico, in molti casi, però, forse più di facciata che di sostanza[3].

Quali dati per la documentazione linguistica?

Trattando di documentazione linguistica, un aspetto che viene spesso sollevato è quello relativo alla naturalezza del dato. A questo proposito vanno fatte alcune precisazioni, valide in generale per quanto riguarda la raccolta di dati linguistici, ma particolarmente significative nell'ambito della documentazione linguistica. Innanzitutto, seguendo ad esempio Himmelmann (Himmelmann 1998), è necessario tenere separati due piani che vengono spesso confusi: quello della naturalezza e quello della spontaneità. Mentre con naturalezza ci si riferisce al grado di controllo da parte del rilevatore nei confronti del dato, con spontaneità il controllo è esercitato da chi parla verso ciò che viene detto (o scritto). Nel primo caso, dunque, i dati si possono collocare su una linea continua che va da un polo di massima naturalezza con i dati registrati nell'ambito di eventi comunicativi reali, cioè eventi che avrebbero luogo indipendentemente dall'attività di documentazione, al polo opposto dove si collocherebbero compiti elicitativi formali, quali ad esempio la traduzione di parole o la formulazione di giudizi di grammaticalità. Per quanto riguarda invece la spontaneità, una produzione linguistica spontanea è tale in quanto poco controllata, o meglio, poco pianificata. Inoltre, la dimensione della

---

2 *Walser* da *Walliser*, cioè vallesani. Si tratta, di fatto, dei dialetti di tipo *Höchstalemannisch* (letteralmente 'alemannico altissimo') parlati nelle vallate del Vallese germanofono, in parti della Svizzera centrale, e nelle comunità walser italiane, grigionesi e nella ticinese Bosco Gurin (Zinsli 2002). Chiaramente la specificità dei dialetti alemannici parlati da secoli in contesto italoromanzo o galloromanzo pone questi su un piano di maggiore interesse sia dialettologico, sia più in generale della linguistica.

3 Alcuni esempi di questo mutato atteggiamento, sia interno alle comunità walser, sia esterno, sono discussi in Dal Negro (2020). Si segnala in particolare la crescita, nel corso del tempo, di marchionimi walser, così come dell'uso del walser nel paesaggio linguistico, istituzionale e privato.

pianificazione del discorso tende a sovrapporsi, seppure non del tutto, con i livelli di variazione diafasica relativi a gradi maggiori o minori di formalità.

Nella Figura 1 si forniscono alcuni esempi di eventi comunicativi risultanti dall'intersezione delle due variabili considerate, e cioè la naturalezza e la spontaneità.

	+ NATURALE	- NATURALE
+ SPONTANEO	*conversazione in famiglia* *ritrovo informale fra amici*	*task narrativi* *task dialogici*
- SPONTANEO	*discorso pubblico* *narrazione tradizionale*	*lettura lista di parole* *task di completamento*

Figura 1: Tipologia di eventi comunicativi

La distinzione fra naturalezza e spontaneità è particolarmente importante ai fini di una valutazione del dato raccolto in relazione agli obiettivi del lavoro di ricerca o di documentazione, ma meno netta di quanto possa apparire ad un primo sguardo. Ad esempio, come argomentano convincentemente Klamer e Moro (2020) a proposito dell'utilizzo di stimoli visivi (vignette, albi illustrati, videoclip) per elicitare il parlato, le narrazioni che ne scaturiscono sono spontanee (in quanto non pianificate), ma non naturali, mentre la narrazione di leggende e racconti tradizionali, sebbene più naturale in quanto contestualizzata nelle pratiche comunitarie, non è necessariamente spontanea, trattandosi di testi almeno parzialmente codificati e ripetuti infinite volte.

Infine, è importante tenere presente che ad un aumento progressivo del controllo sul dato, da parte sia del ricercatore che dell'informante, aumenta anche il peso che le teorie implicite ed esplicite sulla lingua (e sulle lingue) esercitano sulla natura del dato stesso. Ciò riguarda *in primis* che cosa si aspettano entrambi in un determinato contesto in termini di scelta della lingua, o della varietà di lingua, o, infine, di una variante, come ben sintetizzato in Iannàccaro (2001) e più estesamente in Iannàccaro (2002).

Certamente entrambe le dimensioni, quella della naturalezza e quella della spontaneità, vanno intese come continue, per cui non è difficile immaginare gradi diversi sia di naturalezza (un evento organizzato ai fini di una raccolta dati, come ad esempio una cena, è naturale o non lo è?), sia di spontaneità. Si pensi in quest'ultimo caso al diverso effetto prodotto da un intervistatore che è parte della comunità rispetto a un intervistatore sconosciuto[4].

Provando ad applicare la tipologia di eventi comunicativi schematizzata nella Figura 1 al contesto di lingue caratterizzate da condizioni sociolinguistiche molto critiche, come lo sono le parlate walser in Italia, ci si rende subito conto che molto probabilmente in nessuna di queste comunità una raccolta dati caratterizzata da spontaneità e naturalezza sarebbe oggi fattibile, se non forse in rarissime eccezioni (su questo cfr. anche Adamou 2016 relativamente a contesti simili). Questo restringe fortemente il ventaglio di possibilità di documentazione della lingua minoritaria[5] limitandole (nel migliore dei casi e progressivamente sempre di meno) a situazioni costruite *ad hoc* (anche di carattere

[4] Su questo cfr. alcune osservazioni in Ciccolone e Dal Negro (2021: 36-39).

[5] Questo porterebbe invece alla documentazione delle altre lingue presenti nella comunità e, in particolare, delle varianti locali di queste.

informale) nelle quali sia esplicitamente richiesto l'uso della parlata locale, ad esempio proprio a scopo documentaristico, oppure didattico, cioè finalizzate a rafforzare la pratica della lingua stessa.

Un esempio del primo tipo è ben rappresentato dall'estratto riportato in (1), registrato ormai più di due decenni fa nel corso di un progetto di documentazione del walser parlato nelle comunità piemontesi di Formazza e Rimella (Dal Negro 2006). Dal comportamento verbale e non verbale (risa) delle partecipanti è evidente come la situazione etichettabile come "trovarsi per parlare il titsch[6]" risulti a tutte abbastanza artificiosa. Come esempio del secondo tipo riporto invece parte del testo di una locandina pubblicizzata sui *social network* nel gennaio 2024 relativa ad una iniziativa promossa dal "Walser Kulturzentrum" di Gressoney in Valle d'Aosta (2). Sia il titolo dell'iniziativa, *Tieber zéeme redò* 'parliamo insieme', sia la descrizione "si parlerà solo in titsch", lasciano intendere l'eccezionalità e l'artificiosità di un uso spontaneo della lingua minoritaria e, almeno nelle intenzioni, monolingue (ci torneremo sotto). D'altra parte, dati raccolti in contesti di questo tipo permettono di ottenere una documentazione comunque più naturale dell'elicitazione a partire da stimoli controllati dal ricercatore e più vicina a pratiche comunicative reali, o almeno realistiche.

(1) ADR: ah z Tschuljé riwut
'ah arriva la Giulia'
SIL: [RIDE]
GIN: ja, götän abä Tschuljé
'sì, buona sera Giulia'
SIL: chuntsch grat zittléch
'arrivi proprio al momento giusto'
ADR: ja, séwär zwäk titsch äm béts zellä
'sì, stiamo parlando un po' titsch'
GIU: ah [RIDE]
(Formazza – Archivio Sonoro Walser)

(2) Il Walser Kulturzentrum
organizza
«TIEBER ZÉEME REDÒ»
incontri in cui si parlerà solo in titsch
gratuiti e aperti a tutti,
che si terranno una volta a settimana.

Due *bias* nella documentazione di piccole lingue
L'amodalità

Per la maggior parte, le lingue a rischio di estinzione sono lingue primariamente, se non esclusivamente, parlate, nel senso che un'eventuale resa grafica delle stesse è molto recente, talvolta spontanea (nel senso di Iannàccaro, Guerini e Dal Negro 2015), più spesso elaborata da altri per scopi scientifici e/o descrittivi, in ogni caso appannaggio di pochi membri della comunità[7]. Ciò vale, in Italia, per i dialetti italoromanzi (inclusi

[6] *Titsch*, termine locale corrispondente a *Deutsch* 'tedesco', è il glottonimo comunemente usato dalla comunità per riferirsi alla propria parlata, tradizionalmente in opposizione a *wälsch*, cioè italiano. Il tedesco (standard) viene invece denominato *rächti titsch* 'tedesco giusto' (Dal Negro 2004: 67-68).

[7] Ovviamente nel caso italiano (o più in generale europeo) non si tratta di comunità ad oralità primaria (cioè analfabete), nel senso che oggi tutti i membri che ne fanno parte sono comunque alfabetizzati nella lingua di maggioranza. Va anche osservato, con Cardona (1983: 25), che "[u]na volta instaurata la pratica dello scrivere, si avvia nella comunità un processo irreversibile; la forma scritta assume un valore definitorio e irrinunciabile,

quelli che dispongono di tradizione scritta anche illustre) e, *a fortiori*, per le alloglossie senza copertura linguistica (cosiddette *dachlos*), che non si riconoscono cioè in nessuna lingua standard di riferimento, come è il caso delle parlate walser di cui trattiamo in queste pagine.

Il fatto, dunque, di essere lingue *solo* parlate, e in particolare parlate in contesti dialogici di vicinanza e familiarità tra interlocutori, cioè di *kommunikative Nähe* nei termini di Koch e Oesterreicher (2007), dovrebbe costituire un aspetto centrale della documentazione, e poi della descrizione, di lingue piccole e piccolissime. Come è noto (Voghera 2017; ma già Cardona 1983), la modalità, cioè il meccanismo semiotico dato dalla combinazione di canale (fonico-acustico, mimico-gestuale, grafico-visivo), grado di interattività tra produttore e ricevente del messaggio, e tempi di produzione e ricezione (sincroni o asincroni), si accompagna a correlati di tipo funzionale che determinano quli strutture linguistiche siano più efficaci, o comunque più attese. Di conseguenza, le lingue primariamente orali rivestono un interesse teorico notevole nella prospettiva dell'evoluzione di strutture grammaticali che siano compatibili con un uso della lingua esclusivamente orale, e anche sotto questa angolatura andrebbero forse valutati fenomeni linguistici apparentemente 'esotici' spesso riscontrati in lingue piccole e senza tradizione scritta[8]. Inoltre, dal momento che la componente interattiva è centrale per una definizione di modalità, aspetti sociolinguistici relativi alla densità e compattezza territoriale delle reti sociali entro le quali la lingua può essere effettivamente usata potrebbero svolgere un ruolo non secondario in alcune aree particolarmente sensibili del sistema, come ad esempio la codifica linguistica della deissi (per il contesto italiano cfr. ad esempio Prandi 2015; Dal Negro 2018) e, più in generale, di altri fenomeni relativi alla pragmatica e al discorso (cfr. anche Grenoble 2014).

Tutto ciò dovrebbe portare a dare priorità alla documentazione dell'uso parlato, e ad una riflessione che non sia amodale, che tenga cioè in conto delle specificità di produzione e ricezione della lingua in questione, al netto delle oggettive difficoltà, cui si accennava sopra. Difficoltà che riguardano in primo luogo l'elicitazione di parlato naturale in contesti di *language shift*, ma anche relative al trattamento di dati di parlato, e cioè alla trascrizione e annotazione, passi necessari per la compilazione di un *corpus*, ma particolarmente onerosi nel caso di lingue di questo tipo (Adamou 2016: 15-16).

I due esempi che seguono riportano produzioni linguistiche della stessa persona, una donna anziana e molto competente nella varietà di walser formazzino, elicitati in contesti e con obiettivi molto diversi. In un caso (esempio 3) si tratta di un compito traduttivo (dall'italiano al walser) proposto da me stessa, studiosa esterna e (all'epoca) sconosciuta alla comunità; in quel contesto è il figlio a fare da tramite. L'elicitazione di un parlato totalmente decontestualizzato, più vicino allo scritto (si noti il riferimento a un questionario: *devi dire come c'è scritto qui*), o comunque astrattamente amodale, mette la donna in serie difficoltà, anche solo per comprendere quello che le viene richiesto di fare[9].

---

giungendo a costituire il modello ideale di ogni produzione, anche orale". Da questo punto di vista, la lingua minoritaria, non scritta, viene quindi automaticamente svalutata rispetto alla lingua maggioritaria, scritta e normalizzata.

[8] Discutendo dei riflessi linguistici del *continuum* scritto-parlato in "highly literate societies", Coulmas (2013: 50) osserva giustamente che "[a] grammar of a language that has no written form – to which, obviously, the notion of an oral-literate continuum is not applicable – is an altogether different matter".

[9] Si noti che non è sempre questo il caso e una frequentazione intensa fra ricercatore e informatore può portare quest'ultimo a trasformarsi in una sorta di "professionista" di indagini linguistiche, certamente facilitando il compito di chi conduce l'inchiesta, ma costringendo comunque sempre ad interrogarsi seriamente sulla natura del dato elicitato (Iannàccaro 2001).

Nell’altro caso (esemplificato dall’estratto 4), la stessa parlante (indicata come L) è coinvolta da due conoscenti (una coetanea e la figlia di questa, incaricata di registrare e qui contrassegnata con R) nella narrazione di eventi realmente accaduti ai tempi in cui l’informatrice era una ragazzina. Gli eventi raccontati sono inseriti in una fitta rete di fatti e di persone più o meno noti, ma al tempo stesso già parte di quell’insieme di racconti basati su credenze e superstizioni, tramandati di generazione in generazione in ambito famigliare e di vicinato[10]. Nel caso specifico, l’estratto è parte di un *corpus* di conversazioni raccolte per un progetto di documentazione linguistica delle parlate walser (Dal Negro 2006) che vedeva coinvolti in prima persona membri delle comunità stesse, incaricati di reperire informanti, registrare conversazioni e, in parte, trascrivere i dati, cui si è accennato sopra a proposito dell’esempio (1). A differenza delle difficoltà osservate in (3), il frammento di discorso trascritto in (4) presenta un parlato fluente e linguisticamente complesso nel quale si notano anche tratti interessanti sul piano pragmatico quali l’uso dell’articolo con i nomi propri e il genere neutro nelle forme di accordo con referenti (noti) di sesso femminile.

(3) Int: [LEGGE DA UN FOGLIO STAMPATO] la mamma va in chiesa
Figlio: mü/ das müsse in pumattertitsch sägä jetz
‘questo si dovrebbe dire in titsch formazzino adesso’
Madre: ja, wir gängen zer mess
‘sì, noi andiamo a messa’
Figlio: nee aber dü möss säge wi hir isch gschribes: *la mamma va in chiesa*
‘no, ma tu devi dire come c’è scritto qui: ...’
Madre: ich ga zer mess
‘io vado alla messa’
Figlio: nee, dü möss
‘no, tu devi’
Madre: chappala
‘chiesa’
Figlio: dü möss säge di möter get in chappala
‘tu devi dire la mamma va in chiesa’
Madre: eh di möter get in chappala
‘eh la mamma va in chiesa’

(4) L: un das éscht äs zeichä ksé
‘e questo è stato un segno’
R: un der tag wa t mötär het kheiratä
‘e il giorno in cui si è sposata la mamma’
L: wa t mötär het kheiratä désch tagsch éscht är kschtorbä
‘quando si è sposata la mamma, quel giorno lui è morto’
L: un eis jarsch [TOSSISCE] äs éscht im tüsukdrihu/ tüsuknihunnärtdriutrisk ksé, dö hentsch z indschum hüs tanzut, éch bé drizäjärigs ksé, hentsch z indschum hüs tanzut un da sintsch ksé di di metjé älli da, un da éscht ksé z Pitsch Dschuditti, ér het=s nit pchennt, dü hescht=s chunnä pchennä, un hen tanzut un ... di ganz nacht im altä jar, un.. un im niwä jar, dö ... és hei kgangä z Dschuditti, wé älli, un éscht ärséchät, het der schtéch percho, in sébä tagu és kschtorbä.
‘e un anno, era il 1933, allora si ballava [per il Capodanno] in casa nostra, io avevo tredici anni, si ballava in casa nostra e là c’erano tutte le ragazze, c’era la Giuditta del P., voi non l’avete conosciuta, tu hai potuto conoscerla e hanno ballato tutta la notte

[10] Alcuni dei quali poi raccolti in pubblicazioni locali a scopo didattico o documentaristico (Zertanna e Dal Negro 2015).

nell'anno vecchio e nell'anno nuovo, poi la Giuditta è andata a casa, come tutti. E si è ammalata, ha preso la polmonite, dopo sette giorni è morta'

Il monolinguismo

Un secondo *bias*, molto radicato nella tradizione e nelle pratiche di documentazione linguistica, è quello del monolinguismo[11], e cioè il *focus* selettivo su una singola lingua oggetto di documentazione, come se questa esistesse in un *vacuum* sociolinguistico. Per definizione, invece, le lingue minoritarie fanno sempre parte di repertori multilingui dei quali normalmente costituiscono il tassello più fragile: l'attenzione andrebbe dunque spostata dalla lingua in sé (in isolamento) alle dinamiche tra le lingue e alla struttura di questi repertori complessi, al limite per individuare quali siano gli spazi comunicativi naturali della lingua minacciata (e soprattutto per verificare se ce ne siano ancora). Con le parole di Grenoble e Martin (2023: 266-267):

> The documentation of endangered languages has to date primarily focused on the creation of monolingual documentary corpora […] This research, although invaluable, may often fail to document the larger language ecosystem, the multilingual setting in which language shift occurs.

Tra gli obiettivi di un progetto di documentazione linguistica dovrebbero perciò rientrare anche gli usi mistilingui, la cui varia fenomenologia (stili alternanti, insertivi, interdizione degli stessi in determinati contesti, ecc.) costituisce un tratto costitutivo delle diverse comunità linguistiche, non di rado caratterizzandole in maniera significativa.

Purtroppo, questo *bias* è spesso difficile da evitare per diverse ragioni, non tutte riconducibili alla volontà di chi compie la ricerca. Una prima ragione ha a che fare con quanto già detto a proposito della difficoltà a documentare il parlato conversazionale in contesti naturali, cioè dove più ci si aspetta di riscontrare l'uso combinato di più lingue. Chiaramente, un maggiore controllo dell'informante sul suo parlato comporta anche un maggiore controllo sulla scelta della lingua da usare che, nel caso di compiti elicitativi molto mirati (si pensi alla traduzione di frasi), coincide con l'uso monolingue della lingua da documentare. Questo dipende in parte dalle aspettative del ricercatore-rilevatore, ma soprattutto dall'immagine che il parlante vuole dare della propria lingua, e solo di quella, e di se stesso come parlante linguisticamente 'leale' (*loyal*) verso la lingua ereditata, a maggior ragione in quanto minacciata di estinzione. Tale atteggiamento, spesso eccessivamente purista, è ben noto a chi lavora nel campo delle *endangered languages* (Dorian 1981) e riguarda, oltre a *code-switching* e *code-mixing*, anche prestiti e *insertions* lessicali. Questi vengono spesso 'glossati' oralmente come negli estratti (5), (6) e (7), quasi in forma di giustificazione, da ricondurre addirittura alla storia pluricentenaria dell'insediamento alloglotto, come in (6).

(5) dopo woljontsch gucku oi *la tele/ ti dico la televisione*
'poi vogliono anche guardare la tele/ ti dico la televisione'
[Inchieste sul campo a Rimella: descrizione di vignette]

(6) mh te/ … hier sägä *televisiun* … wäge … eh das ischt äso …. wenn eh wenn sind di fa/ fom de/ fa der schwitz hier cho in pomatt, de sin no chei fernseh gsi, *televisione* … di roschtog, modernisch roschtog isch da cheine ksé, un de hen di lit nit gwisst was ischt *television*

[11] Un *bias* che, a dire il vero, è sempre più spesso messo in discussione (Lüpke 2010; Lüpke e Storch 2013; Adamou 2016; Dal Negro 2021; Good 2023; Grenoble e Martin 2023).

'mh te/ qui diciamo televisiun perché eh è così: quando [i walser] sono venuti qui a Formazza dalla Svizzera, allora non c'era la televisione, non c'erano le cose moderne e allora la gente non sapeva cosa fosse la televisione'
[Inchieste sul campo a Formazza: traduzione di frasi]

(7) P: un bé z pomatt giborä der zwettu ökschtä *millenovecentoquarantuno*, *si può fare un po'* titsch *e un po'* wältsch?
'e sono nata a Formazza il due agosto 1941, si può fare un po' tedesco e un po' italiano?'
R: ja, *si può* [RIDE]
[Archivio Sonoro Walser: Formazza]

Viceversa, la volontà di evitare prestiti, anche occasionali, nello svolgimento di compiti traduttivi percepiti come complessi in quanto distanti dall'uso linguistico reale (come si è già osservato sopra a proposito dell'esempio 3), porta il parlante insicuro a scegliere alternative lessicali, come nel caso della risposta data dall'informatrice formazzina ad uno degli stimoli del progetto AlpiLink (Rabanus et al. 2023) e riportato qui come esempio (8). Si noti che le difficoltà riscontrate dai parlanti possono essere anche inattese e difficilmente prevedibili in fase di progettazione di un questionario, soprattutto se pensato per elicitare dati provenienti da contesti linguistici e sociolinguistici molto diversi fra loro, come è il caso di AlpiLink[12].

(8) stimolo da tradurre: *in piazza non ci sono alberi alti*
im dorf sén nit .. gross eh .. tanne
'in paese non ci sono grandi abeti'
[S19_wae_U0038 (F_70_Formazza)]

In questo brevissimo esempio una delle difficoltà riscontrate dall'informante è di tipo referenziale, per cui la donna commenta, prima di tradurre[13], *"piazza" non l'abbiamo*, nel senso che sembra proprio mancare il referente nell'insediamento tradizionale di Formazza. La parlante risolve il problema con un'alternativa che ritiene passabile e che le permette di evitare un prestito dall'italiano: *dorf* 'villaggio, frazione'. Una seconda difficoltà è rappresentata da una parola apparentemente non problematica il cui referente in questo caso dovrebbe essere ben presente nell'ambiente naturale circostante: 'alberi' (nel walser locale *beim*). In quello che sembra un paradosso, dopo qualche incertezza la parlante seleziona invece un iponimo, *tanne* 'abeti', corrispondente all'essenza arborea più diffusa sul territorio. L'abete è dunque l'albero per eccellenza, quello di cui è più probabile che si parli nel quotidiano di una lingua che si usa quasi esclusivamente per riferirsi al contesto più immediato e non per parlare in astratto, ad esempio di alberi.

Discussione e conclusioni

In queste pagine si sono messe in evidenza soprattutto le insidie sottese alle attività di documentazione linguistica, e più in generale alla ricerca sul campo, soprattutto se svolta in contesti caratterizzati da *language shift*, come è il caso delle comunità di minoranza walser nell'Italia nord-occidentale. Le conclusioni che se ne possono trarre sono che una documentazione linguistica da intendersi in senso stretto (*à la* Himmelmann 1998) di lingue come il walser sarebbe da escludersi a priori. Tuttavia, è

12 Scopo del progetto AlpiLink (cfr. https://alpilink.it/) è la raccolta di dati comparabili nel più alto numero possibile di varietà dialettali e alloglotte dell'Italia settentrionale. Per questo motivo la scelta degli stimoli (frasi da tradurre, descrizioni di immagini e altri task di elicitazione più mirati) non poteva tenere conto dei diversi contesti culturali e sociolinguistici nei quali la raccolta dati si sarebbe svolta.

13 Lo svolgimento del questionario online è stato effettuato con la mia assistenza nell'agosto 2023 per cui ho potuto verificare in prima persona le difficoltà incontrate dalla parlante (di fatto una parlante abbastanza fluente di 70 anni).

proprio nell'ambito delle comunità linguistiche più fragili e meno vitali (come le diverse alloglossie in Italia) che una qualche forma di documentazione linguistica sarebbe auspicabile e forse anche possibile con alcuni accorgimenti di cui proverò a fornire qualche esempio partendo dai punti discussi nelle sezioni precedenti: naturalezza del dato, centralità della modalità parlata, rilevanza del plurilinguismo.

Una documentazione delle pratiche discorsive interne alla comunità che sia quanto più naturale possibile porta a interrogarsi su che cosa abbia sostituito la lingua minoritaria nel momento in cui questa è uscita dall'uso. Le variabili in gioco, anche limitandosi al panorama italiano, non sono poche e includono la lingua nazionale, le varietà regionali della stessa, i dialetti italoromanzi o altre lingue di minoranza parlate nella regione e dotate di maggior prestigio e diffusione. Un aspetto particolarmente interessante di cui tenere conto riguarda il grado di assimilazione etnico-culturale della comunità dopo che la lingua minoritaria ha smesso di essere il principale mezzo di comunicazione quotidiano, anche come *we-code*, cioè internamente alla comunità. La volontà (non necessariamente esplicitata) di mantenere la comunità comunque distinta dal contesto sociolinguistico circostante può tradursi nello sviluppo di tratti linguistici marcati nella varietà locale della lingua di maggioranza. Ma il desiderio di distinzione può essere anche molto consapevole e arrivare a permeare atteggiamenti e discorsi, oltre a lasciare tracce visibili nello spazio pubblico (ad esempio nel paesaggio linguistico).

Nel caso in cui esista ancora una comunità di parlanti, per quanto minuscola e ridotta a pochi nuclei familiari, parallelamente al lavoro di elicitazione di dati direttamente finalizzati alla descrizione di fenomeni linguistici mirati, dovrebbe essere prioritario cercare di documentare anche interazioni che siano quanto più possibile vicine al polo del dato naturale. La registrazione di dialoghi (semi)spontanei fra parlanti (e non solo fra parlante e rilevatore) permetterebbe di tenere traccia della lingua parlata in contesto naturale, magari rivelando fenomeni linguistici inattesi.

Per quanto riguarda i *task* elicitativi, anche per stimolare il dialogo fra parlanti andrebbero esplorate tecniche che si avvicinino al vissuto della comunità al fine di rendere il compito più significativo. Ad esempio, per l'elicitazione di forme pronominali e di altre espressioni linguistiche usate per il riferimento personale, le fotografie di famiglia si prestano certamente meglio della descrizione di vignette o di albi illustrati nei quali sono rappresentati personaggi sconosciuti o immaginari. Vedute paesaggistiche locali si prestano per compiti descrittivi, ad esempio per l'elicitazione di preposizioni locative e di verbi stativi e di movimento, meglio di illustrazioni create *ad hoc* o di fotografie di luoghi non conosciuti. Si tratta, in questi casi, di aumentare la motivazione da parte del parlante, soprattutto se anziano, fornendo degli stimoli significativi e non astratti e lontani dal proprio vissuto.

In conclusione, quando il tempo è limitato – e non tanto quello da dedicare alla ricerca sul campo, quanto il tempo che resta prima che la lingua si estingua del tutto – la documentazione è necessariamente varia, frammentaria e non sempre ottimale: ma, come si sa, ogni dato è comunque migliore di nessun dato. Diventa a questo punto cruciale la responsabilità di chi raccoglie affinché i dati siano non solo archiviati e resi disponibili per future ricerche, ma anche corredati di tutte le informazioni relative al contesto di elicitazione, necessarie per farne in seguito un uso consapevole.

Bibliografia

ADAMOU, Evangelia (2016) *A Corpus-Driven Approach to Language Contact: Endangered Languages in a Comparative Perspective*. Berlin-Boston: De Gruyter.

CARDONA, Giorgio Raimondo (1983) "Culture dell'oralità e culture della scrittura". In ASOR ROSA, Alberto (ed.), *Letteratura italiana. Vol. II.* Pp. 25-101. Torino: Einaudi.

CICCOLONE, Simone, and Silvia DAL NEGRO (2021) *Comunità bilingui e lingue in contatto*. Cesena: Caissa Italia.

COULMAS, Florian (2013) *Writing and Society: An Introduction*. Cambridge: Cambridge University Press.

DAL NEGRO, Silvia (2004) *The Decay of a Language: The Case of a German Dialect in the Italian Alps*. Bern: Lang.

DAL NEGRO, Silvia (ed.) (2006) *Parlare walser in Piemonte: archivio sonoro delle parlate walser*. Vercelli: Mercurio.

DAL NEGRO, Silvia (2011) "Walser, comunità". In SIMONE, Raffaele (ed.), *Enciclopedia dell'Italiano*. Pp. 618-621. Roma: Istituto dell'Enciclopedia Italiana.

DAL NEGRO, Silvia (2018) "Lingue minori e deissi sociale". In COSTAMAGNA, Lidia, Elisa DI DOMENICO, Alejandro MARCACCIO, Stefania SCAGLIONE, and Barbara TURCHETTA (eds.), *Mutamento linguistico e biodiversità. Atti del XLI Convegno della Società Italiana di Glottologia*. Pp. 45-66. Pisa: Il Calamo.

DAL NEGRO, Silvia (2020) "The dilemmas of 'saving' a dying language: Walser German between documentation and planning", *Language Problems and Language Planning* 44(3): 273-286.

DAL NEGRO, Silvia (2021) "Documentazione linguistica e lingue in contatto". In FAVILLA. Maria Elena, and Sabrina MACHETTI (eds.), *Lingue in contatto e linguistica applicata: Individui e società*. Pp. 9-22. Milano: Officinaventuno.

DORIAN, Nancy C. (1981) *Language death: the life cycle of a Scottish Gaelic dialect*. Philadelphia: University of Pennsylvania Press.

FIORENTINI, Ilaria (2022), *Sociolinguistica delle minoranze in Italia*: *un'introduzione*. Roma: Carocci.

GOOD, Jeff (2023) "Adapting methods of language documentation to multilingual settings", *Journal of Language Contact* 15(2): 341-375.

GRENOBLE, Lenore A., and Jack B. MARTIN (2023) "Documenting multilingualism and contact", *Journal of Language Contact* 15(2): 265-270.

GRENOBLE, Lenore A. (2014) "The importance and challenges of documenting pragmatics", *Language Documentation and Description* 12: 145-162.

HIMMELMANN, Nikolaus P. (1998) "Documentary and descriptive linguistics", *Linguistics* 36: 161-195.

IANNÀCCARO, Gabriele (2001) "Alla ricerca del dato". In ALBANO LEONI, Federico, Eleonora STENTA KROSBAKKEN, Rosanna SORNICOLA, and Carolina STROMBOLI (eds.), *Dati empirici e teorie linguistiche. Atti del XXXIII Congresso della Società di Linguistica Italiana*. Pp. 23-36. Roma: Bulzoni.

IANNÀCCARO, Gabriele (2002) *Il dialetto percepito: sulla reazione di parlanti di fronte al cambio linguistico*. Alessandria: dell'Orso.

IANNÀCCARO, Gabriele, Federica GUERINI, and Silvia DAL NEGRO (2015) "Esperienze spontanee di elaborazione ortografica: premesse per un'analisi (socio)linguistica". In DAL NEGRO, Silvia, Federica GUERINI, and Gabriele IANNÀCCARO (eds.), *Elaborazione ortografica delle varietà non standard*. Pp. 7-50. Bergamo: Sestante.

KLAMER, Marian, and Francesca MORO (2020), "What is "natural" speech? Comparing free narratives and Frog stories in Indonesia", *Language Documentation & Conservation* 14: 238-313.

KOCH, Peter, and Wulf OESTERREICHER (2007), "Schriftlichkeit und kommunikative Distanz", *Zeitschrift für Germanistische Linguistik* 35(3): 346-375.

LÜPKE, Friederike, and Anne STORCH (2013) *Repertoires and Choices in African Languages*. Boston-Berlin, De Gruyter.

LÜPKE, Friederike (2010) "Research methods in language documentation", *Language Documentation and Description* 7: 55-104.

PRANDI, Michele (2015) "Varieties in Italy 2: Alpine Varieties". In JUNGBLUTH Konstanze, and Federica DA MILANO (eds.), *Manual of Deixis in Romance Languages*. Pp. 114-139. Boston-Berlin, De Gruyter.

RABANUS, Stefan, Anne KRUIJT, Birgit ALBER, Ermenegildo BIDESE, Livio GAETA, and Gianmario RAIMONDI. 2024. *AlpiLinK Corpus 1.0.4*. DOI: 10.5281/zenodo.10533936.

VOGHERA, Miriam (2017) *Dal parlato alla grammatica: costruzione e forma dei testi spontanei*. Roma: Carocci.

UNESCO Ad Hoc Expert Group on Endangered Languages (2003) *Language Vitality and Endangerment*. https://unesdoc.unesco.org/ark:/48223/pf0000183699.

ZERTANNA, Jolanda, and Silvia DAL NEGRO (2015) *Filastrocche, proverbi e racconti nel titsch di Formazza*. Domodossola: Grossi.

ZINSLI, Paul (2002) *Walser Volkstum, in der Schweiz, in Vorarlberg, Liechtenstein und Italien: Erbe, Dasein, Wesen*. 7. ergänzte Auflage. Chur: Bündner Monatsblatt.

Some preliminary observations on the preparation of a Saho treebank

Andrea Di Manno Università di Napoli «L'Orientale» – adimanno@unior.it
Paolo Milizia Università di Napoli «L'Orientale» – pmilizia@unior.it

ABSTRACT

In this paper we present some preliminary remarks on the preparation of a morphosyntactically annotated treebank for Saho within the Universal Dependencies (UD) framework, an annotation scheme used in a large collection of corpora, where Cushitic languages are still marginally represented. We discuss some issues concerning tokenization, morphology, and syntax, with particular reference to cases where deviations from the descriptions proposed in the literature on Saho are deemed necessary, in order to accommodate Saho data within the UD architecture, and cases where some peculiarities of the Saho language require an expansion of the inventories put forward in UD.

Keywords: *Saho language; treebank; syntax; morphosyntax; Universal Dependencies*

ISO 639-3 code: ssy

DOI: 10.23814/ethn.20.24.dim-mil

1. Introduction[1]

Saho is a Cushitic language spoken in central and south-eastern Eritrea and in the neighbouring area of Tigray, Ethiopia, and together with Afar constitutes the Saho-Afar group within Eastern Cushitic. The two languages form a dialectal continuum, and three major groups can be identified for Saho: Northern Saho, Central Saho, and Southern Saho. The latter is the closest to Afar, while for Central Saho one distinguishes varieties that share isoglosses with the northern varieties (CS1) and varieties that share features with the southern ones (CS2). Though in the last decade several contributions have improved our understanding of the language[2] and provided new material (including the Saho Corpus, cf. Jama Musse Jama 2022), a syntactically annotated treebank of the language is still missing.

This paper presents some preliminary remarks on the preparation of such a treebank within the Universal Dependencies (UD) framework (de Marneffe et al. 2021). UD adopts a dependency grammar perspective, meaning that it presumes direct links between linguistic units. Since it is based on a lexicalist view of syntax, dependency relations hold between words and there is no attempt to segment words into morphemes. Each word has a head on which it depends (except for the main predicate, which is considered the root) and can have one, zero or multiple dependents. In diagrams, dependency is represented through an arrow from the head to the (head of the) dependent and relation labels indicate the syntactic functions. Although providing a closed (universal) inventory of morphosyntactic categories for consistent cross-linguistic

[1] The authors thank Axmadsacad Maxammad Cumar, Giorgio Banti and Moreno Vergari for their support in the preparation of the present paper. Of course, the authors alone are responsible for all errors and omissions. The paper results from joint work by the authors. Paolo Milizia is responsible of Sections 2 and 3; Andrea Di Manno is responsible for the rest of the paper.

[2] See Banti – Vergari 2010 and Banti – Vergari 2023: 294-95 for previous contributions on the language. A comprehensive bibliography on the language and the people is in Vergari et al. 2023: 31-50.

annotation, UD allows language-specific extensions for the treatment of language-specific phenomena. Being an ongoing project, new treebanks are constantly added and, consequently, the guidelines keep being updated in order to deal with previously unattested phenomena. Although in the current version (2.13) UD consists of 259 treebanks in 148 languages, Cushitic languages are underrepresented, since only one treebank for Beja (Kahane et al. 2021) is available.

Annotations are encoded in plain text files in the CoNLL-U format. Each sentence consists of one or more word lines, and each word line contain the following ten fields: ID (the word index in the sentence), FORM (the word form), LEMMA (the lemma of the word form), UPOS (the universal part-of-speech tag), XPOS (the optional language-specific part-of-speech tag), FEATS (the list of morphological features), HEAD (the word index of the head of the current word), DEPS (the enhanced dependency graph, with additional relations and nodes to make implicit relations between words more explicit) and MISC, containing any other annotation. We use the MISC field to provide an English gloss of the Saho word; complying to the UD guidelines, we place an underscore whenever a field is empty[3].

2. Tokenization and lemmatization

An issue concerning tokenization is how to deal with words which correspond to a fused sequence of personal pronoun plus postposition: e.g. *yedde* (or *yodde*) = *yi* + *dde* ‘on me’. A solution is to consider the fused word as a supertoken, e.g.:

```
1-2   yodde
1     yo    yi    PRON  _     Case=Acc|Number=Sing|Person=1|PronType=Prs
      _     _     _     _
2     dde   dde   ADP   _     _     1     case  _     Gloss=on
```

A similar treatment can be applied to the so-called “free forms” of postpositions. According to Banti – Vergari (2023: 311), these are special forms of postpositions which are not combined with a pronoun and which, in their interpretation, entail an understood third-person pronoun. E.g., besides *-dde* ‘on’, a free form *adde* is found meaning ‘on him/her/it’. According to a different analysis, such free postpositions can be viewed as portmanteau words comprising a bound pronoun *a-*/*aka-*/*aa-* plus the bound postposition[4]. Along these lines, an annotation of the following type may be proposed:

```
1-2   adde
1     a     a     PRON  _     Case=Acc|Number=Sing|Person=3|PronType=Prs
      _     _     _     _
2     dde   dde   ADP   _     _     1     case  _     Gloss=on
```

A further issue concerns the nominalizing suffix *-m*. It is important, indeed, to distinguish the instances in which a final *-m* can be considered as an element of the morphosyntactic structure of a phrase from those in which a word ending in *-m* can be considered, from the synchronic point of view, as a non-analysable element even if that final *-m* is arguably to be identified with the nominalizing *-m* from the etymological point of view. Thus, for instance, adverbs such as *mangum* ‘very much’ and *dagum* ‘a little, few’ are better treated as simple ADV elements, though they are clearly historically connected with the stative verbs *mango* ‘be much’ and *dago* ‘be few’ (cf.

[3] See https://universaldependencies.org/format.html All the cited urls have been last accessed in March 2024.
[4] This pronoun might be etymologically connected with the anaphoric demonstrative *ay* (Banti – Vergari 2023: 309), *aa* in Southern Saho (Esayas Tajebe 2015: 161), *a-* both in Afar and Northern Saho according to Morin (1995: 88).

Banti – Vergari 2005: 124) and probably contain the nominalizing *-m*, etymologically. Analogously, some postpositional phrases with *-h* which have become fixed expressions can be considered as single ADV words. As a rule of thumb, they can be treated as single elements if Vergari – Vergari (2003) has a dedicated entry for them, as occurs, e.g., in the case of *rummah*, 'truly', but literally 'with truth'.

Sticking to Vergari – Vergari (2003), nouns and pronouns are lemmatized using the absolute form; verbs of classes 1, 2 and 4 are lemmatized using the first person singular of their perfect, since it is closer to the verbal stem and can be easily recognized; verbs of class 3 are lemmatized according to the form of the third person singular of the present.

3. Features and feature values

We prefer not to introduce new parts of speech (*xpos*) and therefore, when relevant, differences in the sub-parts of speech will be dealt in the field FEATS, which, in the UD scheme, contains pieces of information about a word's part-of-speech and morphosyntactic properties.

In order to account for the complexity of the number system of Saho nouns, we propose to expand the feature value inventory of UD by adding the options "Number = Gnrl" for the general number and "Number = Sgtv" for the singulative (a value which has also been proposed for the Beja treebank[5]). Indeed, Saho nouns fall into two subsystems: one exhibits the common opposition "singular : plural"; the other opposes a general number form, which is unspecified as to the "singular : plural" opposition, to a corresponding singulative which is used when the noun refers to one item, but is structurally marked as compared to the general number form (see Zaborski 1986: 21-53; Banti – Vergari 2023: 307). As a rule of thumb, the subsystem to which a noun belongs can be inferred from Vergari and Vergari's dictionary (2003): if the English rendering of a noun exhibits an optional "(s)" marker or some analogous expedient (e.g. "*zizzaale nf* honeybee(s)") then the noun belongs to the "general : singulative" subsystem, otherwise it can be treated as a common "singular : plural" noun. It should be mentioned, however, that the status of the "general : singulative" opposition as inflectional rather than derivational may be questioned[6]. Note that agreement targets such as verbs only exhibit the opposition "singular : plural".

The core cases of Saho are nominative, absolute and genitive, for which, following the UD guidelines[7], we use the values Nom, Acc and Gen respectively.

As far as verbal morphology is concerned, according to the use of UD, the specifications "Tense = Pres"[8] and "Tense = Past" will be used: these will also be applied to simultaneous and sequential converbs, respectively.

We use the polarity feature to distinguish negative verb forms from positive ones: while "Polarity = Neg" should be mandatorily present with negative verb forms such as the negative past or the negative imperative, the complementary specification "Polarity = Pos" may be omitted throughout.

---

[5] In fact, the Beja treebank adds a singulative feature, which is used to distinguish collective nouns that designate a single entity with a boolean value (Sgtv = Yes if it is a singulative).
See https://universaldependencies.org/bej/index.html

[6] For instance a "singular : plural" noun such as *fan* (M.SG) : *faanon* (M.PL) can also have a corresponding singulative, as *fanta* (F). Note also that a "general : singulative" noun such as *zizaale* can have more than one singulative counterparts: *zizzaaletto* (M) and *zizzaalettö* (F).

[7] See https://universaldependencies.org/u/feat/Case.html

[8] Note that in UD the use of the label "Pres" is also prescribed for non-past tenses.

Verbs of classes 1, 2 and 4 exhibit a gender distinction in the third singular: this will be accounted for by the values Masc or Fem of the feature Gender, that will be omitted for other persons and for verbs of class III, which have a single word form for the third singular both masculine and feminine.

The feature specification "VerbForm = Rel" may be reserved for cases in which the relative verb is formally different from the corresponding indicative mood. Otherwise, the field DEPREL will suffice to disambiguate between relative and non-relative uses of verbs.

Similarly to what has been proposed for the standard Romanian infinitives and negative imperatives[9], for which a short and a long variant exist, the distinction between the short and the long negative past (e.g. *malifo* vs. *maalifinna*, both meaning 'he/she did not close') can be accounted for by using the feature specification "Variant = Short" and "Variant = Long". The specification "Variant = Short" may also be used for signalling other verb variants, such as the common short form *le* of *leya* 'have' (Class 3 stative verb), e.g.:

4 le leya VERB _ Mood = Ind|Number = Sing|Person = 3|Tense = Pres|**Variant = Short**|VerbClass = 3 5 acl:relcl _ Gloss = have

A further use of "Variant = Short" is with the short variants of the enclitic postpositions (e.g. *-d* 'on, in' instead of *-dde*; cf. Banti – Vergari 2005, 123f.), while "Variant = Long" is used for the emphatic absolute forms of personal pronouns (*yoyya* as opposed to *yi*, cf. Banti – Vergari 2005: 115; Banti – Vergari 2023: 308).

3.1. Features related to part-of-speech sub-classes

As concerns the part-of-speech classes, distinctions worth being accounted for are those between non-stative and stative verbs and between bound and free cardinal numerals. Indeed, stative verbs (class 3 verbs in Banti – Vergari 2023 subgrouping) can be considered as a distinct sub-part-of-speech since they have a reduced inflectional potential: they inflect only for the non-past indicative, the simultaneous converb and the infinitive and do not have separate forms for masculine and feminine third singular.
Although not directly separating stative verbs from non-stative ones, this difference can be accounted for by the VerbClass feature: the values range from 1 to 4, each indicating the corresponding verb class according to the subgrouping proposed in Banti – Vergari (2023).

5 mango mango VERB _ Mood = Ind|Number = Sing|Person = 3|Tense = Pres|**VerbClass = 3** 6 acl:relcl _ Gloss = be.much

As for numerals, we use the value Bound of the feature Variant to account for the distinction between free and bound elements (tagged as "Variant = Bound"). Since the distribution between the two types is syntactically determined, the appropriate tag can be assigned even if the numeral is a Roman or Arabic digit. E.g. for *5* in *5 iggida* 'five years':

_ 5 koona NUM _ NumForm = Digit|NumType = Card|**Variant = Bound** _ nummod _ Gloss = five

[9] See https://universaldependencies.org/ro/feat/Variant.html

It can be noted, incidentally, that there seems to be no need for a special class of ideophones, as far as UD morphosyntactic annotation is involved. Indeed, even if Banti – Vergari (2023: 300) posit such a word class for Saho, its members only appear in compound verbs and other derivatives and never show as autonomous syntactic nodes: an example is *sik* in *sik-erhxe* 'to be silent'. One may also add that the boundary between such ideophone class and the noun class is blurry since full-fledged nouns may also appear as first elements of compound verbs: e.g., the feminine noun *cafuu* 'forgiveness' appears in the compound verb *cafuu-erhxe* 'to forgive'. On the other hand, the ideophone *sik* also appears in the derived noun *sikko* 'silence'.

It is, moreover, advisable to stick to Banti's terminology (2010) and to classify as converbs not only the invariable converbs of Central and Southern Saho but also the subject-agreeing simultaneous converbs of Northern Saho. Indeed, treating them as participles, as in Banti – Vergari (2005: 105), would miss the fact that they are typically used as elements depending on the predicate of the matrix clause rather than on the element with which they agree: e.g. the structure *aliftak ku juble* opening.2SG 2SG.OBL saw.3SG.M 'he saw you while you were opening it' (cf. Banti – Vergari 2023) can be compared with the Italian infinitival clause of *ti ha visto aprirlo*, in which *ti… aprirlo* functions as a complement clause (cf. Mensching 2017: 382). This does not mean that the language is not able to express object clauses proper; as for the Irob dialect, an example is that given by Reinisch (1878: 18 = 104): *kāy yígdifa-m úbḙla* 3SG.M.OBL kill.PST.3SG.M-NMLZ see.PST.1SG 'I saw that he killed him', where the complement clause is marked by the presence of the nominalizing suffix *-m* on the verb (which is also its last word; see §5.2). As concerns the dependency relation between the converb and its head, it will be classified as advcl ("adverbial clause") independently of the issue of whether it forms a complement clause or a true adverbial clause, in accordance with the use prescribed in the UD guidelines, where, e.g., in the annotation of the sentence *They heard about you missing classes* the gerund *missing* is treated as an advcl dependent of the main verb[10].

Note that for converbs we use the "VerbForm = Conv" value and the agreeing or non-agreeing nature of a converb is inferable from the presence or absence of agreement features in the field FEATS.

4. Auxiliaries

According to UD guidelines, we posit a separate part-of-speech class for auxiliaries. The lexemes *ine* 'to be, exist', *kinni* 'to be', *leya* 'to have' and *waye* 'to lack, miss' belong here when used as copulas and when, together with the subjunctive, the infinitive or a converb, they form compound tenses.

Ine and *kinni* are the only two lexemes that can be used as copulas: e.g., in *roble abraahim barha kinni* 'Roble is Ibrahim's daughter' *kinni* is tagged as AUX and is a dependent, through the aux relation, of the nominal predicate *barha* 'daughter'.

The past tense of the class 1 verb *ine* is also used, together with the simultaneous converb, to form the past tense of class 3 verbs: e.g., in *kixinii yine* 'he loved' the converb is considered the root of the sentence, while *yine*, tagged with upos = AUX, depends on it, through the aux relation.

[10] See https://universaldependencies.org/u/dep/advcl.html

Class 3 verbs *kinni* and *leya* are used, together with subjunctive forms, for the two futures of Saho, and are accordingly considered auxiliaries: e.g., *yamaatoona kinon* / *linon* 'they will come'.

Finally, in negative relative clauses, while Northern Saho uses a special negative relative paradigm unmarked for time reference (these forms are marked as "VerbForm = Rel"), Southern Saho (like Afar) uses the infinitive with the auxiliary *waye*; the infinitive followed by a form of *waye* is also used for the negative jussive and to negate verbs in conditional sentences (Banti – Vergari 2023: 314).

As already stated, these lexemes are ambiguous between AUX and VERB: in the sentence *zizzaale fantat lacnale makaano kixina* 'Bees like places that have intermediate temperatures' the form *le* (in *lacnale*) is a VERB (see the CoNLL-U line in §3), depending, through the acl:relcl relation, on the feminine noun *makaano* 'place', and having the feminine noun *lacna* 'temperature' as its direct object dependent.

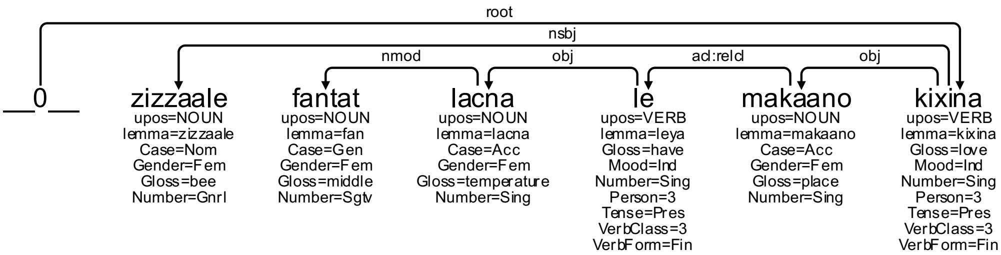


Figure 1. *zizzaale fantat lacnale makaano kixina* 'Bees like places that have intermediate temperatures' (from Banti - Axmadsacad Maxammad Cumar 2009: 106).

5. M-nominalisations

A relevant issue concerns the nominalizing suffix *-m* and the indefinite *tiya* m. (*tiyä* f., *mara* pl.), which are found both in what Parker and Hayward (1985: 287) call "M-nominalised clauses" and as the second element of independent pronoun forms. Although they are arguably historically interconnected, we propose to treat them separately.

5.1. Determiners and pronouns

We distinguish in Saho a class of determiners (upos = DET), which are only used as attributes and do not inflect for gender and number, from a class of pronouns (upos = PRON), which can be used independently. Each of these classes consists of possessive, demonstrative, indefinite and interrogative elements (distinguished through the different values of the feature "PronType" and through the feature "Poss", for possessives). Thus, for example, the possessives *ku* in *ku migac* 'your name' and *kuti* in *kuti yiggidile* 'yours is broken' are treated as two distinct lexemes (*ku* and *kutiya* respectively) with different POS tags:

1 ku ku DET _ Number = Sing|Person = 2|Poss = Yes|PronType = Prs 2 det:poss _ _
1 kuti kutiya PRON _ Case = Nom|Gender = Masc|Number = Sing|Number[psor] = Sing|Person = 2|Poss = Yes|PronType = Prs 2 nsubj _ _

As far as the feature layer is concerned, while possessive determiners are only tagged for the values of the possessor, possessive pronouns may have two different genders and numbers: that of the possessed object (triggering agreement on the verb if it is a subject) and that of the possessor[11], indicated by *psor* in square brackets after the relevant feature attribute. In this way we can account for the differences between, e. g., *teetiya* (Case = Acc|Gender = Masc|Gender[psor] = Fem|Number = Sing|Number[psor] = Sing|Person = 3|PronType = Prs|Poss = Yes) and *kaatiyä* (Case = Acc|Gender = Fem|Gender[psor] = Masc|Number = Sing|Number[psor] = Sing|Person = 3|PronType = Prs|Poss = Yes)[12]. In the plural there is no gender distinction, but according to Banti and Vergari (2005: 116-17) forms ending in *-mara* are used for human referents, while forms ending in *-m* are used for non-human ones. In our annotation scheme this can be accounted for by adding the feature "Animacy", with values "Hum" and "Nhum" for humans and non-humans respectively[13], so that *yimara* (Animacy = Hum|Case = Acc|Number = Plur|Number[psor] = Sing|Person = 1|PronType = Prs|Poss = Yes) can be distinguished from *yim* (Animacy = Nhum|Case = Acc|Number = Plur|Number[psor] = Sing|Person = 1|PronType = Prs|Poss = Yes). Some pronouns also show the "singulative : general" opposition, e.g. *aketto* 'another one, m.', *akettö* f., *akim* pl. (Banti – Vergari 2005: 119)[14].

5.2. Subordination

The nominalizing suffix *-m* and the indefinite *tiya* (*tiyä* f., *mara* pl.) are also found in structures corresponding to free relative clauses ("headless relative clauses" in Banti – Vergari 2023: 317) in other languages, although in Saho these can be used in a variety of contexts. The analysis we propose is somehow different from the ones found in previous literature, since we treat the two elements in a different way.

The indefinite *tiya* is analysed as an indefinite pronoun representing the element modified by the relative clause in cases like the one in Figure 2, where the verbal form *orbishshe* is the dependent of the pronoun *tiya* (nominative *ti*), through the acl:relcl relation.

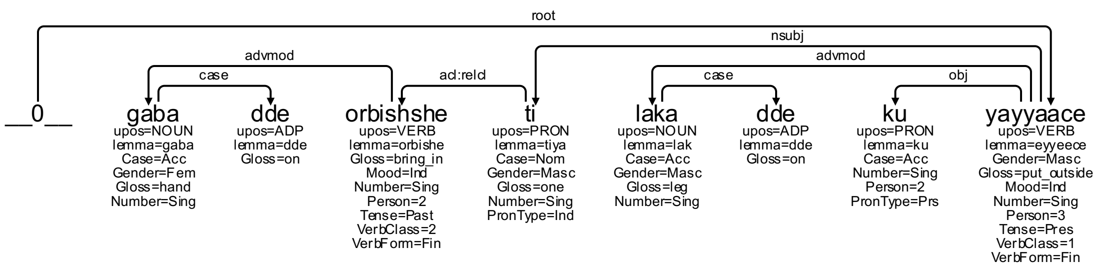


Figure 2. *Gabadde orbishshe ti lakadde ku yayyaashe* 'Who you brought in with your hand kicks you out with his foot' (from Banti – Vergari 2023: 312).

[11] The same distinction could apply to the Person feature, since the Person value tagged is always the possessor's one and hides the fact that nominative possessive pronouns always trigger a third person agreement on the verb. Nonetheless, we do not use it, since its usage is discouraged in UD: the issue is discussed in https://universaldependencies.org/u/overview/feat-layers.html

[12] Note that the "Gender[psor]" feature is relevant only when the possessor is a third person singular.

[13] In Irob (Southern Saho of Ethiopia), there seem to be a parallel distinction also in the singular, where the forms ending in *-tiya/-tiyä* are used for human referents, while a form in *-iyya* is used for non-human ones (Esayas Tajebe 2015: 150).

[14] The corresponding DET is *aki* 'other'. The singulative forms of the pronoun *aketto*, *akettö* seem lexicalized, since the element *-tto*, when it functions as a singulative suffix, does not trigger a change in the vowel, see e.g. *alaaki* 'specie(s) of shrub(s)' ~ *alaakitto* 'seed, fruit'. For similar phenomena in Afar see Parker and Hayward (1985: 237 n.; 238 n.2).

In this way the annotation does not differ from a relative clause modifying a noun (cf. Figure 1 in §4); in fact, the only difference seems to be that relative clauses with *tiya* can be used in pseudoclefts, as in Figure 3, where the copula, according to UD guidelines, is a dependent of the nominal predicate (in this case *miyatto*) and the pronoun in the nominative (*ti*) is considered the subject of the sentence.

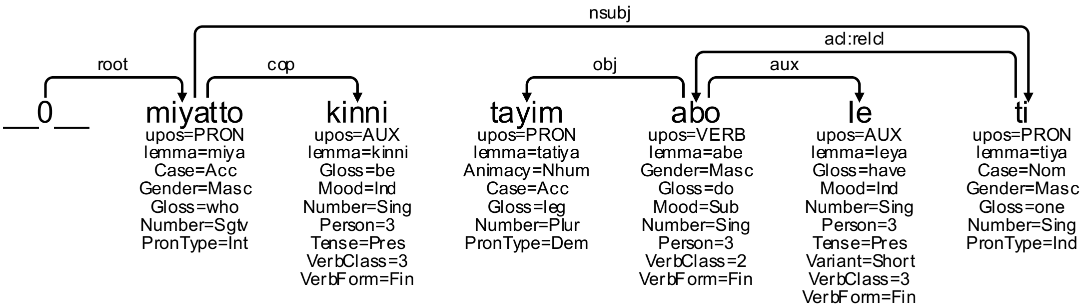


Figure 3. *Miyatto kinni tayim abo le ti* 'Who is it who will do this?' (from Banti – Vergari 2023: 314).

If a relative clause follows its head, *-ya* is suffixed to the verb: this is treated as a relativizing subordinating conjunction and considered a dependent of the verb through the mark relation, as in Figure 4.

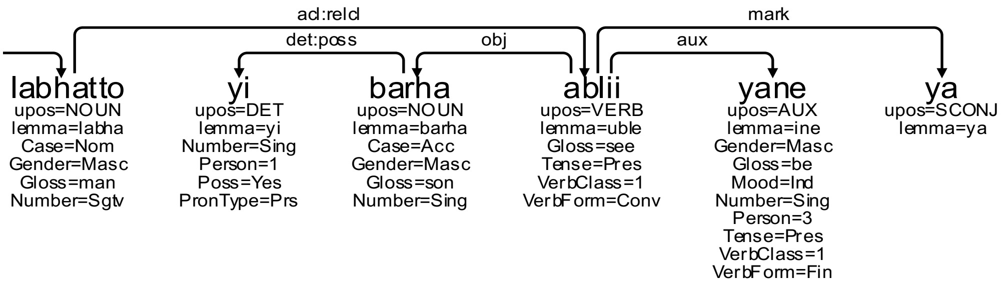


Figure 4. *Labhatto yi barha ablii yaneya* 'The man who is seeing my son' (from Banti – Vergari 2023: 317).

In a similar way, the nominalizing suffix *-m* is treated as a subordinating conjunction (SCONJ), as in Figure 5: note that a verbal form modified by *-m* can have a determiner as a dependent, as is the case with infinitives in Italian (*il dire*).

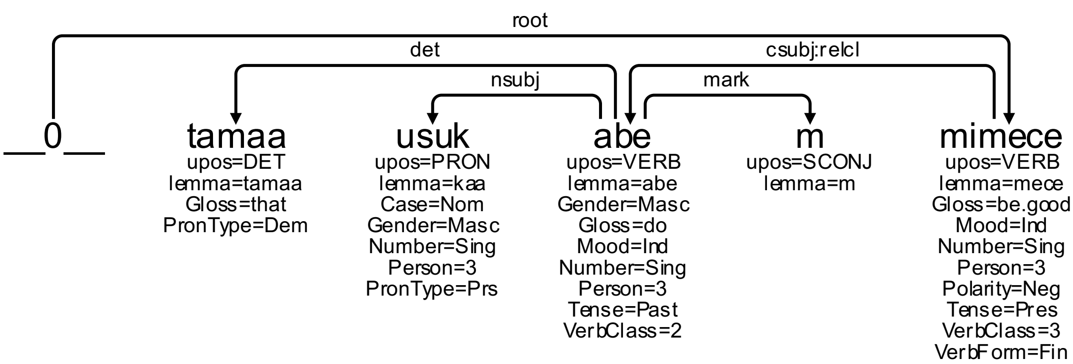


Figure 5. *Tamaa usuk abem mimece* 'What he did isn't good' (from Banti – Vergari 2023: 317).

Most importantly, these relative clauses can be used as sentential complements (see Figure 6 for a subjective clause, and Figure 7 for an objective clause): we use the dependency relations csubj:relcl, ccomp:relcl and xcomp:relcl in order to distinguish them from the complement clauses with the subjunctive (Figure 8). In UD the relation xcomp is reserved to clausal complements whose subject is controlled (that is, must be the same as the higher subject or object, with no other possible interpretation, as is the case in Figure 7 with *aba* and *farha*), while the ccomp relation is reserved to all the other cases (e.g., *mece* in Figure 7).

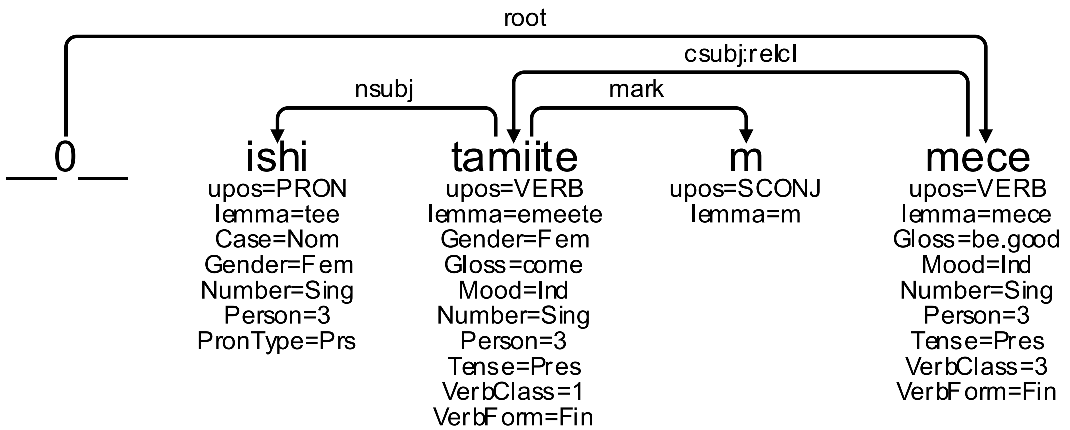


Figure 6. *Ishi tamiitem mece* 'It is good that she comes' (from Banti – Vergari 2023: 318).

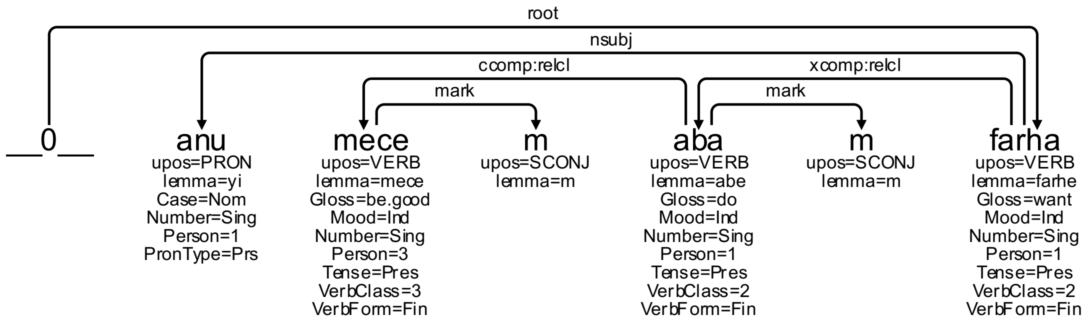


Figure 7. *Anu mecem abam farha* 'I want to do what is good' (from Banti – Vergari 2023: 318).

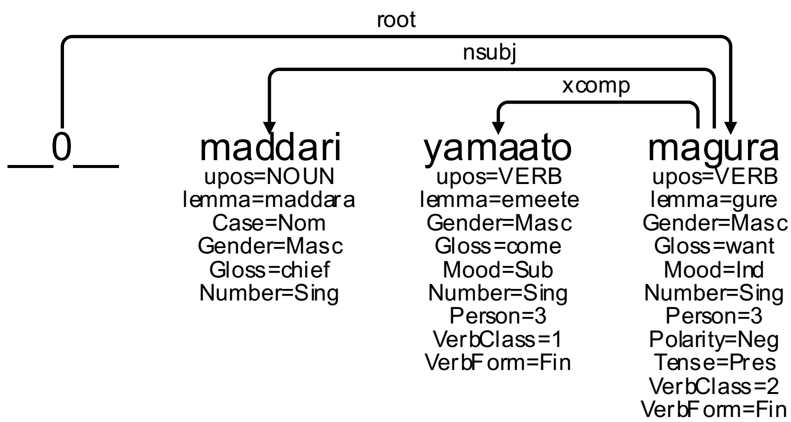


Figure 8. *Maddari yamaato magura* 'The chief doesn't want to come' (from Esayas Tajebe 2015: 327).

Followed by a postposition, free relative clauses with *-m* can function as subordinate adverbial clauses: in these cases, we use the advcl:relcl relation, and both the *-m* and the postposition are considered dependent on the verb, with the mark relation, as in Figure 9.

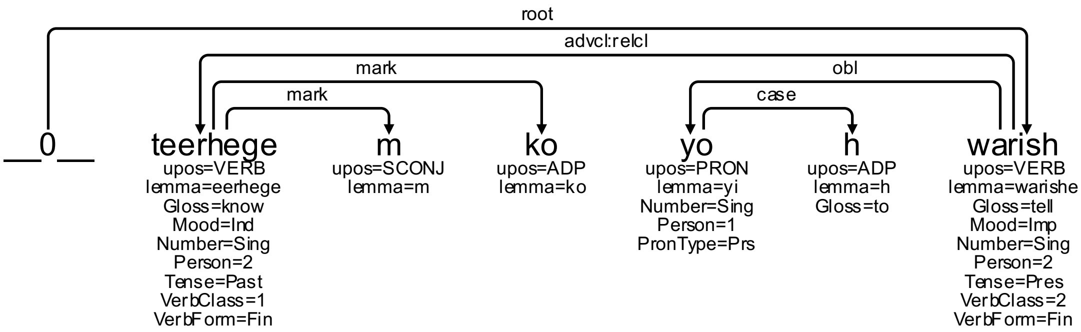


Figure 9. *Teerhegemko yoh warish* 'If you know it, tell me!' (from Banti – Vergari 2023: 318).

Sometimes the nominalizing *-m* is omitted (the phenomenon is widespread in Afar, according to Lamberti 1990: 155): in these cases, the verb of the subordinate clause is considered a dependent of the main verb through the advcl relation and the postposition receives the mark relation (Figure 10).

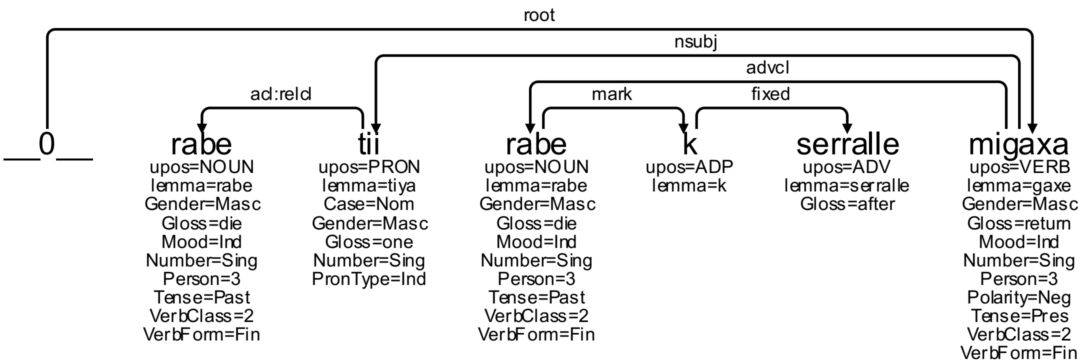


Figure 10. *Rabe tii rabek saralle migaxa* 'Whoever dies will not come back after death' (from Banti – Vergari 2023: 318).

Although temporal clauses with *gedda*/*ged* 'time' in Northern Saho or *gul* in Southern Saho can be treated as relative clauses modifying the noun *gedda/gul* (Lamberti 1990: 151-58; Banti – Vergari 2023: 318), we consider them adverbial clauses introduced by a subordinating conjunction *gedda/gul* 'when' (lemmatised as such in Vergari – Vergari 2003):

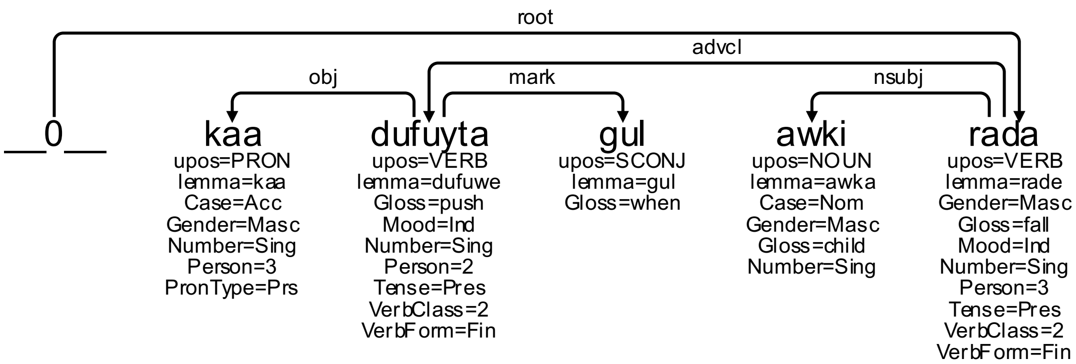


Figure 11. *Kaa dufuyta gul awki rada* 'When you push him, the child falls' (from Esayas Tajebe 2015: 288).

As already discussed in §3.1 temporal clauses are expressed also through converbs.

6. Final remarks

In this paper, we have addressed some preliminary issues on the preparation of a Saho treebank within the framework of Universal Dependencies. We have discussed some problems concerning tokenization, morphology, and syntax. The challenge, in this respect, was twofold. On the one hand, to account for morphological and part-of-speech peculiarities of the Saho language, we had to expand the feature and feature value inventories of UD. In this process, we tried to stick to features already used in other treebanks, especially in the Beja one, since Beja is the only Cushitic language at present available in UD. On the other hand, to stay within the architecture underlying the UD framework, significant deviations from the descriptions proposed in the literature were necessary in some cases.

This effort is useful and necessary both because it will enrich and increase the flexibility of UD, providing new material from a language belonging to an underrepresented family, and because it will make Saho data (for example, the material in the Saho Corpus; cf. Jama Musse Jama 2022) more accessible to the wider scientific community. Finally, although the specimen presented here primarily uses data from the Northern Saho variety described in Banti and Vergari (2023), it is hoped that a treebank of Saho-Afar could contain material from the different dialect groups of these languages: in the CoNLL-U format, the multilingual nature of such a treebank can be easily dealt with through sentence-level comments (lines starting with a hash), where metadata providing information about the dialectal provenance of a text can be stored.

References

BANTI, Giorgio (2010) “Remarks on the typology of converbs and their functional equivalents in East Cushitic”. In VÖLLMIN, Sasha, Azeb AMHA, Christian RAPOLD, and Silvia ZAUG-CORETTI (eds.), *Converbs, Medial Verbs, Clause Chaining and Related Issues*. Pp. 31-80. Köln: Köppe.

BANTI, Giorgio, and AXMADSACAD MAXAMMAD CUMAR (2009) “A few Saho texts about bees and honey”, *Ethnorêma* 5:89-108.

BANTI, Giorgio, and Moreno VERGARI (2005) “A Sketch of Saho Grammar”, *Journal of Eritrean Studies* 1-2: 100-131.

BANTI, Giorgio, and Moreno VERGARI (2023) “Saaho”. In MEYER, Ronny, Bedilu WAKJIRA, and Zelealem LEYEWM (eds.), *The Oxford Handbook of Ethiopian Languages*. Pp. 294-319. Oxford: Oxford University Press.

DE MARNEFFE, Marie-Catherine, Christopher MANNING, Joakim NIVRE, and Daniel ZEMAN (2021) “Universal Dependencies”, *Computational Linguistics*, 47(2): 255-308.

ESAYAS TAJEBE (2015) *Descriptive Grammar of Saaho.* PhD dissertation.

JAMA MUSSE JAMA (2022) “Saho Corpus: Semi-automation of Verb Conjugation in Saho: Verbs Class I”, *Ethnorêma* 18: 69-87.

KAHANE, Sylvain, Martine VANHOVE, Rayan ZIANE, and Bruno GUILLAUME (2021) “A morph-based and a word-based treebank for Beja”. In DAKOTA, Daniel, Kilian EVANG, and Sandra KÜBLER (eds.), *Proceedings of the 20th International Workshop on Treebanks and Linguistic Theories (TLT, SyntaxFest 2021).* Pp. 48–60. Sofia: Association for Computational Linguistics.

LAMBERTI, Marcello (1990) "Some word order principles of Saho-Afar", *Rassegna di Studi Etiopici* 34: 127-168.

MORIN, Didier (1995) *«Des paroles douces comme la soie». Introduction aux contes dans l'aire couchitique (bedja, afar, saho, somali).* Paris: Peeters.

REINISCH, Leo (1878) "Die Sahosprache", *Zeitschrift der Deutschen Morgenländischen Gesellschaft* 32: 415-464.

VERGARI, Moreno, and Roberta VERGARI (2003), *A basic Saho-English-Italian dictionary*. Asmara: Sabur Printing Services.

ZABORSKI, Andrzej (1986) *The Morphology of Nominal Plural in the Cushitic Languages*. Wien.

Identity, speech community, and language concepts in language documentation.
A sociolinguistic approach[1]

Cristina Muru
Università degli Studi della Tuscia
cristina.muru@unitus.it

ABSTRACT

Documentation of endangered languages has fostered a new discipline in linguistics, namely documentary linguistics and it has undergone many changes in its general approach in recent years. Mainly focussed on language description initially, it has evolved towards a deeper documentation of language practices within communities and, more recently, made a concrete contribution of linguists to speech communities for revitalisation of their languages, wherever this is the community's desire. Despite the changes that have occurred in the approach of many language documentation projects, it is still necessary to focus our attention on some issues which are crucial in the documentation of any language. Therefore, before moving towards the main issue of this paper – that is, what counts as language in language documentation – it is necessary to undertake a reflection on some basic concepts such as identity and speech community, as well as on the we-concept. The aim is to clarify for the reader the basic assumptions that I adopt in considering what counts as language in language documentation. I will start with the identity concept: therefore, I will discuss the concepts of 'we' and 'speech community', and I will propose some concluding remarks which offer further reflections on the language description and documentation framework.

Keywords: *identity, speech community, language, variety of language, language documentation and description, linguistic diversity, sociolinguistics, ethnographic approaches in language documentation.*

DOI: 10.23814/ethn.20.24.mur

1. On the concepts of identity and speech community

Identity has been widely discussed and, at the same time, deconstructed within anthropological debates. Authors such as Abu-Lughod (1991), Amselle (1999), Anderson (1983), Fabietti (2000) and Remotti (1996), have proposed the non-existence of an identity concept, at least in the way it was developed in the nineteenth century and

[1] I would like to express my deepest gratitude to Peter K. Austin for his comments on an earlier version of this paper, as well as for his bibliographical suggestions. I want also to acknowledge all the helpful conversations I had with him and other specialists on language documentation and rivitalization such as Candide Simard, Julia Sallabank, and Sheena Shah. Some of these took place during exchanges between University of Tuscia and SOAS, University of London, within the Erasmus programme (when Brexit was far away from our collective imagination). My thinking of this topic has also benefitted from feedback following my presentation at the 10th International Summer School in Language Documentation, at University "L'Orientale", Naples, 10-14 July 2023 (for further details, consult https://www.ethnorema.it/language-documentation/) where I introduced attendees to some of the arguments discussed in this paper. Finally, I thank the Ethnorêma editor and reviewers for their comments and criticisms, which helped me improve the paper for publication.

is often perceived nowadays.[2] Indeed, their arguments have largely demonstrated that identity is an atemporal concept which automatically frustrates itself.[3] Nevertheless, as (Western) human beings, we cannot avoid using the term 'identity' in referring to ourselves; consequently, expressions like "my identity is" or "I belong to", as opposed to "you are" and "you belong to", are very common. However, if we reflect on how the identity concept applies in our own daily life, we must admit that we continuously adapt ourselves to each of the specific moments or communicative events in which we are involved.[4] What really matters is that the specific moment in which we are involved will never be replicated in the same way. Consequently, while adapting to the surrounding circumstances, I can equally define myself as an individual human being and a total part of something (even though Pascal [1962] largely contributed to the demolition of the ego concept). Nevertheless, each time we interact with someone different, we move ourselves within a dimension other than that of our surrounding reality. For instance, I can perceive myself as a citizen of my town when I am in another town within "my" state, whereas I experience myself as a citizen of "my" state when I am in another geographical space far away from my home. Equally, I am a smiling and flexible person among friends and a rigid one when I am in the classroom. Consequently, how is it possible to define an identity concept? I think that we should admit that we are all what we need to be in a specific communicative event in order to interact efficiently with others – if we want! – or to ensure our face (cf. face-work theory by Goffman 1955; 1974 and Brown and Levinson 1987).

This also means that we can define ourselves and our identity only if we have considered the *we*-dimension beforehand, because we are always part of something (a community, a state, a group of friends, etc.). Indeed, we will not be able to define ourselves in any way if we do not consider what is around us, even when we are still convinced by the idea that we do not need others in order to understand who we are. As Remotti (2010: 65) underlines, this means that we cannot renounce the identity concept, even though it is a flexible, always changing, one, since it is a natural tendency of human beings to act as if it does exist. However, what we can do is be aware of the fact that it is a mere illusion to define our own identity concept within fixed boundaries as such and such.

At the same time, this means that the *we*-dimension is another construction since we adapt to different *we*-dimensions at the same time: we are friends, we are workers, we are speakers, we are religious, we are Italian, or we can become Italo-American. We always change the *we*-dimension to that to which we originally belonged. Furthermore, we are something or someone within our perception, while we can be someone else within the perception of others. Indeed, the way others perceive us not only redefines us but can also affect our own perception of who we are.

A similar approach can be adopted to disentangle the perception or the definition of a linguistic or speech community.[5] Silverstein (1996: 285) clearly differentiates them. The former refers to "a culture of standardization" where a group of people "are united

[2] Even though in Western countries there is currently a massive operation of deconstruction of identity, above all among young people, promoting a fluid understanding of gender, along with a long series of labels for classifying all the possible genders available.

[3] A list of further references on this topic, which does not pretend to be exhaustive, is Bayart (2009), Boumard, Lapassade, and Labrot (2006), Sahlins (2000), and Sciolla (1983; 2000).

[4] This means that we can have multiple but defined identities which we constantly build and change according to the context in which we act.

[5] For some traditional definitions of speech communities, refer to Bloomfield (1933: 49), Duranti (1997); Gumperz (1968), Hockett (1958: 8), Le-Page (1968: 189–212), Lyons (1970: 326), Labov (1972: 120), Laitin (2000), Silverstein (1996, 1998).

in adherence to the idea that there exists a functionally differentiated norm for using their "language" denotationally". The term 'speech community' is employed for a group of people having multiple 'languages'. Between the nineteenth and twentieth centuries, the nationalist ideology of European thought developed the concept of space as dominant not only to define new states, but also to apply to the definition of social spaces. In so doing, this nationalist ideology helped to create physical borders between people, and contributed to separating what was previously a continuum of people and consequently of speakers. Through this procedure, language communities were delimited, and this is particularly evident also in countries colonised by Europeans, such as in Africa and Asia. In these contexts, in fact, not only were new 'artificial' communities created, but European values were imported. For example, in India, Nehru and Gandhi, the leaders of the struggle for independence, had both studied in Europe. While conducting their struggle for freedom, they imported to and adopted in their own land the European (nationalist) identity concept: "we must be Indians under one single language, we must use our Indian language" (quoted in Muru 2009).[6] But what was this Indian language? How could it be possible to define one single (language) identity within a country that Khubchandani (1991) defined as a multicultural rainbow? Thus, how could it be possible to mark borders which identify a monolithic language community in this kind of context? And above all, how was it possible to apply the concept of 'one nation, one language' to India? As Anderson (1983: 4-7) points out:

> nation-ness, as well as nationalism, are cultural artefacts of a particular kind […] the nation is imagined as limited because even the largest of it encompassing perhaps a billion human beings, has finite, if elastic, boundaries, beyond which lie other nations […] finally it [a nation] is imagined as a community, because regardless of the actual inequality and exploitation that may prevail in each, the nation is always conceived as a deep, horizontal comradeship.

The result is that the only possible way to identify a nation is to determine an A which is different from a B. This means that A starts and ends at a specific point, and the same holds for B. Consequently, the nineteenth-century idea of an organic nature of a group of people arose and was identified with the idea of community itself (Morpurgo 1996; see also Pennycook 2005: 60-75). At this point, it was necessary to find elements which could help the community to be solid and tied together, and language was one of the strongest tools available. Indeed, through language, one can unify and at the same time mark differences between people, and thus (physical) borders between groups (Gumperz 1982). However, this idea is not common everywhere in the world. For example, Le-Page and Tabouret-Keller (1985: 209-215) demonstrated the flexibility of language perception among speakers of Creole, Belizean, and Spanish, which changes according to context, time, and people. Indeed, they claimed that language was used for self-ethnic identity only sporadically. Furthermore, members of the younger generation seemed to be able to establish their identity separately from their language.

Taking into consideration the definition of a speech community in using the language of science, it is possible to observe how much it is anchored in the idea of language, that is: a group of people who speak the same language. Only within sociolinguistic and anthropological frameworks have speech community definitions been unanchored from

---

[6] For the original debate about this issue, see Bipan, Mukherjee, and Mukherjee (1999); Khubchandani (1991), and Pattanayak (1981).

the idea of language, and definitions nowadays may be more or less strictly correlated to it. For example, in Labov (1966), a speech community is a group of people with shared norms and common evaluations of linguistic variables. In Eckert (2000), rather than concentrating on patterns of interaction, the focus is on why people come together, in what practices speakers engage, and how they shape and are shaped by their linguistic usages: she refers to a 'community of practice' rather a 'speech community'. Finally, for Duranti (2001), the speech community is not an entity that is observable in a historical, social, and linguistic dimension as a unitary system, but rather, as a set of practices that are socially shared.

Indeed, as already stated above quoting Silverstein (1996), speech communities are not linguistically homogeneous (or identifiable and stable) but are organisations of diversity where people may speak several languages having a set of norms about how each is to be used. As Hill (2006: 114-117) states, "speech communities will differ not only in manifesting different kinds of language structures, but in manifesting different patterns of use". Even though there is no 'best' definition of a speech community (rather, it is the scholars' attitudes which count), it is true that in the majority of cases, language is not only a tool for the expression of culture but also of identity, which marks the contraposition between majority languages and minority languages, where the former corresponds to the bigger group of people and the latter to the smaller. Therefore, the linguistic policy that states adopt can determine the nature of the relationships among these majority and minority groups. Within this perspective, languages assume a specific nature; they become perceived as objects useful to define (monolithic) groups, each with a historical and social valence. Once again, the language concept, as well as the identity concept, and the *we*-concept discussed above, are all considered as labels identifying discrete categories, while all these concepts are flexible within each individual. Not only is it "important to realise that communities are not monolithic: there are frequently conflicting beliefs and ideologies within speech (and language) communities regarding language, its status, domains, functions, policy – and who has the authority or legitimacy to decide any of these" (Austin and Sallabank 2014: 14), but one has also to keep in mind that some attempts have been made to define 'a language' and these have shown that any useful definition is heavily dependent upon context (Bobaljik, Pensalfini, and Storto 1996: 3). Indeed, as Hill (2006: 114-117) asserts, "speech communities are not linguistically homogeneous (or identifiable and stable!) but are organizations of diversity".

2. What counts as language in language documentation?

This section discusses the language concept further, focusing on how it is used in language documentation and how it could, or maybe should, be used instead. Assuming that each speaker adapts their ways of speaking to the context of situation and to their interlocutors as well as the topic, it is possible to affirm that rather than referring to 'language', one should use the term 'variety', as has been sustained within sociolinguistics (Weinreich 2010 [1979]). Therefore, the first issue related to the language concept and its identification or definition is a series of questions:

- How do the speakers use the language varieties?
- For what purposes, with whom, and in what contexts?
- What kinds of speakers use different language forms?
- How much can speakers control the varieties they use?
- How do speakers behave towards different language varieties?

These questions are strictly related to observation of language within a sociolinguistic framework, where the focus is on diversity of speech (Hymes 1972: 38).

Therefore, the questions concern what is the speaker's own language – what language do they speak, with whom, for discussing what, in what context, and when do their ways of speaking change?[7] This demonstrates that even for a single speaker it is not possible to ascertain one language, but rather a set of varieties which may partially overlap or may totally differ. This is quite common everywhere in the world, above all in plurilingual contexts, but it is perceived differently in Western countries where the nationalist political idea is strong and the idea of 'the nation' is usually associated with an unrealistic monolingualism. For example, it is evident in Eritrea and Ethiopia, where, for example, Saho speakers – whose population includes several clans, most of them with their distinctive territories – refer to their own language with different glottonyms (*Saahot luqha* or *Saahot waani* – 'language of the Saho' – with variants *Sāho, Saḥo, Šaho, Šaḥo, Šoho*).[8] In this regard, Fill and Mühlhäusler (2006) claim that both the process of identification and the way languages are named are far from being an act of objectivity. The Saho language does not differ from other examples. Previously it was an oral but unwritten language, which was later recognised as one of the nine national languages of Eritrea. This policy change brought the need to select a national writing system and thus Saho is written today in a Latin-based orthography. However, reality counts three major dialect groups of Saho (Northern, Central, and Southern), along with three writing systems and four major groups of poetry genres which are functionally differentiated (Banti and Vergari 2010; 2014). Therefore, to what does today's 'Saho national language' correspond? Does it correspond to an abstract variety selected to represent the social group at a national level, or to the ways of speaking that speakers really use in their daily interactions?

A similar situation is found in Sardinia, where the language named Sardo was included among the languages to be protected by Italian law 482/1999 on minority languages despite the fact that no Sardinian spoke it since it does not exist: rather, two main varieties (*logudurese-nuorese* and *campidanese*) represent what Sardinian people really speak.

Other possible contexts include situations where language is not perceived as a tool for identifying someone or something. This is the case among the Bantu in Southern Africa, where the concept of ethnic or tribal identity may be strong but the concept of a tribe 'owning' a variety of Bantu as its own language has been definitively recognised as an invention of the British colonisers, "whereas the Shona themselves had been accustomed to thinking of the linguistic behaviour of all Bantu as one continuum of 'language'" (Le Page and Tabouret-Keller 1985: 240). Finally, as Childs, Good, and Mitchell (2014: 173) state, "multilingual diversity […] is not merely about how many languages an individual may have competence in. A single language may show a similarly complex diversity of functions, taking on variable significance depending on where it is used and with whom". Therefore, multilingual diversity can be addressed when referring to a single language. It is clear that a language (or a dialect) should be seen to be for the most part a cultural construct. For example, in this regard, Gal (2006) asserts that languages are (also) a European invention, while Duranti (1997: 332)

[7] This is the series of questions asked in during my lesson at the Summer School in Language Documentation held in Naples, 10-14 July 2022.

[8] Saho varieties have been largely studied and documented by Moreno Vergari (*Ethnorêma*) and Giorgio Banti (formerly University of Naples "L'Orientale"). See Banti and Vergari (2010: 83-108; 2014: 133-144).

proposes not only rethinking language as an object of study but also deconstructing the idea of language itself. He even goes beyond this call and claims that his aim "represents an interdisciplinary effort to improve on the notion of linguistic communication currently used or implied in the social sciences and the humanities", where the focus is not on the language used in communication but rather on the relationships among people,[9] and then on the basic need that people experience during a communicative act – that is, to be recognised by others.

In conclusion, consider that:

(1) speakers adapt their language to each context, topic or interlocutor involved in speech acts;

(2) language serves to negotiate spaces, identities, topics, borders, etc.;

(3) language is the link between social categorisation and stratification and it is used to create meaning through social interactions.

This means that the first important step that should be undertaken in language documentation (LD) is to document the contrasting uses of multiple varieties by a given group of people. Consequently, analysis related to languages should focus on linguistic repertoires and practices.[10] Indeed, as Lüpke (2010: 60) states, "Language Description and Documentation (henceforth LDD) research often takes place in multilingual and multilectal settings in contexts where no standard variety exists" and diglossic and triglossic situations should also be exhaustively documented because they are included among the linguistic practices of a community.[11] Given these premises, we can move to discuss how LD should incorporate a sociolinguistic approach.

3. Traditional approaches in language documentation

A typical definition of LD refers to systematically recorded representations of both spoken and written forms of a language in their appropriate sociocultural context.[12] As Himmelmann (1998) states, LD includes all those methods, theoretical frameworks and tools that are deemed necessary for the recording of a natural language or one of its varieties (see also Austin 2006; Austin and Sallabank 2011).

According to Austin (2006: 87-112), language documentation begins with the development of a project to work with a speech community on a language, and its major goal is the creation of lasting multi-purpose records of languages or linguistic practices through audio and video recordings of speakers and signers, and annotation, translation, preservation and distribution of the resulting materials.

Taking these definitions of language documentation and description into account, it becomes apparent that the focus of research is on the language Y (or on a variety of it– that can be endangered or not), of which one ideally covers a diverse range of genres and contexts, which is spoken by the community X and which must be recorded, described, preserved, and distributed.

---

[9] "For linguistic anthropologists the question of the nature of language cannot be separated from the question of the use of language by particular individuals at a particular time and with time as one of its fundamental dimensions" (Duranti 1997: 337).

[10] For further reading on this issue, refer to Calvet (2006), Gal (2006), Pennycock (2005), Riciento (2005), Sallabank (2011).

[11] At the same time, the social significance of code-switching, like all uses of varieties of language as acts of identity, should be considered. In this regard, see Gumperz (1977) and Le-Page and Tabouret-Keller (1985).

[12] For further reading about language documentation, refer to Austin (2006; 2010), Austin and Grenoble (2014), Grenoble (2010), Grenoble and Furbee (2010), Himmelmann (1998; 2006), Lehmann (2001), and Woodbury (2003; 2011).

What does this assumption imply?

Considering the previous paragraphs, it should clearly emerge that the foundation on which LDD relies implies the theoretical assumption that speakers can be grouped within a (stable) delimited speech community[13] and within this community they learn to identify themselves through the language they speak. Therefore, "language is the label used to distinguish one group from another, mainly to reinforce the notion of cultural identity and often, unfortunately, superiority or inferiority" (Bobaljik, Pensalfini, and Storto 1996: 3). Hence, there is usually a community that is at a disadvantage with respect to another; this community uses a language that may be threatened by the language of the strongest community, and this language turns into one of the only means of expressing a group identity. Consequently, the loss of the language within the minority group would be equivalent to a loss of identity for the community of speakers itself and, as a consequence, the loss of its culture, since it is their language that expresses it.

Consequently, by identifying boundaries defining language, identity, and speech community, these concepts are turned into stable concrete elements which exist within a specific reality and at a specific time, being so transformed into a unified organic element. On the contrary, language can be a component of identity construction but not necessarily the only one or the central one.[14]

Consequently, also considering the multilingual diversity characterising speakers' repertoires, the first difficulty that LDD has to face is the question of what should be documented and how.[15] Secondly, one should wonder if and how much a LD carried out with these principles can contribute to reinforcing the idea of preservation. As Austin and Sallabank (2011) claim, traditional LDD is typically based on data elicited from elderly people instead of younger ones. The younger generation is usually kept apart from the process (unless the focus is on language acquisition which looks at how children acquire their verbal repertoires), or is taken into consideration only in order to demonstrate that the language is changing or is becoming lost, and that it is progressively becoming endangered. This seems automatically to imply that linguists should document a language in order to preserve it in the way it is spoken by the elderly people[16] – since they are more representative of the 'pure' concept of language we want to document or describe. If one applies this perspective, then it seems that each language in the world should be considered as a threatened language, since all of the oldest speakers will eventually be lost and young people will be the innovators. In this way, there are speakers of minority languages who become 'fossilised entities' that must be studied and preserved in the way they are without giving them the chance to participate in the changes the world is undergoing. On the other hand, it pushes small communities of speakers living in the same area to fight in order to keep and defend their own language which is perceived as a tool for their identity affirmation.[17] In my opinion, LDD could bring unfair results if the relativism related to the concepts of language, identity, and

[13] For example, when discussing research methods in LDD, Lüpke (2010: 60) talks about the need for LDD to define the target population and its boundaries—indeed, the need for identification of the speech community. On the documentation of endangered languages and speakers, see Grinevald (2003) and Dobrin and Berson (2011: 187-211).
[14] I am grateful to Peter Austin for this observation on the issue language = identity.
[15] For further discussion on research methods in language documentation, refer to Lüpke (2010: 55-104).
[16] Peter Austin (personal communication and 2014: 13) underlines how this view is commonly part of the ideology of many communities also.
[17] For further reading about the use of languages as tools for the creation of identity, refer to Amselle (1999; 2001), Amselle and M'Bokolo (2008), Fabietti (2000), and Gallissot, Kilani, and Rivera (2007). For a discussion on the construction of identities, refer to Assam (1997), Kroskrity (2000), and Said (1978). For a reflection on cultures, see Abu-Lughod (1991), Geertz (1973), and Kubchandani (1991).

speech community is not taken seriously. For example, it could enable puristic ideologies and attitudes which are harmful to endangered languages rejecting innovations such as loan words, code mixing, and code switching.[18] Furthermore, according to the kind of data which are collected, speakers and their languages can become stereotypes which are kept alive in a way that is more folkloristic than real. For example, Himmelmann (2006: 4) talks about the risk of "data graveyards", stating that there will be

> Large heaps of data with little or no use to anyone […] language documentation is not a theory-free or anti-theoretical enterprise. Its theoretical concerns pertain to the methods used in recording, processing, and preserving linguistic primary data, as well as to the question of how it can be ensured that primary data collections are indeed of use for a broad range of theoretical and applied purposes.

Over the last few years, the contribution of sociolinguistics to LDD has pointed out how arbitrary the decision is, when taken alone by linguists, concerning the kinds of data and uses of language they are going to collect. Sociolinguistics has highlighted how LDD carried out up to this point has tended to focus its recordings mainly on those elements considered necessary for the creation of grammar, dictionaries, and texts (the so-called Boasian trilogy, which remains central to LDD grammar). Therefore, in order to avoid the risk of documenting fossils, as well as not to put pressure on a group of speakers, there is a need to listen to the speakers and to understand if and in which ways their ways of speaking are threatened. Secondly, it needs to be understood whether there is a real desire and will among the speakers to maintain a specific language. This is the only case in which linguists should be allowed to intervene and make efforts to help them revitalise their way of expression. As Florey (2004) claims, language documentation and language maintenance do not mean preserving the language untouched like an artefact in a museum.

4. A sociolinguistic approach to language documentation

Migge and Legalise (2012) point out that is not possible to think about language documentation if the language is not observed from different perspectives; this particularly holds true in multilingual societies.[19] Hence, Migge and Legalise (2012: 308) have carefully investigated the nature of the Takitaki language from various angles (i.e. language usage, speakers' perception of it, typology of speakers, etc.) "using different types of data, data collection methods and analytical methods" and have "stressed the idea that without all this, any linguistic description would be incomplete and simplistic at worst." In fact, what emerges in their study is that knowledge of the Takitaki language often "appears to be indispensable for carrying out some activities in [the speakers'] everyday life, but people's feelings toward it are ambivalent. For most of them, the aim is not to learn or speak and understand it perfectly, or to emulate Maroon practices, but to get along and to interact with people locally" (Migge and Legalise 2012: 310). Austin and Sallabank (2011; 2014) and Grinevald (2003) have also demonstrated how important it is to adjust the scheduled programme to collect the right corpus for language documentation in the light of this new awareness – indeed, what

[18] I am in debt to Peter Austin for this observation.
[19] See also Le-Page (1968).

speakers, what language, what context, when, how, etc. ways of speaking should be documented.

This points once again to the notion that defining a language is not easy, since it does not correspond to a specific entity. Migge and Legalise's (2012: 310) analysis of the social profiles of people saying that they practice Takitaki indicates that what is commonly referred to by the term Takitaki "consists of different language practices of varieties."

Therefore, before carrying out any kind of recording and preparing a corpus that might be representative of the description and analysis of a language, one should understand:

- What is it that speakers think their language is?
- Whether the language is a real ethnically identifying criterion for those speakers with whom we are going to collaborate.
- Whether it is possible to analyse languages without taking human beings' identity negotiations into proper consideration. Is it possible to exclude and not consider in which situations and kinds of interactions the speakers exchange and negotiate their 'identity'?

As Childs, Good, and Mitchell (2014: 171) claim, "sociolinguistic documentation can be understood as extending our conception of language documentation beyond its typical, nostalgic emphasis on specific ancestral codes to the sociolinguistic contexts and patterns of language use in a given community". In fact, sociolinguistic documentation should collect linguistic data in a carefully considered range of contexts reflecting the social features that characterise the different social groups. Furthermore, the collected data should be associated with a satisfying representation of the sociolinguistic contexts where various codes are used. This means that the traditional sociolinguistic interview[20] is not enough. Rather, other methodologies and methods coming from anthropological studies should be included, such as participant observation which is also "a constructive methodological resource for those who are committed to conducting their research in an ethical manner" (Schwarz and Dobrin 2018: 256). At this regard Schwarz and Dobrin (2018: 260) while underlining that "linguists are more aware than ever before that language documentation does indeed involve "linguistic social work"" emphasise "how the ethnographic method of participant observation can help documentary linguistics establish positive social relationships with language communities" (Schwarz and Dobrin 2018: 264).

This also implies that a language documentation project should be carefully planned in terms of time to be spent within the community. In conclusion, as Harrison (2005: 22) states, a sociolinguistic documentation practiced with concerns for socio-cultural variables should also aim to be an "ethnographically informed language documentation", which advocates for "the inclusion of ethnographic methods […] a restored balance between structuralist concerns and attention to [the] cultural content of speech".

Embedding the practices of LDD within a sociolinguistic approach should be understood as a valuable theoretical and methodological framework whose purpose is to promote change and/or social development in human communities. Indeed, sociolinguistics, both micro and macro, can contribute to a better understanding of how

[20] On the sociolinguistic interview and analysis of language variation, as well as on sociolinguistic and ethnographic approaches in fieldwork, see Briggs (1986; 1999), Meyerhoff *et al.* (2011), Milroy and Gordon (2003), Schilling-Estes (2007; 2013), and Tagliamonte (2006).

the LDD can lead towards a particular language policy[21] instead of another one. Sociolinguistics also reflects on the perception the speakers have of themselves or want to show to others when they are involved in communicative acts. Furthermore, it analyses the language usage differences among speakers – i.e. what variety is spoken in a certain situation and with whom. All these analyses can contribute to a better understanding of what should be documented and, above all, how this documentation should be carried out. Sociolinguistics can provide the mental flexibility that it is necessary to maintain in a LDD project when dealing with people and varieties of languages. Furthermore, by considering a sociolinguistics approach along with ethno-linguistic and the pragmatic ones, LDD can be included as a valuable means through which to promote the revitalisation of a language.[22] As Flores Farfán and Ramallo (2010: 1-12) argue, one should carry out a "[s]ociolinguistic of development, in which the revitalisation of linguistic communities is the priority"; this "opens [a] new perspective for the emerging field of linguistic documentation, in which the societal aspects of the research have frequently been marginal".

The basis should be that LDD should aim at bringing benefits to a group of speakers in the first instance, and to think about the merely (ethno-)linguistic aspects secondly. Indeed, it is necessary first to make a choice and decide what kind of approach one wants to adopt while conducting a documentation. We can pursue a passive (and utilitarian) documentation, by which we obtain a simple collection of linguistic data that may contribute to linguists' scientific knowledge of the world's languages and to linguistic theory. This approach essentially ignores what are going to be the effects on the group of speakers we plan to work with, or what is going to happen to them or the records after we leave. In this way, we can simply work by stressing the idea of the existence of a language which has the right to be preserved and possibly transmitted to future generations with the aim of keeping diversity alive; whenever this is not possible, we can try to maintain a picture of the language through its description. The risk with this approach is of marking the concept of alteration/diversity[23] negatively and excessively stressing the concept of identity.

Another approach linguists can adopt is an active and supportive/sympathetic LDD. Indeed, by discarding the idea of what it means to document a language as a linguist,[24] it focuses on how the linguists' studies can contribute to the people with whom they work, and to the broader society as well (see Austin in this volume).

I do not want to say that LDD must become a new means of revolution for the existence of human beings, but rather we can assume that linguistics as a discipline can cooperate with other disciplines in order to bring advantages and benefits to others, and specifically to those working with us as consultants. We maintain that linguistics can actually do that without renouncing its first intent toward a broader understanding of the

[21] For a discussion on the connection between language documentation and language policy, refer to Sallabank (2007: 144-171). For a discussion on language ideologies and beliefs about languages, refer to Austin and Sallabank (2014) and Gal and Irvine (2000: 35-84).
[22] On language revitalisation, see Austin and Sallabank (2014), Grenoble and Whaley (2006), and Fill and Mühlhäusler (2006).
[23] For further discussion about a theory in the field of language documentation, refer to Louanna Furbee (2010: 3-24) and Mosel (2006: 67-85).
[24] Berge (2010: 51-66) discussed the concept of adequacy in documentation: "we can derive some general guiding principles, many of which suggest that, ultimately, adequacy in documentation means letting go of preconceived notions of what it means to document a language as a linguist" (2010: 52).

theory of language; if that is the case, why should we not consider this opportunity more seriously?

In conclusion, to adopt a sociolinguistic framework in language documentation means to de-emphasise the ancestral code and the idea of an ideal speaker/signer and of a pure language that has a greater right to be documented. It means moving away from a language-as-system view to a language-as-practice view, since this contributes to emphasising situated language use. Therefore, the best practice is to apply a transdisciplinary approach whose goal is integration of approaches, including not only sociolinguistics but also ethnographic methods (i.e. life histories, naturalistic conversation, communicative modalities including gestures and other non-verbal resources). Within this framework, we emphasise situated language use and focus on the linguistic habits of speakers, including all their linguistic varieties and their multilingual competencies (e.g. code-switching and code-mixing; local slang and 'secret' languages; the language of everyday activities as opposed to traditional narratives). It also suggests that researchers remember that "language practices are indeed influenced by socioeconomic and political changes which constantly take place" (Sugita 2007: 243), inviting us to understand the dynamics that lead people together as part of a group, the activities in which they are engaged, the common experiences they share and, above all, the needs and interests of those who identify with a given language or way of speaking.

5. Concluding remarks

The previous sections have discussed concepts such as language, community of speech, identity, *we*- and others, which are crucial to language documentation. Not only has it been argued that it is difficult to apply these labels to reality, but also that misapplication could lead to unfair results in LDD. Therefore, it has been proposed not only to adopt a sociolinguistic approach in LDD but to contemplate different methods, including ethnographic ones, arguing that this would give the necessary feasibility to any LDD project. Here we discuss further what has been proposed above with the aim of offering some conclusive statements about what it would be necessary to do.

Firstly, our concept of language, whether endangered or not, should change and it should be treated as a flexible concept. Therefore, LDD should consider this possibility in greater depth, and languages which are considered endangered should no longer be transformed into roses to be treated with extreme delicacy or butterflies to be collected in a box. In this light, that which should be preserved is not a particular language, as if it were a living organism, but the diversity of ways of speaking among human beings, because, as Edwards (2009: 232) claims, "language is not organic. Languages themselves obey no natural imperatives; they have no intrinsic qualities that bear upon any sort of linguistic survival […] they possess no 'inner principle of life'" (see also Fill and Mühlhäusler 2006; Mühlhäusler 1996, 2000; Mufuwene 2002). Rather, speakers are alive and they make choices about the ways they want to speak. The problem is that this choice often – above all with regards to threatened languages – is not a true choice, since political, social, and economic factors threaten both linguistic and environmental-biological diversity, pushing people to abandon their languages and their style of life, not only because of the unfavourable social and economic conditions in which threatened language speakers usually live but also because, as Austin and Sallabank (2011: 10) claim, there is "a common assumption, especially among non-linguists […] that the usage of a single language would bring peace, either in a particular country or worldwide". For example, Brewer (2001) assumes that linguistic diversity contributes to interethnic conflict, while examples like the former Yugoslavia show how linguistic divergence (and then diversity) were more a consequence than the actual cause of conflict. Furthermore, linguistic divergence there was the ultimate tool created to delimit and maintain borders. Indeed, what was formerly known as Serbo-Croatian is now split into Croatian, Serbian, Bosnian, and Montenegrin (Austin and Sallabank 2014; Brewer 2001; Greenberg 2004). On the contrary, many other realities demonstrate that multilingualism and the desire that speakers have to maintain distinctive linguistic realities is a common pattern. For example, the acquisition of multilingual competence has been a normal part of every child's socialisation among Cape Keerweer people in northern Australia (Le-Page and Tabouret-Keller 1985: 240-242). Indeed, it is much more a modern assumption that language diversity is a problem, both in the capitalistic view as well as in the 1930s USSR perspective. Anderson (1983: 43) claims that the print languages laid the bases for national consciousness and led to a crucial impact on the diversity of languages and of human beings.[25] Indeed, before capitalism arose and before the invention of print as tool for a nation-state ideology that idealised the axiom one people = one language promoting a monoglot mass reading public, mutual incomprehensibility among people was not a problem. Language diversity was of only slight importance historically. For example, in the medieval period it was religion rather than languages which played a role of identity or belonging.

Consequently, one should not document and describe languages but rather peoples' ways of communication; one should not talk about endangered languages and the loss of identity of communities of speakers because, as Austin and Sallabank (2014: 15) have already pointed out, 'language' as a concept might mean something different to different people. As a consequence, language may not be the only thing representing the ways in which people identify or express themselves. Therefore, establishing an identity to be preserved through a language can led to the idea that the recognition of that language is the only means through which human beings can perceive themselves or be recognised by others. Furthermore, emphasising the concept of identity may lead to the idea that too many identities (and, as a consequence, too many languages) within the same shared reality can become a problem for the maintenance of peace within a society. On the contrary, documentation and description should enhance the discovery and reconsideration of diversity—that is the normal status—and should recognise the possibility of accustoming people to diversity itself as their main aim.

Consequently, in my opinion, the focus of LDD should move ahead to document the diversity of language, trying to contribute further to discarding the negative idea of the *we*- dimension and *other*-dimension. Indeed, the contraposition between different *we*- implies that if we could all communicate through one language, our collaboration and co-habitation could improve. In contrast, if we move towards other fields of research (i.e. environment, agriculture, nutrition, etc.), it is clear that the general assumption to which everybody is headed is that diversity is not only better but necessary for our survival. Where can this diversity be found? It could be embedded within varieties of

[25] In this regard, one can also stress how much writing is crucial to the empowerment of a language, much more if it is printed writing. For example, Remotti (1996: 54) underlines how the power of writing offers a way to bolster the concept of identity: "the written text is something that fixes the identity, removing it from the 'flux' and the different 'alternative possibilities' in order to fix it in a perpetual (or almost) form" [il testo scritto è qualcosa che inchioda l'identità che la stacca dal "flusso" e dal turbinio delle "possibilità alternative" per fissarla in una forma perenne (o quasi)].

languages. Consequently, it is not necessary to document and preserve the world in the way it was, since changes are normal and have always occurred, but to learn from what was good and positive in a previous stage of the world, mainly before capitalism, and to readapt that knowledge (also available through the description of small cultural realities) to the society of today.

Secondly, LDD made with the aim to preserve diversity through the documentation of a specific language perceived as a symbol of identity for a community can lead to unfair results, emphasising the minority of some groups or denying them a chance of change. Should linguists stop documenting languages? Should we simply sit and wait for the linguistic diversity in the world to disappear? As Krauss *et al.* (1992: 8, 28) stated long time ago, in my opinion we should not. On the contrary, we should carry out the kinds of research and studies that aim to maintain and enhance the value of linguistic diversity, and we should help people to live together within this diversity.

One first step towards this approach is to carry out an LDD oriented towards communities, specifically by linking the needs and interests of speakers/signers to the documentation of practices (Mosel 2006: 67-85) and communicative understandings.

LDD should describe how speakers pick up the varieties available to them or create new ones. "If a language is action, as proposed by Malinowski, and the ways we speak provide us with ways of being in the world, as suggested by Sapir, Whorf, and many others, linguistic communication is part of the reality it is supposed to represent, interpret and evoke" (Duranti 1997: 232). What should be documented is not the language, but the linguistic communicative mechanisms, that process through which human beings interact and share the world and reality, and negotiate its recognition in the society.

In order to do that, it is necessary to plan the documentation and the kinds of data to be obtained from a different perspective (e.g. by taking into account the possibility of documenting conversations instead of narrative stories told by elderly people: at least the proportion of conversations should be higher than that of narrated stories, which are still important for achieving other results), therefore adopting the methods and techniques suggested by sociolinguistic and ethnographic studies (Lüpke 2010: 67-96).

We should stop talking about the almost obsessively stated 'endangered languages', or at least, we should consider the idea of talking about 'denied recognition of human beings' along with 'endangered languages'.

Instead of academic discussions (only and merely) about endangered languages as if they were monolithic, linguists should take into consideration what is actually under threat: as soon as a different variety of languages disappears, relationships among human beings loosen and worsen. What is endangered is the ability to be flexible, which may make us able to tolerate and adapt to the real diversity instead of conforming to an artificial uniformity.

LDD should contribute to promoting the spread of diversity and should reveal to the world that there is no reason to be scared about it. On the contrary, it is only within this diversity that we can exist. In doing this, LDD could contribute to the improvement of those groups which are representative of a small piece of diversity and to whom it is usually required to choose between being what they would like to be and what it is better to be according to our model of society, or political and socioeconomic pressure, leading them to become uninterested in their own way of existence.

Finally, I would argue that LDD should be carried out with the purpose of achieving the recognition and then the acceptance of people and their ways of expression. If language is one of the means through which these human beings can find the strength to allow others to be discovered, through which they can affirm their existence as different human beings, then the language is welcome. However, people should have the freedom to express themselves in the ways they choose: therefore, one should not protect languages but rather linguistic rights. If we want to consider the concept of language as a tool for communication through which we build our relationships, we can say that this objective may be lost, change, disappear, or reappear in a different form, or change just a little. Knowledge of a language that undergoes one of these transformations should not represent a problem or lead the speakers to choose between two entities, the *we*-concept imposed from the outsider on the group or the in-perceived *we*-concept as it is understood within the group, when it can simply be an inclusive and exclusive *we* at the same time.

LDD should become a campaign for sustaining diversity as a natural and, above all, necessary pattern of, to, and for the world.

References

ABU-LUGHOD, Lila (1991) "Writing against culture". In FOX, Richard G. (ed.), *Recapturing anthropology. Working in the present.* Pp. 137-162. Santa Fe, New Mexico: School of American Research Press.

AMSELLE, Jean-Loup (1999) *Logiche meticce. Antropologia dell'identità in Africa e altrove.* Torino: Bollati Boringhieri.

AMSELLE, Jean-Loup (2001) *Branchements. Anthropologie de l'universalité des cultures.* Paris: Flammarion.

AMSELLE, Jean-Loup, and Elikia M'BOKOLO (2008) *L'invenzione dell'etnia.* Roma: Meltemi.

ANDERSON, Benedict (1983) *Imagined communities: Reflections on the origin and spread of nationalism.* London/New York: Verso.

ASSAM, Jan (1997) *La memoria culturale. Scrittura, ricordo e identità politica nelle grandi civiltà antiche.* Torino: Einaudi.

AUSTIN, Peter K. (2006) "Essentials of language documentation." In GIPPERT, Jost, Nikolaus P. HIMMELMANN, and Ulrike MOSEL (eds), *Essentials of language documentation: 178 (Trends in linguistics. Studies and monographs).* Pp. 87-112. Berlin/New York: Mouton de Gruyter.

AUSTIN, Peter K. (ed.) (2010) "Current issues in language documentation". *Language Documentation and Description, Volume 7.* Pp. 12-33. London: SOAS.

AUSTIN, Peter K. (2014) "Introduction". In AUSTIN, Peter K., and Julia SALLABANK, *Endangered languages: Beliefs and ideologies in language documentation and revitalization.* Pp. 1-26. Cambridge, UK: Cambridge University Press.

AUSTIN, Peter K., and Lenore A. GRENOBLE (eds.) (2014) "Current Trends in Language Documentation". *Language Documentation and Description, Volume* 4. Pp. 12-25. doi: https://doi.org/10.25894/ldd258

AUSTIN, Peter K., and Julia SALLABANK (2011) "Introduction". In AUSTIN, Peter K., and Julia SALLABANK (eds.), *The Cambridge handbook of endangered languages.* Pp. 1-24. Cambridge, UK: Cambridge University Press.

AUSTIN, Peter K., and Julia SALLABANK (eds.) (2014) *Endangered languages: Beliefs and ideologies in language documentation and revitalization.* Oxford: Oxford University Press.

BANTI, Giorgio, and Moreno VERGARI (2010) "The Saho of Eritrea and the documentation of their language and cultural heritage". In LUSINI, Gianfrancesco. (ed.), *Annali. Current trends in Eritrean studies*, Vol. 70. Pp. 83-108. Napoli: L'Orientale.

BANTI, Giorgio, and Moreno VERGARI (2014) The Nazme: A genre of religious poetry of the Eritrean Saho. *Incontri Linguistici* 37: 133-144.

BAYART, Jean-François (2009) *L'illusione identitaria.* Troina: Città aperta.

BERGE, Anna (2010) "Adequacy in language documentation". In GRENOBLE, Lenore A., and Louanna N. FURBEE (eds.), *Language documentation: Practices and values.* Pp. 51-66. Amsterdam/Philadelphia: John Benjamins Publishing.

BIPAN, Chandra, Mridula MUKHERJEE, and Aditya MUKHERJEE (1999) *India after independence 1947-2000.* New Delhi: Penguin Books.

BLOOMFIELD, Leonard (1933) *Language.* New York: Holt, Rinerhart & Winston.

BOBALJIK, Jonathan David, Rob PENSALFINI, and Luciana STORTO (eds.) (1996) "Papers on language endangerment and maintenance of linguistic diversity". *The MIT working papers in linguistics, Vol. 28.* Cambridge, MA: MiTWPL.

BOUMARD, Patrick, Grogers LAPASSADE, and Michel LABROT (2006) *Il mito dell'identità. Antologia della dissociazione.* Dogliani: Sensibili alle foglie.

BREWER, Marilynn B. (2001) "Intergroup identification and intergroup conflict: When does ingroup love become outgroup hate?" In ASHMORE, Richard D., Lee JUSSIM, and David WIDILDER (eds.), *Social identity, intergroup conflict and conflict reduction.* Pp. 17-41. Oxford: Oxford University Press.

BRIGGS, Charles (1986) *Learning how to ask. A sociolinguistic appraisal of the role of the interview in social science research.* Cambridge, UK: Cambridge University Press.

BRIGGS, Charles (1999) "Interview". *Journal of Linguistic Anthropology* 9(1/2): 137-140. https://www.jstor.org/stable/43102448B

BROWN, Penelope and Stephen C. LEVINSON (1987) *Politeness: some universals in language use. Studies in interactional sociolinguistics 4.* Cambridge: Cambridge University Press.

CALVET, Louis-Jean (2006) *Towards an ecology of world languages.* Cambridge, UK: Polity Press.

CHILDS, Tucker, Jeff GOOD, and Alice MITCHELL (2014) "Beyond the ancestral code: Towards a model for sociolinguistic language documentation". *Language Documentation and Conservation* 8: 168-191.

DOBRIN, Lise M., and Josh BERSON (2011) "Speakers and language documentation." In AUSTIN, Peter K., and Julia SALLABANK (eds.), *The Cambridge Handbook of Endangered Languages*. Cambridge Handbooks in Language and Linguistics. Pp. 187-211. Cambridge, UK: Cambridge University Press.

DURANTI, Alessandro (1997) *Anthropological linguistics.* Cambridge, UK: Cambridge University Press.

DURANTI, A. (2001) *Linguistic anthropology*. Oxford: Blackwell.

ECKERT, Penelope (2000) *Linguistic variation as social practice*. Oxford: Blackwell Publishers.

EDWARDS, John (2009) *Language and identity.* Cambridge, UK: Cambridge University Press.

FABIETTI, Ugo (2000) *L'identità etnica. Storia e critica di un concetto equivoco.* Roma: Carocci. [1995, prima edizione].

FILL, Alwin, and Peter MÜHLHÄUSLER (2006) *The ecolinguistic reader: Language, ecology and environment.* London: Continuum.

FLORES FARFÁN, José Antonio, and Fernando RAMALLO (eds.) (2010) *New perspectives on endangered languages: Bridging gaps between sociolinguistics, documentation and language revitalization.* Amsterdam/Philadelphia: John Benjamins Publisher Co.

FLOREY, Margaret (2004) "Countering purism: Confronting the emergence of new varieties in a training program for community language workers". *Language Documentation and Description, Volume 2.* Pp. 9-27. London: SOAS.

FURBEE, Louanna N. (2010) "Language documentation. Theory and practices". In GRENOBLE, Lenore A., and Louanna N. FURBEE (eds.), *Language documentation: Practices and values.* Pp. 3-24. Amsterdam/Philadelphia: John Benjamins Publishing.

GAL, Susan (2006) "Contradiction of standard language in Europe: Implications for the study of publics and practices". *Social Anthropology* 14(2): 163-181.

GAL, Susan, and Judith T. IRVINE (2000) "Language ideology and linguistic differentiation". In KROSKRITY, Paul V. (ed.) *Regimes of language: Ideologies, polities and identities.* Pp. 35-84. Santa Fe: School of American Research Press.

GALLISSOT, René, Mondher KILANI, and Annamaria RIVERA (2007) *L'imbroglio etnico in quattordici parole-chiave.* Bari: Dedalo.

GEERTZ, Clifford (1973) *The interpretation of cultures.* New York: Basic Books.

GOFFMAN, Erving (1955) "On face-work: an analysis of ritual elements in social interaction". *Psychiatry: Journal for the Study of Interpersonal Processes.* 18(3): 213-231.

GOFFMAN, Erving (1974) *Les Rites d'interaction.* Paris: Editions de Minuit.

GREENBERG, Robert D. (2004) *Language and identity in the Balkans: Serbo-Croatian and its disintegration.* Oxford: Oxford University Press.

GRENOBLE, Lenore A. (2010) "Language documentation and field linguistics: The state of the field". In GRENOBLE, Lenore A., and Louanna N. FURBEE (eds.), *Language documentation: Practices and values.* Pp. 289-309. Amsterdam/Philadelphia: John Benjamins Publishing.

GRENOBLE, Lenore A., and Louanna N. FURBEE (eds.) (2010) *Language documentation: Practices and values.* Amsterdam/Philadelphia: John Benjamins Publishing.

GRENOBLE, Lenore, A., and Lindsay J. WHALEY (2006) *Saving languages: An introduction to language revitalization.* Cambridge, UK: Cambridge University Press.

GRINEVALD, Colette (2003) "Speakers and documentation of endangered languages". *Language Documentation and Description, Volume 1.* Pp. 52-72. London: SOAS.

GUMPERZ, John (1968) "The speech community". In SILLS, David L. (ed.), *International encyclopedia of the social sciences*. Pp. 381-386. London: Macmillan.

GUMPERZ, John (1977) "The Sociolinguistic Significance of Conversational Code-Switching". *RELC Journal,* 8: 1-34.

GUMPERZ, John J. (1982) *Language and social identity.* Cambridge, UK: Cambridge University Press.

HARRISON, David K. (2005) "Ethnographically informed language documentation". *Language Documentation and Description* 3: 22-41. http://works.swarthmore.edu/fac-linguistics/74

HILL, Jane H. (2006) "The ethnography of language and language documentation". In GIPPERT, Jost, Nikolaus P. HIMMELMANN, and Ulrike MOSEL (eds.), *Essentials of language documentation: 178 (Trends in linguistics. Studies and monographs).* Pp. 113-129. Berlin: Mouton de Gruyter.

HIMMELMANN, Nikolaus P. (1998) "Documentary and descriptive linguistics". *Linguistics* 36: 161-195.

HIMMELMANN, Nikolaus P. (2006) "Language documentation: What is it and what is it good for?" In GIPPERT, Jost, Nikolaus P. HIMMELMANN, and Ulrike MOSEL (eds.), *Essentials of language documentation: 178 (Trends in linguistics. Studies and monographs).* Pp. 1-30. Berlin: Mouton de Gruyter.

HOCKETT, Charles (1958) *A course in modern linguistics.* New York: Macmillan.

HYMES, Dell (1972) “Models of the interaction of language and social life (revised from 1967 paper)”. In GUMPERZ, John J., and Dell HYMES (eds.), *Directions in sociolinguistics: The ethnography of communication*. Pp. 35-71. Oxford: Blackwell.

KHUBCHANDANI, L. M. (1991) *Language, culture and national building. Challenges of modernization*. New Delhi: Indian Institute of Advanced Study Shinla in association with Mahohar Publications.

KRAUSS, Michael, Ken HALE, Lucille J. WATAHOMIGIE, Colette CRAIG, LaVerne Masayesva JEANNE, and Nora C. ENGLAND (1992), “Endangered Languages”. *Language*. 68(1): 1-42.

KROSKRITY, Paul V. (2000) *Regimes of language: Ideologies, polities and identities.* Santa Fe: School of American Research Press.

LABOV, William (1966) *The social stratification of English in New York City*. Washington DC: Center for Applied Linguistics.

LABOV, William (1972) *Sociolinguistic patterns.* Oxford: Blackwell.

LEHMANN, Christian (2001) “Language documentation: A program”. In BISANG, Walter (ed.), *Aspects of typology and universals*. Pp. 83-97. Berlin: Akademie Verlag.

LE-PAGE, Robert (1968) “Problems of description in multilingual communities”. *Transactions of the Philology Society* 67(1): 189-212.

LE-PAGE, Robert Brock, and Andrée TABOURET-KELLER (1985) *Acts of identity: Creole-based approaches to language and ethnicity.* Cambridge, UK: Cambridge University Press.

LÜPKE, Friederike (2010) “Research methods in language documentation”. *Language Documentation and Description, Volume 7.* Pp. 55-104. London: SOAS.

LYONS, John (1970) *New horizons in linguistics.* Harmondsworth: Penguin.

MEYERHOFF, Miriam, Chie ADACHI, Golnaz NANBAKHSH, and Anna STRYCHARZ (2011) “Sociolinguistic fieldwork”. In THIEBERGER, Nicholas (ed.), *The Oxford handbook of linguistic fieldwork*. Pp. 121-146. Oxford: Oxford University Press.

MIGGE, Bettina, and Isabelle LEGALISE (2012) *Exploring language in a multilingual context: Variation, interaction and ideology in language documentation.* Cambridge/New York: Cambridge University Press.

MILROY, Lesley, and Matthew GORDON (2003) *Sociolinguistics. Methods and interpretation*. London: Blackwell.

MORPURGO, Davis (1996) *La Linguistica dell’Ottocento.* Bologna: Il Mulino.

MOSEL, Ulrike (2006) “Fieldwork and community language”. In GIPPERT, Jost, Nickolaus P. HIMMELMANN, and Ulrike MOSEL (eds.), *Essentials of language documentation: 178 (Trends in linguistics. Studies and monographs).* Pp. 67-85. Berlin/New York: Mouton de Gruyter.

MUFUWENE, Solikoko S. (2002) “Competition and selection in language evolution”. *Selection* 3(1): 45-56. http://www.akkrt.hu/journals/select

MÜHLHÄUSLER, Peter (1996) *Linguistic ecology: Language change and linguistic imperialism in the Pacific region.* London/New York: Routledge.

MÜHLHÄUSLER, Peter (2000) “Language planning and language ecology”. *Current Issues in Language Planning* 1(3): 306-367.

MURU, Cristina (2009) “Una varietà di anglo-indiano: Il Butler English”. In TURCHETTA, Barbara (ed.). *Pidgins e Creoli. Introduzione alle lingue di contatto.* Pp. 105-140. Roma: Carocci.

PASCAL, Blaise (1962) *Pensieri.* Torino: Einaudi.

PATTANAYAK, Debi Prasanna (1981) *Multilingualism and mother tongue education.* New Delhi: Oxford University Press.

PENNYCOOK, Alastair (2005) “Postmodernism in language policy”. In RICIENTO, Thomas (ed.), *An introduction to language policy: Theory and method.* Pp. 60-76. Oxford: Blackwell.

REMOTTI, Francesco (1996) *Contro l'identità.* Roma-Bari: Laterza.

REMOTTI, Francesco (2010) *L'ossessione identitaria.* Roma-Bari: Laterza.

RICIENTO, Thomas (2005) *An introduction to language policy: Theory and method.* Oxford: Blackwell.

SAHLINS, Marshall (2000) “Goodbye to ‘tristes tropes’: Ethnography in the context of modern world history”. *The Journal of Modern History* LXV(1): 1-25.

SAID, Edward (1978) *Orientalism.* London: Penguin.

SALLABANK, Julia (2007) “Language documentation and language policy*”. Language Documentation and Description, Volume 7.* Pp. 144-171. doi: https://doi.org/10.25894/ldd229.

SALLABANK, Julia (2011) “Language policy for endangered languages”. In AUSTIN, Peter K., and Julia SALLABANK (eds.), *The Cambridge handbook of endangered languages.* Pp. 277-290. Cambridge, UK: Cambridge University Press.

SCHILLING-ESTES, Natalie (2007) “Sociolinguistic fieldwork”. In BAYLEY, Robert, and Ceil LUCAS (eds.), *Sociolinguistic variation: Theory, methods and application.* Pp. 165-189. Cambridge, UK: Cambridge University Press.

SCHILLING-ESTES, Natalie (2013) *Sociolinguistic fieldwork.* Cambridge, UK: Cambridge University Press

SCIOLLA, Loredana (ed.) (1983) *Identità. Percorsi di analisi in sociologia.* Torino: Rosenberg & Sellier.

SCIOLLA, Loredana (2000) “Riconoscimento e teoria dell’identità”. In DELLA PORTA, Donatella, Monica GRECO, and Arpad SZAKOLCZAI (eds.), *Identità, riconoscimento, scambio. Saggio in onore di Alessandro Pizzorno.* Pp. 5-29. Roma/Bari: Laterza.

SUGITA, Yuko (2007) “Language revitalization or language fossilization? Some suggestions for language documentation from the viewpoint of interactional linguistics”. In AUSTIN, Peter K., Oliver BOND, and David NATHAN (eds.), *Proceedings of conference in language documentation & linguistic theory*, 7-8 December 2007. London: School of Oriental and African Studies, University of London.

TAGLIAMONTE, Sali A. (2006) *Analysing sociolinguistic variation*. Cambridge, UK: Cambridge University Press.

WEINREICH, Uriel (2010 [1979]) *Languages in contact: Findings and problems*. New York: Mouton.

WOODBURY, Anthony (2003) “Defining documentary linguistics”. *Language Documentation and Description, Volume 1.* Pp. 35-51. London: SOAS.

WOODBURY, A. (2011) “Language documentation”. In AUSTIN, Peter K., and Julia SALLABANK (eds.), *The Cambridge handbook of endangered languages*. Pp. 159-186. Cambridge, UK: Cambridge University Press. doi: 10.1017/CBO9780511975981.009

An account of recent research on Nara: tone, plural and texts

Graziano Savà
Independent researcher
(grsava@gmail.com)

ABSTRACT

The present short paper gives an account of some findings on the Eritrean and Sudanese Nilo-saharan language Nara. The context is a recent four week long research conducted by the writer, partially in collaboration with Prof. Giancarlo Schirru, with two Nara speakers in Addis Ababa. The paper focusses on tone marking, plural formation and documentation. The data are preceded by some basic information on Nara and an overview of past research on the language.

Keywords: *Nara, linguistic fieldwork, tone, plural*

ISO 639-3 code: nrb

DOI: 10.23814/ethn.20.24.sav

1. Basic info on Nara

Nara is a Northern East Sudanic language, like other Nilo-saharan language subgroups such as Nubian, Tama and Nyimang. Its ISO code is nrb. It is spoken by about 100,000 (99,800 according to *Ethnologue* (Eberhard *et al.*, 2020) mainly in Eritrea, in the area of Barentu, and in Sudan, in Kassala. A diaspora exists in town. There are four Nara dialects: Higir, Koyta, Mogoreeb (or Mogoraayeeb) and Saantoorta (“three tents”).

According to Thompson (1976b: 599):

> There is considerable dialect difference among these sections of the tribe and even striking differences within each section from village to village. They can all communicate in direct conversation, but have considerable difficulty in overhearing people of other sections talking among themselves. The differences involve both vocabulary and grammar.

In Sudan, where Mogoreeb is spoken, the Nara are known by the older name *Bāryā* or *al-Bāryā* (الباريا). Barya (ባርያ, Baria, Barea) occurs as an ethnic name in inscriptions of king ʿEzana, who ruled in the 4th century in Aksum. In Eritrea this name has been replaced by Nara in order to avoid the derogatory meaning “slave”, “one who is in service of a demon”, which *barya* has acquired in Ethiosemitic languages such as Geʿez, Amharic, Tigre and Tigrinya. This development of the old name is probably due to the fact that the Nara-speaking communities were frequently raided for capturing slaves. Speakers of Nara link their self designation with *nara* ‘sky, cloud’.

The language is commonly used in the Nara communities at home and for everyday communication, and it is transmitted to children. However, also Arabic and Tigre are widely used for intercommunication, and most literate Naras use Arabic as a written language. The diaspora in Arabic speaking countries like Sudan tends to use Arabic more and more, and also in the Nara-speaking areas in Eritrea Arabic is felt to be spreading.

Since 1988, a written form of Nara has been developed in Eritrea for schoolbooks and other publications addressed to the Nara-speaking community. It has a Roman orthography, and is based on Higir Nara, with a few items from other dialects.

2. Previous research on Nara

The first data on Nara were published by Reinisch (1874), who edited field notes by Werner Munzinger. The dialect of Nara described in Reinisch (1874) has still to be identified in full detail, but it appears to differ both from Higir and from Mogoreeb as they are spoken today. Thompson (1976) is a brief, 10-page, description of the Higir of Mogolo, one of the main Nara towns. The language in his article is called Nera, with *e* translitterating the first order of the Fidel script. The first topic it touches upon is phonology, which besides description of consonantal and vocalic inventories and syllables has two notes, one on stress and one on tone. About tone it is simply stated that there is a "normal" and a "high" tone (Thompson 1976: 484). What is called stress is rather tone in later descriptions by other authors. In terms of nominal morphology, there are a few lines of description for nouns, adjectives, pronouns, demonstratives and postpositions. As for verbal morphology, the description starts from the imperative and the jussive to aspect, tense and mood. Irregular and compound verbs close this section on verbal inflection. Verbs, however, are also treated in the section on verbal derivation. This includes also the formation of verbal nouns and adjectives. The article ends with a short section on word order and a few lines on the lexicon.

Hayward's and Dawd Abushush's studies all focus on Higir Nara. Abushush 1999 is probably the best account of Higir presently available. In 22 pages the author describes briefly the phonology of the language, stating that what Thompson described as stress it is actually tone. The overview continues with nouns, and their plural classes determined by suffixes and verbs with a number of paradigms, which are preceded by an explanation of phonological processes. The paper ends with a list of adjectives and examples of derivation. Abushush (no date) is a short, six-page paper which contains examples of the role of tone in some grammatical context. Hayward (2000, 23 pages) and Abushush and Hayward (2002, six pages) have roughly the same content. After discussing the inventory of Nara consonants and vowels, the papers go into details into the tone system of Higir. It is restated that there are two tones H and L but that some of the words have falling tones on short vowels, such as *bûk* "dove". It is also determined that no HLH sequence is allowed in the language and that a sandhi phenomena adjusts the situation when needed. Abushush and Hayward (2002), moreover, focus on the difference of vowel realisations in closed and open syllables.

Elsadig (2016) is the best account available of Mogoreeb Nara and Nara in general. It is a 76-page unpublished MA thesis which touches upon the phonology, the nominal morphology, the verbal morphology and the adjectives of Mogoreeb. It is the first work that shows the distribution of consonants and vowels in Nara. The section on tone, however, is quite short (only four pages) and shows examples of lexical and grammatical tone. The section on plural marking on nouns is quite rich, but no tones are marked.

Banti and Savà (2021) reports the achievement of other authors on Nara phonology based on Higir data collected by Giorgio Banti. The paper goes in deeper discussion of consonantal distribution and syllable structure.

In terms of lexicon, besides Reinisch's *Wörtenbuch* in his 1874 publication, Rilly (2010) has a comparative lexicon of the North East Sudanic languages to which Nara belongs. There is an indication of the dialect each Nara word comes from. The aim of

Rilly's book is to demonstrate that Meroitic is part of this group and, therefore, Nara and Meroitic would be historically connected to each other.

Savà (2018) mentions lexical research on Higir with data collected by Giorgio Banti which generated an unpublished wordlist of about 900 items.

The main sources on the history and social organisation of the Nara are Munzinger (1864: 373-536), Pollera (1913), and Conti Rossini (1916: 741-802). For more recent sociolinguistic, historical, and anthropological information, see Thompson (1975b: 598 f.), Treiber (2007), and Issa (2018).

2.1. Documentation

The only published texts for scientific aims are in Reinisch (1874). More volumes have been and are being published for pedagogical reasons and on other various topic by the Eritrean Ministry of Education.

In 2023 the present writer collected two texts from the Koyta speaker Issa Adem. The texts, of twenty minutes each, were translated and in a second phase, see next section, were transcribed by another speaker (see below). The next section also mentions the collection of two more texts in 2024.

3. My recent research

In Addis Ababa live two Naras, whose identity and names I do not reveal for privacy reasons. I will refer them as O. cand A. whenever needed. It is with them that I conducted my fieldwork for four weeks in December 2023-January 2024. I normally have been preferring to do data collection in the place where the language is spoken. However, in this case, for logistic, organisational and safety reasons this was not possible. Doing fieldwork in Nara locations in Eritrea is complicated if one has only four weeks to complete it, and in Sudan military actions are going on. Thanks to contacts established for me in Addis Ababa, by Moreno Vergari, Ethnorêma's director, to whom I wish to express my gratitude, I found the proper conditions to do my job on Nara in the Ethiopian capital, which I know very well. In particular, Vergari put me in contact with the two Nara speakers I worked with.

The working language was English and Amharic, the main language of Ethiopia (Savà 2023). O. was better in these languages than A., but A. had a better sense of linguistics analysis. It is important to notice that A. is a Koyta mother tongue speaker, though his mother is Kunama, a nearby ethnic group speaking another Nilo-saharan language. O.'s mother tongue is Mogoreeb. Most of my data, therefore, are Koyta, which is good as this Nara dialect, together with Saantoorta, is the least studied. The two consultants speak Tigrinya to each other because they feel too many divergences between the two Nara dialects they speak.

Logistically, I lived in a guesthouse, furnished with a common living room in which the consultants and me could work. O. and A. were employed as translators in Nara and this fact took time away from our interviews. In order to negotiate their availability, I talked to their supervisor Mr. Mohammed, who is not a Nara mother tongue speaker. I also had to skip working in festivities such as Ethiopian Christmas and Epiphany.

Mr. Mohammed profited a lot by attending our fieldwork sessions, as he came to understand more about the tone system of Nara and the necessity to mark tone in the translation.

The research was part of a three-year postdoctoral project for the creation of a textual database in Nara at the University of Naples l'Orientale. The University provided me not only funds for the research but also the necessary recording tools and a Macbook[1]. In detail, the recording equipment I had at my disposal for the research is:

1 MacBook Pro 13"
1 Marantz digital recorder PD661MKII
1 Zoom digital recorder H4n Pro
1 Zoom digital recorder H1n
1 Rode binaural microphone NT4
1 Rode omni microphone Reporter
1 Shure headset microphone
2 Quiklok microphone small tripod
1 Sennheiser headphone HD 206
1 Sennheiser headphone HD 65 TV

Out of these I used mainly the Marantz recorder and the Rode NT5 microphone. When I interviewed both speakers, I used the Rode omni Reporter microphone.

3.1. Data collection

To start with, in order to get a grasp on the phonemes of Nara and their realisation(s), I had a look with my consultants at didactic texts produced by the Eritrean Ministry of Culture to promote the language. They are written in Nara standard orthography, as already stated. I acknowledge Giorgio Banti for providing me with some samples. We went through words, paying close attention to tones and plural forms, but without recording to make the consultant comfortable with the new kind of work that they were going to do. I took particular note of tonal minimal pairs such as (the transcription here and in the following is in Nara orthography):

làá	spear	*làà*	necklace
túsá	urine	*tùsà*	ears
kábá	kind of disease	*kàbà*	tree species
súm	poison	*sùm*	grass
káásà	shadow	*kààsá*	forest
nóótá	meat	*nóótà*	eyes
dààtá	six	*dààtà*	place for gathering crops
féétà	moon	*fèètá*	mice
káttà	ax	*káttá*	sisters
bállá	legs	*bállà*	leg
báná	language	*bànà*	speech
ásmá	cholostrum	*àsmà*	heart
tíí	child	*tìì*	bee

Eventually I asked A. to record all the words in singular and plural.

I went back to nouns later, because it was time to record texts, a real documentation activity. I asked A. to produce oral narrations of any kind he felt comfortable with. He

[1] For this, I wish to express my deepest gratitude to Alessandra Gallo, head secretary of the Department of Asian, African and Mediterranean Studies, and Prof. Giancarlo Schirru, responsible for my research.

asked to think about it overnight. The day after he provided me with a very first complete text of about ten minutes on family love, followed by a five-minute long joke. We have immediately started transcription and translation of the first two minutes. A. revealed to be really great also at this job.

For text transcription and translation, I used PRAAT, a speech analysis tool developed by computational linguists and phoneticians at the University of Amsterdam. PRAAT allows time-aligned transcripts and translation contained in tiers with the support of various kinds of sound analysis such as wave forms, intensity, pitch and formants. This will greatly assist my phonetic transcription, but for that moment I used it for the orthographic transcription and the translation. At a certain point I asked A. to use PRAAT to do the job on his own. He did it wonderfully. The same occured when he elaborated a text I had recorded in Naples with Issa Adem. It is a personal story about, something funny but also slightly dramatic that happened to him in his youth. Issa's story was translated, therefore A. did only the transcription. It is 15 minutes long.

When we finished with the texts, we continued with identifying tones of nouns. I still had some uncertainty with some words, so I decided to write the words, classifying them according to their syllable structure. Here are some examples (notice that the character <v> indicates the phoneme /ŋ/ in Nara orthography):

CVV		CVVC		CVCV		(C)VVCV	
dùù	sheep	*sààỳ*	summer	*hárà*	warthog	*èèshí*	hut
tíí	child	*fáás*	ax	*wòsò*	snake	*tààfó*	tukan
làá	spear	*gàán*	bowl			*wáárí*	egg
						sóórà	tree sp.

CVCCV	
sóllò	scorpion
lávvá	louse
ndàrtí	mat

In the last week of my stay I was joined by the research convenor Prof. Giancarlo Schirru. We worked together on plural formation, while he worked alone on personal pronouns, with particular attention to the first person plural inclusive and exclusive.

As for the plurals, we checked the lists found in Elsadig (2016), marking tones and plural formation differences. Our main collaborator was O., the Mogoreeb speaker, as, as already stated, this is the dialect described by Elsadig.

Nara has plural forms that can be organised into classes distinguished by suffixes. Elsadig (2016) identifies the following suffixes: *-a, -ta, -ka, (C)-Ca, -la, (t)-ta, (k)-ka, (n)-na.* There are also the following replacive phenomena: *i-/-a, -e/-ta, -e/-ka, -o/-ta.* Elsadig does not mark tones.

Going through the pluralised nouns, we found out several differences in plural formation as compared to Elsadig (2016). For example:

	Elsadig	O.	
dòvgòr	*dòvgòrà*	*dòvgòrná*	sister
kodóór	*kodóóra*	*kòdóórná*	hyena
tús	*túsá*	*tússà*	ear
túl	*túlá*	*túllà*	cheek
fàrà	*fàràtà*	*fààràttà*	horse

dùù	*dùùtà*	*dùùttà*	sheep
nòssál	*nòssálà*	*nòssálkà*	face
dàbál	*dàbállà*	*dàbálkà*	leather sack
bòò	*bòòkkà*	*bòòttá*	kind of tree
jàmmù	*jàmmúkkà*	*jàmmúttà*	cat
sèrèmé	*sèrèmká*	*sèrrèmmèttá/sèrèmtá*	gazelle
hàìmé	*hàìmká*	*hàìmettá*	deer
tóórò	*tóórtà*	*tóórà*	veranda (Arabic)

3.2. Post-fieldwork activity

In a post-fielddwork period I created a database of the material collected together with relevant metadata. The metadata entries in the table are structured in the following way:

ID	sound file	Nara orthography	phonetic form	English gloss
2024_01_22_NRB_lex_27	1099.wav	kavvi	[kaŋːi]	termite

Each identifier shows the date of the recording first, then the Nara iso-code, the kind of data ("lexeme" in the example) and a serial number for the *lex* category. The second column hosts the raw audio file names assigned by the recorder. Then, there is the utterance in Nara orthography followed by the IPA transcription. Finally, the gloss in English.

As for the texts, I have exported them from PRAAT into ELAN. This is a tool developed at the Max Planck Institute initially for the study of sign languages, but that is nowadays widely used for glossing. The text in this case, as in PRAAT, is time-aligned, with tiers containing various grammatical annotation, such as morphological glosses and an indication of part of speech. In particular, I will make use of a special version of ELAN, ELAN-Corpa, created and maintained at the French CNRS research laboratory LLACAN by Christian Chanard and his collaborators. This version of ELAN allows semi-automatic glossing. In short, one associates a lexeme or a morpheme to a gloss on first occurrence and this automatically appears when the lexeme or the morpheme occurs again. This is valid for all the tiers and has the result of creating a grammatical glossary. It is a fantastic tool that allows saving a lot of time in a quite tedious operation. The glossary can be exported to other tools for the creation of a proper dictionary. These are FLEx (Fieldwork Language Explorer) and WeSay. Both software tools were created by SIL and run only on Windows (PRAAT and ELAN have versions for MacOS, Windows and Linux). In order to create a comfortable interface, I plan to install an emulator in my Mac so that I have one machine with two virtual machines, one with MacOS and one with Windows.

Another tool for sound editing is Audacity, an intuitive software widely used by linguists.

All mentioned software is freely downloadable and we have to be grateful to creators and developers, as well as those who provide funds, to make them available to everyone. I, myself, have contributed to spreading their use when I was teaching how to use them in courses at Addis Ababa University in 2008-2010, with great profit for my students.

4. Conclusions

This brief description of my research on Nara is not to be taken as an exhaustive analytic account. In fact, there is a lot to do in the analysis of the data and of the texts, in particular. Therefore, the paper is rather an example of how field research can be properly organised to improve our knowledge on a little-known language such as Nara. The friendly environment of the research was crucial for its success and for this I am deeply grateful to O. and A. The research could not have been organised also without the collaboration of Moreno Vergari and my host Yemamu Ahmed, to whom I am indebted.

References

BANTI, Giorgio, and Graziano SAVÀ (2021) "Nara phonology", *Rassegna di Studi Etiopici*, 5 (third series): 235-268.

CONTI ROSSINI, Carlo (1916) *Principi di diritto consuetudinario dell'Eritrea*. Rome: Ministero delle Colonie.

DAWD, Abushush (no date) *Effectiveness of tone in Nara*. Asmara: unpublished ms.

DAWD, Abushush (1999) *A brief essay on the phonology and morphology of the Nara language*. London, SOAS: unpublished ms.

DAWD Abushush, and Richard HAYWARD (2002) *Nara*, *Journal of the International Phonetic Association* 32:2: 249-255.

EBARHARD, David M., Gary F. SIMONS, and Charles D. FENNIG (eds.) (2020) *Ethnologue: languages of the world, 23rd edition*. Dallas: SIL International. [Online version: http://www.ethnologue.com].

ELSADIG, Omda Ibrahim Elnur (2016) *Major word categories in Nara*. Cologne: University of Cologne. Unpublished MA dissertation.

HAYWARD, Richard J. (2000) "Observations on tone in the Higir dialect of Nara". In VOSSEN, Reiner, A. MIETZNER, and A. MEISSNER (eds.), *Mehr als nur Worte... Afrikanistische Beiträge zum 65. Geburtstag von Franz Rottland*. Pp. 247-267. Cologne: Rüdiger Köppe Verlag.

ISSA, Adem (2018) *The Nara culture: continuity & change – A case study on rituals & means of livelihood*. Riga: Lambert Academic Publishing.

KIEVIT, Dirk (2007) "Nara". In UHLIG, Siegbert (ed.), *Encyclopaedia aethiopica, Vol. 3 (He-N)*. Pp. 247-267. Wiesbaden: Harrassowitz Verlag.

MUNZINGER, Werner (1864) *Ostafrikanische Studien*. Schaffhausen: Fr. Hurtersche Buchhandlung.

POLLERA, Alberto (1913) *I Baria e i Cunama*. Rome: Reale Società Geografica.

REINISCH, Leo (ed.) (1874) *Die Barea-Sprache – Grammatik, Text und Wörterbuch nach den handschriftlichen Materialien von Werner Munzinger Pascha*. Wien: Wilhelm Braumüller.

RILLY, Claude (2005) "Classification of Nara language in Eritrea". *Journal of Eritrean Studies* 4/1-2: 1-27.

RILLY, Claude (2010) *Le méroitique et sa famille linguistique*. Louvain and Paris: Peeters.

SAVÀ, Graziano (2018) “The Nara project: achievement and perspectives”. *Rassegna di Studi Etiopici* 2 (third series): 145-156.

SAVÀ, Graziano (2023) *Grammatica introduttiva della lingua amarica*. Naples, UniorPress.

THOMPSON, E. David (1976a) “Nera”. In BENDER, M. Lionel (ed.), *The non-Semitic languages of Ethiopia*. Pp. 484-494. East Lansing and Carbondale: African Studies Center of the Michigan State University and Southern Illinois University.

THOMPSON, E. David (1976b) “Languages of northern Eritrea”. In BENDER, M. Lionel (ed.), *The non-Semitic languages of Ethiopia*. Pp. 597-603. East Lansing and Carbondale, African Studies Center of the Michigan State University and Southern Illinois University.

TREIBER, Magnus (2007) “Nara ”. In UHLIG, Siegbert (ed.) *Encyclopaedia aethiopica, Vol. 3 (He-N)*. Pp. 1144a-1147a. Wiesbaden: Harrassowitz Verlag.

Conversation Analysis for Language Documentation: Interactional Perspectives on Lesser-described Languages

Daniela Veronesi
Free University of Bozen-Bolzano
daniela.veronesi@unibz.it

ABSTRACT

The paper offers a preliminary overview of studies on lesser-described languages carried out in recent decades within the theoretical and methodological framework of Conversation Analyis, an ethnomethodologically inspired approach to the investigation of language use and social interaction. Reported results of large-scale comparative investigations on fundamental conversational mechanisms like turn-taking, sequential organization and repair thereby show the universal character of such practices, while highlighting peculiarities of languages like ǂĀkhoe Hai||om, Cha'palaa, Murrinh-Phata, Yélî-Dnye and Tzeltal, in comparison with widespread Western European languages. It is thus discussed how Conversation Analysis – with its data-driven, emic, situated, and multimodal perspective on spoken interaction – can fruitfully complement language documentation work on lesser-described languages and speech communities.

Keywords: *Conversation Analysis, lesser-described languages, language use, ordinary interaction*

DOI: 10.23814/ethn.20.24.ver

1. Introduction

Over more than five decades of research, Conversation Analysis (cf. Sidnell & Stivers 2013) – driven by the assumption that social interaction, from the most casual conversation to institutional encounters, is organized in specific ways that are displayed by and to participants through verbal and non-verbal conduct – has provided extensive insights on the mechanisms that underlie talk-in-interaction in a variety of settings, as well as on language practices and language structures as they are thereby deployed to accomplish social actions. While studies in this field initially focussed on English, conversation analysts' interest gradually expanded to further Western European languages, as well as, in more recent years, to languages from other language families and groups, among which also lesser-described languages.

After a brief characterization of Conversation Analysis (CA) and of the fundamental mechanisms governing of social encounters as outlined within its theoretical and methodological approach – e.g. turn-taking, sequence organization, and repair – (§ 2), in this paper I will go over some large-scale CA-oriented studies on such mechanisms (§ 3.1-3.3.), as well as on further topics explored, for lesser-described languages, from a conversational perspective (§ 3.4). I thus aim at providing a first, exploratory overview of what CA has provided so far in the field, and at the same time at discussing how the emic, naturalistic, and action-oriented approach of CA can offer relevant descriptions of language use, and language structures, for small-scale, indigenous, lesser-described languages, this way possibly complementing established methods in the field of language documentation.

2. Investigating social interaction from the perspective of Conversation Analysis

Conversation Analysis (Sacks, Schegloff & Jefferson 1974; Sacks 1992a) is an ethnomethodologically inspired approach to language use and human action, studied inductively on the basis of naturalistic data – audio and videorecordings of naturally occuring interactions, which are subsequently transcribed in their finest details (Jefferson 2004; Hepburn & Bolden 2013) –, with the aim of describing the procedures used by social actors to produce and recognize interactional conduct, and, thereby, the competences that "ordinary speakers use and rely on in participating in intelligible socially organized interaction" (Heritage & Atktinson 1984: 1). Since its inception in the late 1960s, CA – progressively gaining attention within a variety of disciplines like linguistics, anthropology, communication studies, and psychology, to name but a few – has provided detailed investigations of fundamental mechanisms governing talk-in-interaction such as turn-taking – with the basic rule "one speaker at a time" – and the phenomenon of "repair", that is, the way in which social actors address problems in "speaking, hearing and understanding" (Schegloff, Sacks & Jefferson 1977). A further, primary organizational principle studied by CA scholars is the sequential organization of conversation in "adjacency pairs" – a first action by a speaker 'calling for' a second action by the interlocutor (such as "greeting-greeting", "question-answer", "request-compliance" and the like, cf. Schegloff & Sacks 1973) –, and the underlying system of (social) "preference" governing the relationship between the two actions. A request, for instance, is expected to be fulfilled, a question makes a particular type of answer relevant, a first assessment is typically followed by an agreeing second assessment, and the like (cf. Pomerantz & Heritage 2013). Conversation analysts have thereby examined how such mechanisms are differently deployed in ordinary, mundane conversations (for instance, among friends) and in "institutional" settings – classrooms, hospitals, courtrooms, radio and TV studios etc. –, showing how in these latter ones participants typically orient to institution-specific goals and to restrictions on the nature of their interactional contributions (Drew & Heritage 1992; Heritage 2005).

Interested in the sequential and temporal organization of talk and action, and, more and more in recent decades, in the way different semiotic modalities (talk, gesture, gaze, body movements etc.) are integrated so as to form coherent courses of action (cf. Goodwin 1981; Mondada 2009; Streeck, Goodwin & LeBaron 2011; Haddington, Mondada & Nevil 2013; Goodwin 2013; Hazel, Mortensen & Rasmussen 2014; Mondada 2019; Depperman & Streeck 2018; Haddington et al. 2023, among others), as an analytical perspective CA focuses on the collaborative and interactional aspects of speakers' conduct, thus going beyond the boundaries of isolated, verbal "sentences" or "utterances" – for a long time the traditional focus of descriptive linguistics and pragmatics (cf. Drew 2018) –, to investigate turns-at-talk as well as visible and audible behaviour in their sequential context within interaction, and as a joint accomplishment of conversational partners.

As mentioned above (cf. §1), in its initial phase Conversation Analysis was mainly confined to English, to then expand to further widespread Western European languages – spoken in urban settings, with official status and writing systems –; against this background, as Dingemanse & Floyd (2014: 453) point out, the inclusion of ethnographic understandings to the analysis of language use in interaction was not widely appreciated, and was not part of the CA program, although scholars, as native members of the analyzed societies, were "able to rely on their own assumptions […] for providing cultural context" (Dingemanse & Flyod 2014: 453).

What about, though, small-scale, unwritten languages, often spoken in remote places? While such languages, which are fundamental in terms of the world's linguistic diversity, have historically been the focus of cross-linguistic grammatical typology within descriptive linguistics – but have been thereby rarely compared in terms of interactive structures and practices – and linguistic anthropology has traditionally had a special interest for formalized language use, as Dingemanse & Floyd (2014) argue, it is only recenty that CA and its methodology, complemented by ethnographic fieldwork[1], has been applied for cross-cultural comparative studies of conversation – starting with ordinary interaction – which also include lesser-described languages, as will be seen in the following.

3. CA and lesser-described languages: a preliminary overview

As mentioned above (cfr. §2), turn-taking, sequence organization and repair are considered the three fundamental 'pillars' of social interaction; in the following, I will illustrate the contribution provided by Conversation Analysis as pertains their functioning in some lesser-described languages, by mainly drawing, as a way of example, on results of large-scale comparative studies carried out within the Language and Cognition Department of the Max-Planck Institute for Psycholinguistics under the leadership of Stephen C. Levinson (1994-2017). Further investigations on selected topics, explored from a conversational perspective, will also be mentioned here.

3.1. Turn-taking

As a mechanism for coordinating verbal interaction, turn-taking is fundamental for the regulation of who is to speak next and when, in any given social encounter. As outlined in the seminal work by Sacks, Schegloff & Jefferson (1974), in ordinary, informal conversations – by now largely studied by CA for English and further Western European languages and beyond –, turn alternation among speakers follows some basic organizational principles, that is, a current speaker can select the next speaker ("next speaker selection", as in a question addressed to a specific interlocutor), while, if this is not the case, anyone can take the floor; if nobody self-selects, the current speaker may but need not continue. Furthermore, speakers tend to avoid overlapping talk ("one speaker at a time") and to minimize gaps and silence between turns-at-talk: since utterances (e.g. "turn-constructional units", cf. Sacks, Schegloff & Jefferson 1974) can be grammatically or intonationally complete, as well as pragmatically embody a given action (cf. Clayman 2013), turn completion is to some extent predictable and recognizable (e.g. "projectable") by participants, which makes 'smooth' turn transition (no overlaps, no gaps) possible.

From a pragmatic typological perspective, the question thus arises as to whether such turn-taking machinery has a fundamentally universal character: to test this hypothesis – against the background of widespread anthropological claims of radical cultural variability in the timing of conversational turn-taking – CA and pragmatics scholars participating in the "Multimodal Interaction Project" (Language and Cognition Group, Max-Planck Institute for Psycholinguistics, cf. above) examined selected data from a corpus of video recordings of informal, natural conversation (dyadic and multi-participant interactions) in 10 languages from all over the world, namely, Germanic languages (English, Danish, Dutch), Italian (Romance), Japanese (isolate), Korean

[1] On the relationship between CA and ethnography, cf. also Warfield Rawls & Lynch (2024).

(Ural-Altaic), Lao (Tai), as well as languages from traditional indigenous communities like ǂĀkhoe Hai||om in Namibia (Koisan language group) Tzeltal in Mexico (Mayan), and Yélî-Dnye (isolate, spoken in Papua New Guinea).

To allow comparability, the study (Stivers et al. 2009) was based on the analysis of turn transition between "polar questions" (questions that expect a 'yes' or 'no' answer, as in *Have you heard from her?* and *Have you heard from her yet?*, respectively, and the most common question type in 9 of 10 languages in the corpus (corpus, e.g. 350 questions for each language) and measurements in milliseconds of the temporal relation between a question and its response (response timing). Possible factors for delayed answers, well documented in the existing literature for English, were also taken into account: this is the case of responses that do not answer the question (nonanswer responses, as *I don't know*) or that run against the bias of the question (disconfirmations, e.g. A: *Is that your car?* B: *No.*), and which are typically delayed. Furthermore, it was considered that vocal answers may be preceded by nonverbal signals (head nods, head shakes) on the one hand, while a speaker's gaze toward the listener may increase the pressure to respond quickly, on the other.

Results, as Stivers et al. (2009) point out, show strong parallels in turn-taking across languages: responses tend to be neither in overlap nor delayed by more than a half-second (with a continuum that goes from faster to slower responses, see Japanese, +7ms, and Danish, +469ms), while the factors that affect response timing are the same across languages; as a matter of fact, speakers of all languages provide answers significantly faster than nonanswer responses, and confirming answers faster than nonconfirming ones. As for the role of nonverbal conduct, it was noted that visible responses (through head nods, head shakes, shrugs etc.) are faster than speech across the whole corpus – although their inclusion as visible component varies across languages –, and that in 9 out of 10 languages responses are delivered earlier if the speaker is gazing at the recipient while asking the question, with statistical significance in 5 languages, thus hinting at a larger cultural variability of visible conduct as opposed to speech.

All this speaks for a universal organization of turn-taking in informal conversation, which is aimed at minimizing gap and overlap, as well as for a "universal semiotics of delayed response" (Stivers et al. 2009: 10591). Against this background, documented cross-cultural differences in response timing, as mentioned above, would not mine the universal hypothesis, but rather point to differences of "interactional tempo" across cultures, related to what counts as a delay in response because of the specific cultural interactional pace or the overall tempo of social life in a given culture[2].

According to Stivers et al. (2009: 10590), they thus constitute "minor variation in the local implementation of a universal underlying turn-taking system": a conclusion which not only allows to substantiate ethnographic reports on widely studied languages[3], but also to contribute to the description of indigenous languages like the ones included in the corpus, through the examination of spoken language in its natural habitat, that is, as it is used in social interaction (cf.; Schegloff 2006; Couper-Kuhlen & Selting 2018: 3).

[2] By taking into account the relative conversation's rhythm of the examined languages, a silence of 200 ms was judged as a delay in most languages, while a response given after was still considered on time in Danish and Lao (cf. Stivers et al. 2009: 10590).

[3] Japanese speakers, for instance, are said to respond after substantial gaps of silence, while in this study they are "on average, the earliest to respond"; Italian speakers are supposed to be more tolerant of overlap, but Stivers and colleagues found out that only 17% of all responses overlap, and that Italian speakers "leave a slightly longer [+310 ms] than average gap before producing a next turn" (Stivers at al. 2009: 10591).

3.2. Sequence organization: the example of questions and answers

A further area which was investigated by Levinson and colleagues, within the "Multimodal Interaction group's Question and Response Project", is sequence organization (Schegloff & Sacks 1973), namely, the way in which turns-at-talk, as means to perform social actions, are "positioned either to initiate a possible sequence of action or to respond to an already initiated action as part of a sequence" (Stivers 2013: 191); social actions, that is, occur sequentially and are organized in sequences, the most basic one being the adjacency pair (see above, 2).

Scholars (Stivers, Enfield & Levinson 2010) thereby focussed on question-response sequences, with the aim of describing and comparing the way questions are formally coded, via lexical/morphosyntactic and/or prosodic marking, in the 10 examined languages (see above)[4]. The social actions that can be performed through questions were also taken into account, namely "information requests" (so called 'true' questions, as in *What time is it?*), "requests for confirmation" (as in *Is that your car?*), "assessments" which make relevant an agreement (for instance *Isn't it beautiful out today?*), suggestions, offers and requests – typically, in English, oriented to by speakers in terms of acceptance and compliance –, and, last but not least, questions which address conversational problems (so called "repair-initiation", cf. §2). From a sequential perspective, responses were also examined, so as to test the generalizability of normative preferences for responding to questions as they were outlined for English in the CA literature – not only the fact that "a question creates a sequential context where an answer is expected, and makes the addressee accountable if an answer is not forthcoming" (Brown 2010: 2638, cf. Schegloff & Sacks 1973), but also, and more in detail, the specific interplay between question design and 'preferred' answer (cf. Hayano 2013) –, while inquiring about the overall role of non-verbal conduct (e.g. nodding, eye gaze) in question-response sequences.

Against the background of a general consistency across languages as pertains "the strong propensity for questions to be immediately followed by answers" (Brown 2010: 2647), it was thus possible to highlight language-specific peculiarities. In Tzeltal, for instance, as Brown (2010) observes, question-answer sequences strongly diverge from findings based on English conversation, in that Tzeltal speakers (comparatively) minimally deploy gaze in next-speaker selection and do not provide visible-only responses – mutual gaze being very restricted in this community. As for the actions performed through questions in Tzeltal, furthermore, it was noted how suggesting, requesting and offering are relatively infrequent, and how confirmation requests – via full or partial repetition of the interlocutor's prior utterance offering new information – , in turn, are more frequent than information seeking requests. In other terms, as Brown points out, Tzeltal "routinely request confirmation of new information just supplied, not necessarily because they didn't hear or don't believe it, but simply to firmly establish it in common ground before proceeding with the topic" (Brown 2010: 2638). Also, repeats are the most frequent form for positive (affirming) answers to polar questions, while in English these are typically answered with a 'yes' or 'no', and repeats carry additional implications (for instance, a challenge to the question as formulated). This "repeat-as-affirmation response" has been in fact attested in other Mesoamerican languages, as well

[4] A subsequent study into sequence organization (Kendrick et al. 2020) also includes Argentine Sign Language, Mandarin Chinese, Turkmen and Yurakaré (Bolivia).

as in Welsh, Estonian, Finnish and Japanese (cf. Brown 2010: 2640) and is thus not peculiar to Tzeltal only. Nevertheless, for this latter language there is strong evidence, also from Brown's previous studies, that repeats are part of the response system as a *default* way to agree with a prior utterance (not just as the expected response to polar questions, but also in non-question utterances): an insight which could open the door, as Brown (2010: 2647) suggests, to future, quantitative cross-linguistic studies on the association between repeats and agreement.

A further example of linguistic and cultural specificities highlighted by the project, worth mentioning here, comes from ǂĀkhoe Hai||om (a Koe language of the Khoisan family spoken in Northern Namibia), which stands out, as for the distribution of question types, for a predominance of content questions (e.g. questions introduced by a question word, as in *What did you do last night?*), as opposed to polar questions, which were the majority for all other considered languages (Hoymann 2010). As a matter of fact, ǂĀkhoe speakers ask more repair initiating questions than speakers of the other languages, and they do so mostly through content questions; furthermore, they never request confirmation – in other languages requests for confirmation make up between 20% and 50% of all questions –: a fact that explains why polar questions, which would be typically used for requesting confirmation, are fewer for this language.

The reasons for such preference can be found, according to Hoymann (2010), in the social hunter-gatherer culture of ǂĀkhoe speakers, leading them "to pose questions in a way that is less coercive and less restrictive of the answerer than speakers of other languages do" (Hoymann 2010: 2736): indeed polar questions constrain the answerer to a 'yes' or 'no' answer, while content questions provide the interlocutor a greater 'freedom' in choosing how to answer; moreover, requests for confirmation, in the form of polar questions biased towards a yes answer, are highly coercive.

As Hoymann (2010: 2737) points out, anthropological studies on hunter-gatherer societies, particularly in Australia and Southern Africa, have claimed that their conversational style is characterized by the "acceptance of long silences, more overlap and less next speaker selection, which the researchers argue give the conversational partners more freedom, or in other words, are ways of being less coercive". Similarly, in ǂĀkhoe Hai||om "silences are relatively long compared to those of the other languages in the questions project" (Hoymann 2010: 2737) – thus documenting a cultural difference in the speed of the responses, which, though, does not affect the overall structure of the turn-taking system (see above, 3.1.) – and ǂĀkhoe speakers select a next speaker, that is, address questions to a specific individual, "relatively less often than speakers of the other languages do" (Hoymann 2010: 2737). Finally, data show a high number of questions (23%) that obtain no response at all, which also hints at "a greater concern for other speakers' independance", in accordance with anthropological evidence on hunter-gatherers' conversational style (Hoymann 2010: 2739). This, though, does not lead to the conclusion – as Hoymann argues – that there should be a *direct* cause-and-effect relationship between type of society (hunter-gatherers with an egalitarian lifestyle) and a high number of unanswered questions; rather, it is "the way in which these speakers pose the question" – resorting to content rather than polar questions – "that makes it possible not to answer questions" (Hoymann 2010: 2379). The influence of culture on conversational style, that is, is not to be seen at the level of the sequence of utterances (question-response), but rather at that of the function of these utterances, so that for the ǂĀkhoe Hai||om speech community "a reluctance to pose direct questions, or questions that strongly pressure recipients to answer, leads to a higher proportion of content questions or open questions and almost no requests for confirmation" (Hoymann 2010: 2379).

3.3. Repair

As mentioned above, "repair" is one of the fundamental mechanisms of social encounters, allowing to address problems as they emerge in the here-and-now of interaction; indeed, as Digemanse et al. (2015: 1) observe, "there would be little adaptive value in a complex communication system like human language if there were no ways to detect and correct problems". The way in which social actors interrupt the ongoing course of action to address, and solve, possible trouble in speaking, hearing or understanding was one of the mechanisms to be first explored in CA studies by Schegloff and colleagues (Schegloff, Jefferson & Sacks 1977; Kitzinger 2013), who described it in detail, differentiating between "repair initiation" (signalling a problem, that is, a so called "trouble source") and "repair completion" (solving the problem), as well as in terms of who accomplishes repair, namely the trouble source speaker ("self-repair") – preferred in ordinary conversation – or the interlocutor ("other-repair").

Repair was also one of the targets of the Multimodal Interaction group led by Levinson (see above), with investigations focussing on "other-initiated repair": cases in which the recipient of an unclear message signals trouble (*Huh? What?*, as well as through a "candidate solution" as in *You mean x?*), and the current speaker can 'repair' the original message and the trouble source, for instance by repeating this latter one or by confirming the interlocutor's candidate understanding, typically in the form of a question-and-answer exchange between interlocutors.

For this study (cf. Dingemanse et al. 2015; Dingemanse & Enfield 2015), the sample was represented by 12 languages of 8 language families – Dutch, English, Italian, Lao, Yélî-Dnye, but also Icelandic, Russian, Argentinian Sign Language, Cha'paala in Equador (Barbacoan language family), Murrinh-Patha from Northern Australia (Southern Daly) and Siwu in Ghana (Kwa) –, with ca. 50 hours of recordings (4 hours per language); through the examination of more than 2,000 cases of other-initiated repair, it was thus observed that not only this is frequently used (on average about once per 1.4 minutes in any language), but also that it has common properties across languages, thus disconfirming assumptions of radical cultural variation.

Firstly, according to analyses, all languages share three basic types of repair initiator, namely a) "open request", which signals a problem but does not specify its location or nature (as in *Huh? Sorry?*); b) "restricted request", focussing on specification/clarification of a specific element of the trouble source (as in A. *Oh, Sibbie's sister had a baby boy;* B. *Who?*); c) "restricted offer", through which the interlocutor provides a candidate understanding and asks for confirmation (A. *She had a boy?*; B. *You mean x?*).

Secondly, these three basic types of repair initiation, representing the majority of cases in all languages, are accomplished using similar linguistic resources – interjections, question markers, prosody, and repetition of the trouble source turn –, and are drawn upon systematically, across languages, depending on the same contextual factors. So, for instance, open requests are more likely when, due to noise, overlapping talk or distractions to the listener's attention (who might be engaged in a parallel activity), troubles in hearing, or processing what someone just said emerge; if an open repair initiation does not lead to a solution, further repair initiation is then done through more specific types.

In all languages, thirdly, listeners opt for more specific repair initiators over less specific ones whenever possible, this way minimizing joint collaborative work in the resolution of the problem, keeping disruptions to the progression of conversation to a minimum[5], and thereby displaying an altruistic behaviour: a fundamental orientation that was first proposed by CA studies for English conversation, and that, according to the project results, also applies across languages. Findings thus support a pragmatic universal hypothesis: while languages, as Dingemanse et al. (2015: 9) conclude, "may vary in the organization of grammar and meaning, key systems of language use may be largely similar across cultural groups", one of them being other-initiated repair. Against this background, the detailed, qualitative and quantitative analyses carried out for this study have also provided insights on the pragmatics of previously unexplored communicative practices within the repair machinery, in all examined languages (cf. also Enfield et al. 2013); for the purposes of the present article, some findings for Murrinh-Phata, Yélî Dnye and Cha'palaa will be briefly mentioned here.

For Murrinh-Phata, an aborigenal lingua franca spoken by around 2,700 people in communities in Australia's Northern Territory, it has been noted how, against the background of widespread avoidance of certain personal names (recently deceased, certain in-laws and siblings), "candidate repairs" – restricted offers seeking confirmation, see above – are particularly useful "for handling complications relating to the domanis of *place* and *person*, such as person identification when name avoidance is an issue" (Blithe 2015: 295). A further specificity is that the two open formats documented in the Murrinh-Phata data – the interjection "Aa?" (*huh?*) and the question word "thangku" (*what*) –, although occuring in the same interactional contexts, differ in terms of the type of trouble they address and the repair solution they yield, contrarily to "huh" and "what" in English, for which no functional or interactional differences have been documented so far (cf. Drew 1997 and Robinson 2006, quoted in Blithe 2015: 302). As a matter of fact, "Aa?", which is four times more frequent than "tangku", is used as repair initiator when issues of audibility or misaligned recipiency – targeted recipients attending to something/someone other than the speaker, thus not having been listening attentively enough to produce a response, and therefore initiating repair – are at stake (Blithe 2015: 297); "thangku", instead, is more often deployed for dealing with talk that is problematic in terms of speaker's intended meaning, relevance or possible topical disjuncture.

Insights on how other-repair initiators may both conform to well-known European patterns as well as deviate from them are also provided by Levinson's analysis of Yélî Dnye, a 'Papuan', i.e. non-Austronesian language, spoken by ca. 5,000 people on Rossel Island in Papua New Guinea (Levinson 2015: 386). Similarly to Murrinh-Phata, in Yélî Dnye the use of personal names (property of the clan of the father) is restricted, yielding frequent requests for referent clarification; since Yélî Dnye speakers systematically "try minimized reference and escalate only as required, step by step providing additional material in a specific order [e.g. person/number marking on a verb; kinterm; name plus kinterm] until referent identification succeeds" (Levinson 2015: 387, cf. also Levinson 2007), extended repair sequences initiated through restricted formats (wh-questions and polar questions) arise, suggesting that the high level of other-initiated repair previously

[5] Such "conservation principle" (Dingemanse et al. 2015: 7-8) is also documented, in all languages, by the very short duration of repair completion, which lasts about the same temporal lenght of the trouble source turn, hinting at participants' joint effort for efficient problem resolution.

documented for Yélî Dnye (Levinson 2010) "might be partly motivated by culture-specific norms for 'under-telling'"(Levinson 2015: 387).

As for open repair initiators, remarkably, Yélî Dnye does not have forms that are equivalent to *what* or *sorry*; the only open class repair initiator is ":aa?", or ":êê?" (a long nasalized low or central vowel, uttered with rising intonation), which, similarly to other languages, is employed to address acoustic problems. This scarsity of linguistic forms dedicated to open other-initiated repair is compensated, though, by the fact that Yélî Dnye speakers consistently draw upon non-verbal resources both for initiating repair – a "frozen look" (a fixed stare by the recipient), produced without moving – and completing it, as when a trouble source speaker engages in a slight brow rise to confirm a listener's candidate repair, before uttering a verbal confirmation, if any.

That such practices are to be found in other language communities is in fact attested; nevertheless, in Yélî Dnye a heavy reliance on the visual modality seems to be a quite regular feature, made possible also by the community's preference for dyadic face-to-face interactions – which facilitates the use of visual signs such as gestures and facial expressions –, and grounded in the existance of a "culturally conventionalized inventory of facial signals like the affirmatory blink and the affermative eyebrow rise" as well as in the culturally "unusually sustained nature of mutual gaze that is required if visual signals are to be reliable" (Levinson 2015: 407).

An extensive role of non-verbal conduct in accomplishing repair was also documented for Cha'palaa – until recently, a relatively unstudied language –, spoken by the Chachi people (ca. 10,000 speakers) "in small communities and households along the rivers of the Equadorian Province of Esmeraldas between the Andean foothills and the Pacific coast" (Floyd 2015: 467). As a matter of fact, following other-repair initiation related to reference (as in "mu-nu-n", *to whom?*), Cha'palaa speakers can use lip-pointing (towards the referent), while they draw on eyebrown flashes – a brief raising of the eyebrow – to confirm the interlocutor's candidate repair solution (Dingemanse & Floyd 2014: 465); two culturally conventionalized bodily practices which diverge from the ones English speakers would rely upon in similar contexts (e.g. index finger pointing for personal reference, and head nodding for confirmation, cf. Whitehead 2011). In addition, in Cha'palaa conversation *open* other-repair initiation predominates over *restricted* types, as opposed to a general tendency for the opposite in the majority of the other languages examined: Cha'palaa speakers thus show "a preference for displaying hearing problems over understanding problems" (Floyd 2015: 472), possibly because open repair is neutral with respect to responsibily for the problem, while still allowing the trouble source speaker to repeat their turn but also to reformulate it. A further peculiarity of Cha'palaa is the fact that the interjection used for open repair initiation, that is, a long vowel /a/ with slight pre-glottalization (/?a:/), is produced with falling intonation – rather than with rising intonation as in all other languages, with the exception of Icelandic –, this way conforming to Cha'palaa standard falling prosody for content questions, and thus fitting into the wider phonological system of the language.

3.4. Further areas of CA investigations into lesser-documented languages

Investigations carried out by CA scholars on the relationship between cultural and linguistic diversity and pragmatic universals have not been confined to the above mentioned studies, but have extended to a number of further areas of conversational structures: so, for instance, within a larger project led by Nick Enfield (*Human Sociality and Systems of Language Use*, HSSLU 2010-2014) and which also involved members of the Max-Planck Institute (see above), "recruitments" – getting others to do things, as in requests, offers, and suggestions – were examined in Cha'palaa, English, Italian, Lao, Murrinh-Phata, Polish, Russian and Siwu (Floyd, Rossi & Enfield 2020)[6], with findings providing insights on the differences between the pragmatic systems of these languages, but also showing relatively low cross-cultural diversity, "in line with the idea that a species-wide infrastructure for interaction underpins the use of language, largely independent of the specific shape of that language" (Rossi, Floyd & Enfield 2020: 15).

A detailed study within the HSSLU project was also devoted to the expression of gratitude in the same corpus (Floyd et al. 2018), carried out under the assumptions that a) social reciprocity – involving the mutual exchange of goods, services and support, and the ability of individuals to experience gratitudine – is a fundamental aspect of human organization, and that b) to understand the role of gratitude in the maintenance of social reciprocity, a differentiation between the *experience* of gratitude (as an emotion) and the (verbal or non-verbal) *expression* of gratitude (as a linguistic practice) should be drawn, with this latter to be observed in naturally occuring interactions, rather than in controlled laboratory conditions or on the basis of self-report questionnaires, as it was the case in previous research, mostly related to English. Thus, by drawing upon audio-video recorded episodes of ordinary, informal interactions of people who know each other well, and in which someone seeks and obtains a good, support or service from the interlocutor, the study examined whether and insofar the compliance of such "here-and-now" requests – for actions that are relatively straightforward and low-cost, as in *Can you pass me the salt?* – is responded to, by the requester, with verbal expressions of gratitude (as in *thank you*, *sweet* or other positive formulations), as well as with non-verbal acknowledgments like nods or hand gestures. Through the analysis of approx. 200 request and response sequences per language, it was thus shown how – as opposed to social and prescriptive attitudes about politeness found in English-speaking society – "the general norm is to tacitly acknowledge another's cooperative behaviour without explicitely saying 'thank you' […] relying on a shared understanding of the good, service or support received as part of a system of social rights and duties governing mutual assistance and collaboration" (Floyd et al. 2018: 3).

As a matter of fact, against the background of a generalized tendency towards fulfilment of requests across languages, it was noted how expressions of gratitude by requesters are quite infrequent (5.5% of 928 cases of successful requests), with minor but significant variation among languages, which shows how the expression of gratitude is more common in certain languages (here English, 14.5% and Italian, 13.5%) than others (Murrinh-Patha 4.5%, Russian 3.1%, Polish 2.2%, Lao 2%, Siwu 0.8%, and finally 0% in Chaa'paala, which does not even have a conventional way to say 'thank you')[7]. This, on the one hand, speaks once again for the caution that should be used

[6] Thereby it was focussed on recruitments of pratical actions (transfer of an object, performance of a manual task, or alteration of an ongoing bodily movement) to be performed immediately, and the way in which they are accomplished through linguistic resources – imperative, interrogative or declarative sentences, use of modal verbs, explanations, mitigations – as well as non-verbal conduct (gazing at the interlocutor, pointing at a desired object, etc.). Similarly, responses to recruitments – fulfilling, rejecting, or ignoring the recruiting action - were examined as for the inclusion of verbal and visible elements (for instance, "yes", "sure", "no", head nods, head shakes, cf. Rossi, Floyd & Enfield 2020a and Floyd, Rossi & Enfield 2020b).

[7] As Floyd and colleagues remark, though, "Lao, Polish, Russian and Murrinh-Patha are not statistically different from Siwu" (Floyd, Rossi & Enfield 2018: 6), and not affected by social-interactional variables as for instance interlocutors' higher or lower status, while English and Italian frequences are still low (in one out of seven episodes of fulfilled requests); in these latter languages, in fact, expressions of gratitude may occur mostly in institutional

when coming to generalizing conclusions based on speakers of English and other Western European languages; on the other hand, it demonstrates, as the authors conclude, that "[d]espite the attitudes encountered in some cultures that emphasize saying 'thank you' often, such practices do not appear to be necessary for the maintenance of everyday social reciprocity" (Floyd, Rossi & Enfield 2018: 8)[8].

Not to be forgotten here are also the detailed investigations carried out on Australian Aborigenal languages, over nearly two decades, by Rod Gardner and Ilana Mushin, who explored conversational practices in Garrwa – spoken by some remote Aborigenal communities in Northern Australia –, such as turn-taking (Gardner & Mushin 2015; Mushin & Gardner 2009; Gardner & Mushin 2007) and question-answer sequences (Gardner 2010), and joined forces with colleagues to compare various Australian Aborigenal languages as pertains next-speaker selection (Blythe, Gardner, Mushin & Stirling 2018), as well as verbal and visible practices of personal reference and the thereby involved issues of epistemics (Blythe, Mushin, Stirling & Gardner 2022). The notion of epistemics, as outlined in CA (cf. Heritage 2012; Stivers, Mondada & Jakob 2021), has also been fruitful in further studies on languages in small communities, as detailed in a special issue of the *Journal of Pragmatics* dedicated to the topic, and which includes investigations on Datoga in Tanzania (Mitchtell & Jordan 2022), Quechua in Equador (Grzech 2022) and Tzeltal, Yucatec and Zapotec in Mexico (Brown, Sicoli & Le Guen 2022).

4. Concluding remarks

In this exploratory paper I have examined how Conversation Analysis, with its orientation to language, language practices and language structures as situated in the here-and-now of social interaction, can provide a relevant contribution to the study of linguistic diversity, and, within this field, to the investigation of lesser-described languages from a praxeological perspective of language use in context, thereby offering its theoretical approach and its methods to the field of language documentation.

As a way of illustration, it was thus shown how basic mechanisms governing spontaneous, informal conversation such as turn-taking, sequence organization and repair have been compared cross-culturally, within some large-scale CA-oriented projects, through a detailed examination of how turns-at-talk are linguistically constructed, how they are embedded in sequences of actions, and by taking into account the role non-verbal conduct plays in performing social actions.

Further areas of CA investigations that I have mentioned here include recruitments, the expression of gratitude and issues of epistemics; far from being exhaustive, this list hints at the explicative potential of Conversation Analysis when it comes to describe language use in spoken interaction for lesser described languages: an area for which the conversation analytical theoretical and methodological framework, supported by ethnographic work, can fruitfully meet with, and integrate – as Dingemanse & Floyd (2014: 467) suggest for the comparative study of social interaction – linguistic and sociocultural anthropology, descriptive linguistics, corpus linguistics and gesture studies, and thus help to provide investigations that are ecologically valid (e.g. with language and social interaction studied in everyday face-to-face interaction as their

contexts and with strangers, as a comparison carried out by the authors with findings based on service encounters suggests (cf. Floyd, Rossi & Enfield 2018: 8).

[8] See also Zinken, Rossi & Vasudevi (2020), comparing the expression of gratitude in British English, German, Italian, Polish, and Telugu.

natural-cultural habitat), ethnographically enriched, empirically grounded (that is, based on large records of data available for repeated inspection), multimodal, and comparable, that is, with data from comparable settings and sequential environments (cf. Dingemanse & Floyd 2014: 467).

References

BLYTHE, Joe (2015) "Other-initiated repair in Murrinh-Patha", *Open Linguistics* 1: 283-308.

BLYTHE Joe, Rod GARDNER, Ilana MUSHIN, and Lesley STIRLING (2018) "Tools of engagement: selecting a next speaker in Australian Aborigenal multiparty conversations", *Research on Language and Social Interaction* 51(2): 145-170.

BLYTHE Joe, Ilana MUSHIN, Lesley STIRLING, and Rod GARDNER (2022) "The epistemics of social relations in Murrinh-Patha, Garrwa and Jaru conversations", *Journal of Pragmatics* 91: 175-193.

BROWN, Penelope (2010) "Questions and their responses in Tzeltal", *Journal of Pragmatics* 42: 2627-2648.

CLAYMAN, Steven E. (2013) "Turn-Constructional Units and the Transition-Relevance Place". In SIDNELL Jack, and Tanja STIVERS (eds.), 150-166.

COUPER-KUHLEN, Elizabeth, and Margret SELTING (2018) *Interactional linguistics. Studying language in social interaction*, Cambridge: Cambridge University Press.

DEPPERMANN, Arnulf, and Jürgen STREECK (eds.) (2018) *Time in Embodied Interaction: Synchronicity and sequentiality of multimodal resources*, Amsterdam: Benjamins.

DINGEMANSE, Mark, and Nick J. ENFIELD (2015) "Other-initiated repair across languages: Towards a typology of conversational structures", *Open Linguistics* 1: 96-118.

DINGEMANSE, Mark, and Nick J. ENFIELD (eds.) (2015a) Other-initiated repair across languages, *Open Linguistics* 1 (96-577), special issue.

DINGEMANSE, Mark, and Simeon FLOYD (2014) "Conversation across cultures". In ENFIELD, Nick J., Paul KOCHEMANN, and Jack SIDNELL (eds.), *The Cambridge Handbook of Linguistic Anthropology*. Pp. 447-480. Cambridge: Cambridge University Press, Cambridge.

DINGEMANSE Mark, Seán G. ROBERTS, Julija BARANOVA, Joe BLYTHE, Paul DREW, Simeon FLOYD, Rosa S. GISLADOTTIR, Kobin KENDRICK, Stephen C. LEVINSON, Elizabeth MANRIQUE, Giovanni ROSSI, and Nick J. ENFIELD, (2015) "Universal principles in the repair of communication problems", *PLOS ONE* 10(9): 1-15.

DREW, PAUL (1997) "'Open' Class Repair Initiators in Response to Sequential Sources of Troubles in Conversation", *Journal of Pragmatics* 28: 69-101.

DREW, Paul (2018) "The interface between pragmatics and conversation analysis". In in ILIE Cornelia, and Neal R. NORRICK (eds.), *Pragmatics and its Interfaces*. Pp. 59-84. Amsterdam: Benjamins.

DREW, Paul, and John HERITAGE (eds.) (1992) *Talk at Work*. Cambridge: Cambridge University Press.

ENFIELD, Nick J., Mark DINGEMANSE, M., BARANOVA, Joe BLYTHE, Penelope BROWN, DIRKSMEYER, Tyko, Paul DREW et al. (2013) "Huh? What? – A first survey in 21 languages". In HAYASHI, Makoto, Geoffrey RAYMOND, and Jack SIDNELL (eds.), *Conversational Repair and Human Understanding*. Pp. 343-380. Cambridge: Cambridge University Press.

FLOYD, Simon (2015) "Other-initiated repair in Cha'palaa", *Open Linguistics* 1: 467-489.

FLOYD, Simeon, Giovanni ROSSI, and Nick J. ENFIELD (eds.) (2020a) *Getting Others to Do Things: A Pragmatic Typology of Recruitments*, Berlin: Language Science Press.

FLOYD, Simeon, Giovanni ROSSI, and Nick J. ENFIELD (2020b) "A coding scheme for recruitment sequences in interaction". In FLOYD, Simeon, Giovanni ROSSI, and Nick J. ENFIELD (eds.), *Getting Others to Do Things: A Pragmatic Typology of Recruitments*. Pp. 25-50. Berlin: Language Science Press.

FLOYD, Simon, Giovanni ROSSI, Julija BARANOVA, Joe BLYTHE, Mark DINGEMANSE, Kobin H. KENDRICK, Jörg ZINKEN, and Nick J. ENFIELD (2018) "Universals and cultural diversity in the expression of gratitude", *Royal Society Open Science* 5: 180391.

GARDNER, Rod (2010) "Question and answer sequences in Garrwa talk", *Australian Journal of Linguistics*, 30(4): 423-445.

GARDNER, Rod, and Ilana MUSHIN (2007) "Post-start-up overlap and disattentiveness in talk in a Garrwa Community", *Australian Review of Applied Linguistics* 30(3): 35.1–35.14.

GARDNER, Rod, and Ilana MUSHIN (2015) "Expanded transition spaces: the case of Garrwa", *Frontiers in Psychology* 3: 1-14.

GOODWIN, Charles (1981) *Conversational Organization. Interaction between Speakers and Hearers*. New York: Academic Press.

GOODWIN, Charles (2013) "The co-operative, transformative organization of human action and knowledge", *Journal of Pragmatics* 46 (1): 8-23.

GRZECH, Karolina (2021) "Using discourse markers to negotiate epistemic stance: A view from situated language use", *Journal of Pragmatics* 177: 208-223.

HADDINGTON, Pentti, Lorenza MONDADA, and Maurice NEVILE (eds.) (2013) *Interaction and Mobility. Language and the Body in Motion*. Berlin: De Gruyter.

HADDINGTON, Pentii, Tiina EILITTÄ, Antti KAMUNEN, Laura KOHONEN-AHO, Iira RAUTIAINEN, and Anna VATANEN (eds.) (2023) *Complexity of Interaction: Studies in Multimodal Conversation Analysis*, Cham: Palgrave Macmillan.

HAYANO, Kaoru (2013), "Question design in conversation", in SIDNELL Jack, and Tanja STIVERS (eds.), 395-414.

HAZEL, Spencer, Kristian MORTENSEN, and Gitte RASMUSSEN (2014) "Introduction: a body of resources – CA studies of social conduct", *Journal of Pragmatics* 65: 1-9.

HEPBURN, Alexa, and Galina B. BOLDEN (2013) "The Conversation Analytic Approach to Transcription". In SIDNELL Jack, and Tanja STIVERS (eds.), 57-76.

HERITAGE, John (2005) "Conversation Analysis and Institutional Talk." In K. L. FITCH K.L, and Robert E. SANDERS (eds.), *Handbook of Language and Social Interaction*. Pp. 103-147. Mahwah, NJ: Lawrence Erlbaum.

HERITAGE, John, and Max ATKINSON (1984) "Introduction". In HERITAGE, John, and J.M. ATKINSON (eds.), *Structures of social action: studies in Conversation Analysis*. Pp. 1-15. Cambridge: Cambridge University Press.

HOYMANN, Gertie (2010) "Questions and responses in ǂĀkhoe Hai||om", *Journal of Pragmatics* 42: 2726-2740.

JEFFERSON, Gail (2004) "Glossary of Transcript Symbols with an Introduction". In LERNER, Gene (ed.), *Conversation Analysis: Studies from the First Generation*. Pp. 2-23. Amsterdam and Philadelphia: John Benjamins.

KENDRICK, Kobin H., Penelope BROWN, Mark DINGEMANSE, Simeon FLOYD, Sonja GIPPER, Kaoru HAYANO, Elliott HOEY, Gertie HOYMANN, Elizabeth MANRIQUE, Giovanni ROSSI, and Stephen C. LEVINSON (2020) "Sequence Organization: A Universal Infrastructure for Social Action", *Journal of Pragmatics* 168: 119-138.

KITZINGER, Celia (2013) "Repair". In SIDNELL, Jack, and Tanja STIVERS (eds.), 229-256.

LEVINSON, Stephen C. (2010) "Questions and responses in Yélî Dnye, the Papuan language of Rossel Island", *Journal of Pragmatics* 42: 2741-2755.

LEVINSON, Stephen C. (2015) "Other-initiated repair in Yélî-Dnye: Seeing eye-to-eye in the Language of Rossel Island", *Open Linguistics* 1: 386-410.

MITCHELL, Alice, and Fiona M. JORDAN (2022) "Kinship, seniority, and rights to know in Datooga children's everyday interaction", *Journal of Pragmatics* 181: 49-61.

MONDADA, Lorenza (2009) "Emergent Focused Interactions in Public Places: A Systematic Analysis of the Multimodal Achievement of a Common Interactional Space", *Journal of Pragmatics* 41 (10): 1977-1997.

MONDADA, Lorenza (2019) "Contemporary issues in conversation analysis: Embodiment and materiality, multimodality and multisensoriality in social interaction", *Journal of Pragmatics* 145: 47-62.

MUSHIN, Ilana and Rod GARDNER, (2009) "Silence is talk: conversational silence in Australian Aborigenal talk-in-interaction, *Journal of Pragmatics*, 41(10): 2033-2052.

POMERANTZ, Anita, and John HERITAGE (2013) "Preference". In SIDNELL, Jack, and Tanja STIVERS (eds.), 210-228.

RAWLS, Anne W., and Michael LYNCH (2024) "Ethnography in ethnomethodology and conversation analysis: Both, neither, or something else together?", *Qualitative Research* 24(1): 116-144.

ROBINSON, Jeffrey (2006) "Managing Trouble Responsibility and Relationships During Conversational Repair", *Communication Monographs* 73 (June): 137-161.

ROSSI, Giovanni, Simeon FLOYD, and Nick J. ENFIELD (2020) "Recruitments and pragmatic typology". In FLOYD, Simeon, Giovanni ROSSI, and Nick J. ENFIELD (eds.), *Getting Others to Do Things: A Pragmatic Typology of Recruitments*. Pp. 1-23. Berlin: Language Science Press.

SACKS, Harvey (1992) *Lectures on Conversation*. Oxford: Blackwell.

SACKS, Harvey, Emanuel A. SCHEGLOFF, and Gail JEFFERSON (1974) "A simplest systematics for the organization of turn-taking for conversation", *Language* 50: 696-735.

SCHEGLOFF, Emanuel A. (2006) "Interaction: the infrastructure for social institutions, the natural ecological niche for language, and the arena inwhich culture is enacted". In ENFIELD, Nick J., and Stephen C. LEVINSON (eds.), *Roots of Human Sociality*. PP. 70-96. London: Routledge.

SCHEGLOFF, Emanuel A., and Harvey SACKS (1973), "Opening up Closings", *Semiotica* 8(4): 289-327.

SCHEGLOFF, Emanuel A., Harvey SACKS, and Gail JEFFERSON (1977) "The preference for self-correction in the organization of repair in conversation", *Language* 53: 361-382.

SIDNELL, Jack, and Tanya STIVERS (eds.) (2013) *The Handbook of Conversation Analysis*. Chichester: Blackwell.

STIVERS, Tanya (2013) "Sequence organization". In SIDNELL, Jack, and Tanja STIVERS (eds.), 191-209.

STIVERS, Tanya, Nick J. ENFIELD, and Stephen C. LEVINSON (eds.) (2010) "Question-Response Sequences in Conversation across Ten Languages", Special issue, *Journal of Pragmatics* 42(10): 2615-2860.

STIVERS, Tanya, Nick J. ENFIELD, Penelope BROWN, C. ENGLERT, Makoto HAYASHI, Trine HEINEMANN, Gertie HOYMANN, Federico ROSSANO, Jan Peter DE RUITER, Kyung-Eun YOON and Stephen C. LEVINSON (2009) "Universals and Cultural Variation in Turn-taking in Conversation", *PNAS* (*Proceedings of the National Academy of Sciences*) 106(26): 10587-92.

STREECK, Jürgen, Charles GOODWIN, and Curtis LEBARON (eds.) (2011) *Embodied Interaction: Language and Body in the Material World*. Cambridge: Cambridge University Press.

WARFIELD RAWLS, Anne, and Michale LYNCH (2024) "Ethnography in ethnomethodology and conversation analysis: Both, neither, or something else alltogether?" *Qualitative Research* 24(1): 116-144.

WHITEHEAD, Kevin A. (2011) "Some Uses of Head Nods in Third Position in Talk-in-Interaction", *Gesture* 11(2): 103-22.

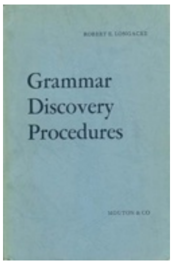
ZINKEN, Jörg, Giovanni ROSSI, and Reddy VASUDEVI (2020), "Doing more than expected: Thanking recognizes another's agency in providing assistance". In TALEGHANI-NIKAZM, Carmen, Emma BETZ, Peter GOLATO (eds.) *Mobilizing Others: Grammar and Lexis within Larger Activities.* Pp. 253-278. Amsterdam and Philadelphia: John Benjamins.

Language Documentation and Linguistic Fieldwork

A list of books, journals and websites[1]

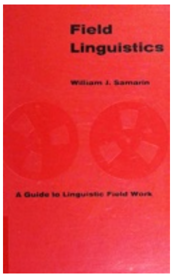
Compiled by Moreno Vergari

BOOKS



Grammar Discovery Procedures. A Field Manual
Longacre, Robert E.
Mouton & Co., 1964

Download at: https://archive.org/details/grammardiscovery0000long

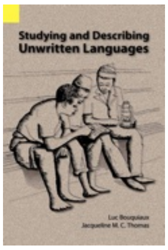


Field Linguistics. A guide to Linguistic Field Work
William J. Samarin
Holt, Rinehart and Winston, 1967

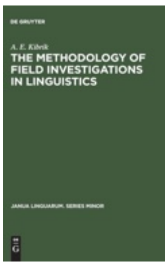
Download at:
https://archive.org/details/fieldlinguistics0000sama/page/n5/mode/2up

Enquête et description des langues à tradition orale
L. Bouquiaux, J. M.C. Thomas
SELAF, 1971

English version: see here below.



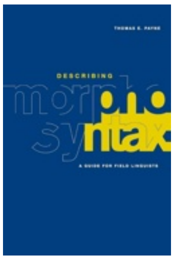
Studying and Describing Unwritten Languages
L. Bouquiaux, J. M.C. Thomas
SIL International, 1992, 2003



The methodology of field investigations in linguistics
A. E. Kibrik
De Gruyter, 1977

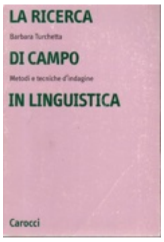
Download at: https://archive.org/details/methodologyoffie0000kibr

[1] Thanks to Peter Austin for suggesting some additions to this list.

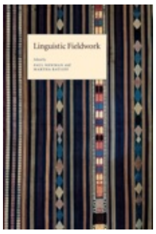


Describing Morphosyntax. A Guide for Field Linguists
T. E. Payne
Cambridge University Press, 1997

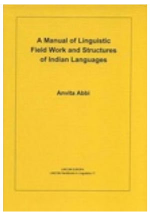
Introduction to Linguistic Field Methods
B. Vaux, J. Cooper
Lincom, 1998



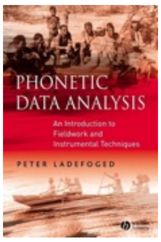
La ricerca di campo in linguistica. Metodi e tecniche di indagine
B. Turchetta
Carocci, 2000



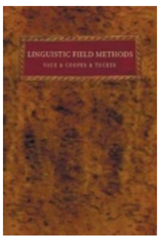
Linguistic Fieldwork
P. Newman, M. Ratliff
Cambridge University Press, 2001, 2008



A Manual of Linguistic Field Work and Indian Language Structures
A. Abbi
Lincom, 2001



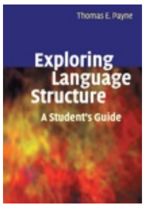
Phonetic Data Analysis. An introduction to Fieldwork and Instrumental Techniques
P. Ladefoged
Blackwell, 2003



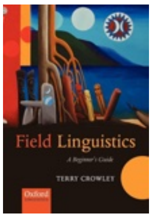
Linguistic Field Methods
B. Vaux, J. Cooper
Wipf & Stock, 2006

Essentials of Language Documentation
J. Gippert, N. P. Himmelmann, U. Mosel (eds.)
De Gruyter, 2006
Download at: https://linggwistiks.files.wordpress.com/2009/04/essentials-of-language-documentation.pdf

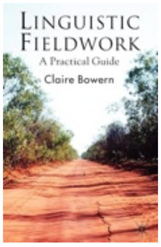
Catching Language. The standing challenge of grammar writing
F. K. Ameka, A. Dench, N. Evans (eds.)
De Gruyter, 2006



Exploring Language Structure. A Student's Guide
T. E. Payne
Cambridge University Press, 2006



Field Linguistics. A Beginner's Guide
T. Crowley
Oxford University Press, 2007



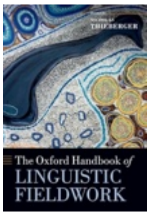
Linguistic Fieldwork. A Practical Guide
C. Bowern
Palgrave Macmillan, 2008

Linguistic Fieldwork. A Practical Guide
C. Bowern
Palgrave Macmillan, 2nd edition 2015

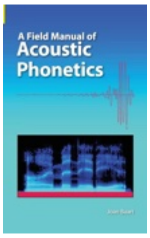
Language Documentation. Practice and values
L. A. Grenoble, N. L. Furbee (eds.)
John Benjamins, 2010

Handbook of Descriptive Linguistic Fieldwork
S. L. Chelliah, W. J. de Reuse
Springer, 2011

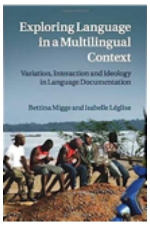
Linguistic Fieldwork
J. Sakel, D. L. Everett
Cambridge University Press, 2012



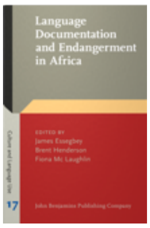
The Oxford Handbook of Linguistic Fieldwork
N. Thieberger (ed.)
Oxford University Press, 2012



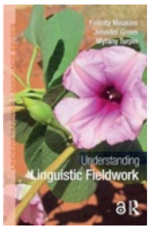
A Field Manual of Acoustic Phonetics
J. Baart
SIL International, 2012



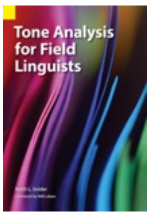
Exploring Language in a Multilingual Context.
Variation, Interaction and Ideology in Language Documentation
B. Migge, I. Léglise
Cambridge University Press, 2013



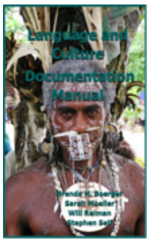
Language Documentation and Endangerment in Africa
J. Essegbey, B. Henderson, F. Mc Laughlin
John Benjamins, 2015



Understanding Linguistic Fieldwork
F. Meakins, J. Green, M. Turpin
Routledge, 2015

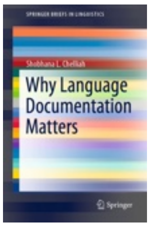


Tone Analysis for Field Linguists
K. L. Snider
SIL International, 2017

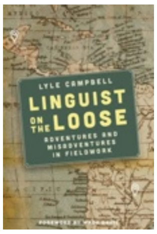


Language and Culture Documentation Manual
B. H. Boerger, S. R. Moeller, W. Reiman, and S. Self
Leappub, 2017

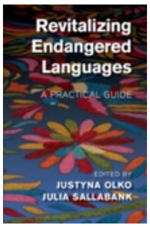
A Manual of Linguistic Fieldwork
F. Rozhanskiy
Tartu, 2021
Download at: https://copius.univie.ac.at/files/fieldwork_2021.pdf



Why Language Documentation Matters
S. L. Chelliah
Springer, 2021



Linguist on the Loose. Adventures and Misadventures in Fieldwork
L. Campbell
Edinburgh University Press, 2021



Revitalizing Endangered Languages. A Practical Guide
J. Olko, J. Sallabank (eds.)
Cambridge University Press, 2021
Download at: https://www.cambridge.org/core/books/revitalizing-endangered-languages/ADCBBA31190F259BA13525C769E92A9A

JOURNALS & WEBSITES

ENDANGERED LANGUAGES DOCUMENTATION PROGRAMME	https://www.eldp.net/
Language Documentation and Description	https://www.lddjournal.org/issues/
LANGUAGE DOCUMENTATION & CONSERVATION	https://nflrc.hawaii.edu/ldc/
SIL	https://www.sil.org/language-culture-documentation/language-documentation https://www.sil.org/resources/publications/lcdd

MAX-PLANCK-GESELLSCHAFT	https://www.eva.mpg.de/linguistics/past-research-resources/documentation-and-description/
The Language Archive	http://fieldmanuals.mpi.nl/
DOBES DOCUMENTATION OF ENDANGERED LANGUAGES	https://dobes.mpi.nl
ELP Endangered Languages Project	https://www.endangeredlanguages.com
elp EL Publishing	https://elpublishing.org https://elpublishing.org/documentation-theory-and-practice/
SOROSORO So the languages of the world may live on!	https://www.sorosoro.org/en/the-sorosoro-programme/
verba africana	http://verbafricana.org
rlexf Reference Lexicon of the languages of Africa	https://reflex.cnrs.fr/Africa/ see also: https://llacan.cnrs.fr/index_en.php
ELA Endangered Language Alliance	https://www.elalliance.org/about/how-we-work/language-documentation
LIVING LANGUAGES	https://www.livinglanguages.org.au/resources
PARADISEC	https://www.paradisec.org.au/resources/ https://www.paradisec.org.au/resources/downloads/
LDaCA Language Data Commons of Australia	https://www.ldaca.edu.au/about/principles/

MONDOFOTO/PHOTOWORLD

Some scenes from fieldwork activities

Fieldwork is like heart surgery:
you can learn it to do it
only by practicing on someone
Peter Ladefoged

Graziano Savà doing fieldwork among the Ongota (Ethiopia), August 2009.
Photos are taken from the video documentary by Robert Weijs, *Ongota, a dying language.*

Translating children's stories into Diyari. From left to right: Reg Warren, Peter Austin, Michelle Warren, Rene Warren, Port Augusta, South Australia, May 2024. Photo by Peter Austin

Michelle Warren facilitating discussion of how to say "Hello" in Diyari, Broken Hill, NSW, November 2023.
Photo by Peter Austin.

Michelle Warren interviews her grandmother Rene Warren about Diyari plant names, Killalpaninna, South Australia, November 2022. Photo courtesy Jan Scott.

Fieldwork in Kaaribossa (Eritrea), January 2011: Moreno Vergari and Ahmedsaad Mohammed Omer interview Mohammed Ali Ahmed. Photo by Roberta Vergari.

Roberta Vergari and Ahmedsaad Mohammed Omer interview Maryam Omar Ali in Buyya (Eritrea), January 2010. Photo by Moreno Vergari.

Traditional Paniya house. Sullimola, Tamil Nadu, India. December 2013. Photo by Cristina Muru.

Paniya husband and wife with Cristina Muru after the interview session. Sullimola, Tamil Nadu, India, December 2013. Photo courtesy Prakash - Centre for Tribal and Rural Development [CTRD] Trust, Ellamannala, Tamil Nadu, India.

Silvia Dal Negro’s fieldwork among the Walser community in Macugnaga, Italy, 2009. Third from the left: Roberto Marone of the local Walser association.

Experiment of online fieldwork with students of the Master's program in Applied Linguistics, with an interview with Jolanda Zertanna, an elderly Walser woman from Formazza, and her granddaughter Beatrice.
Free University of Bozen/Bolzano (Brixen branch), academic year 2021-2022.

DOCUMENTI

The International Decades of Indigenous Languages (2022-2032)

The United Nations General Assembly proclaimed the period of 2022-2032 as the International Decade of Indigenous Languages (IDIL2022-2032).

What is the IDIL 2022-2032

International Decades are important cooperation mechanisms dedicated to raising awareness on a particular topic or theme of global interest or concern and mobilizing different players for coordinated action around the world.

Following the 2016 Resolution 71/178 on the Rights of Indigenous Peoples, in 2019, the United Nations General Assembly adopted a resolution proclaiming the period of 2022-2032 the International Decade of Indigenous Languages, based on a recommendation by the Permanent Forum on Indigenous Issues. The proclamation of the IDIL2022-2032 is a key outcome of the 2019 International Year of Indigenous Languages (IYIL2019).

In 2016, the Permanent Forum stated that 40 per cent of the estimated 6,700 languages spoken around the world were in danger of disappearing. The fact that most of these are Indigenous languages puts the cultures and knowledge systems to which they belong at risk.

In addition to this, Indigenous peoples are often isolated both politically and socially in their countries of residence by virtue of geographical remoteness or historical, cultural, and linguistic inequalities. The COVID-19 pandemic has exacerbated the marginalization of Indigenous peoples around the world, having a negative impact on the world's linguistic diversity.

Indigenous peoples are not only leaders in protecting the environment, but their languages represent complex systems of knowledge and communication and should therefore be recognized as a strategic national resource for sustainable development, peacebuilding and reconciliation.

Indigenous languages also promote local cultures, customs and values which have endured for thousands of years. Indigenous languages add to the rich tapestry of global cultural diversity. Without them, the world would be a poorer place.

The IDIL2022-2032 will help promote and protect Indigenous languages and improve the lives of those who speak and sign them and will contribute to achieving the objectives set out in the United Nations Declaration on the Rights of Indigenous Peoples.

A Global Call for Action

The International Decade of Indigenous Languages offers us a unique opportunity to collaborate in policy development, to ensure continuity and coherence of actions and stimulate an intercultural dialogue in the true spirit of multi-stakeholder engagement, to contribute to making human rights a reality, and to take necessary measures in an interdisciplinary manner to support and strengthen Indigenous languages around the world. The scope of work envisaged during the International Decade is beyond the capacity of any single nation, country, stakeholder group, generation, scientific discipline, policy framework or set of actions. So, the International Decade presents a unique framework for convening a wide range of stakeholders collectively to align their efforts, accelerate development plans, make strategic investments, set research and legislative agendas, and launch concrete initiatives around common goals.

Last update: 9 January 2024

Source:
https://www.unesco.org/en/decades/indigenous-languages/about/idil2022-2032?hub=67103

Links: https://idil2022-2032.org
https://idil2022-2032.org/about-2022-2032/

See also the United Nations Declaration on the Rights of Indigenous Peoples:
https://www.un.org/development/desa/indigenouspeoples/wp-content/uploads/sites/19/2018/11/UNDRIP_E_web.pdf
For the Italian version:
https://www.ethnorema.it/wp-content/uploads/3-09-Documenti.pdf

I contributori possono inviare i loro articoli (in italiano, inglese, francese o tedesco) a: Ethnorêma – Viale Druso, 337/A – 39100 Bolzano – oppure a: *info@ethnorema.it.*
Per ulteriori informazioni: *www.ethnorema.it/rivista*

Ethnorêma adotta un sistema di referaggio a "doppio cieco".

Contributors are requested to submit their articles (in Italian, English, French or German) to: Ethnorêma – Viale Druso, 337/A – 39100 Bolzano – Italy, or to: *info@ethnorema.it.*
Additional information is available at *www.ethnorema.it/en/journal*

Ethnorêma uses a "double blind" peer-review process.