Doing Phonetics in the Rift Valley : Sound Systems of Maasai, Iraqw and Hadza

Ghio[1] A., Demolin[2] D., Karani[3] M., Meynadier[1] Y.

(1) Aix Marseille Univ, CNRS, LPL, Aix-en-Provence, France
(2) LPP (UMR7018, CNRS/Sorbonne Nouvelle), Paris, France
(3) Centre for Communication Studies, College of Humanities, Univ of Dar es Salaam, Tanzania

{alain.ghio,yohann.meynadier}@univ-amu.fr,didier.demolin@sorbonne-nouvelle.fr, karanim@udsm.ac.tz

Abstract

This article discusses the contribution of experimental techniques to the recording of phonetic data in the field. Only a small part of the phonological systems of African languages is described with precision. This is why it is important to collect empirical data in the form of sound, video and physiological recordings. This allows research questions such as patterns of variation to be addressed. Analytical methods show how to interpret data from physical principles and integrate them into appropriate models. The question of linguistic contact between different language families is also addressed. To achieve these general objectives, we present the way we design corpora, and the different ways of recording data with crucial technical considerations during fieldwork. Finally, we focus on 3 languages spoken in the Great African Rift Zone, which includes several linguistic areas belonging to the four major linguistic families of the continent. (1) Hadza is a click language with a very complex consonant system. (2) Iraqw is a Cushitic language with ejective consonants. (3) Maasai is a Nilotic language with implosive consonants and a very elaborate set of interjections, ideophones and animal calls that include sounds not described in the International Phonetic Alphabet.

Keywords: experimental phonetics, fieldwork, hadza, iraqw, maasai

1. The need for empirical data

More than 2000 languages are spoken on the African continent (Güldemann, 2018). Only a small part of them has been described in grammar form. Among those that have been described, the part of grammar devoted to phonetics and phonology is often less than 10% of the content of the full grammar (Maddieson, 2002). In addition, Maddieson indicated that "while syntactic patterns are documented with example sentences, often from natural discourse or texts, the phonetic facts are rarely if ever documented by the presentation of hard evidence". The phonological descriptions to which we have access are sometimes imprecise and there is often an ambiguity of symbols used. To illustrate the lack of data, we can compare the map of languages described in the PHOIBLE database (Moran, S., & McCloy, 2019) which contains 2186 distinct languages and the actual map of languages spoken in Tanzania (Figure 1). Although this database is the most extensive in the field, only 30 languages are proposed in PHOIBLE among the 125 ones spoken in Tanzania. Not only is the data reduced, but it also contains inaccuracies and even errors. For example, if you select the Iraqw Language (Glottocode: iraq1241), the Iraqw sound inventory proposed by the Stanford Phonology Archive (SPA) or by the UCLA Phonological Segment Inventory Database (UPSID) includes implosive consonants in this language as [ɓ] and [ɗ], which is false considering the detailed work by Mous (1993). Furthermore, still in the Iraqw Language, the consonant /q/ is most of the time described as a voiceless uvular plosive but the realization of this phoneme appears not as pulmonic but with an ejective mechanism, which would merit a phonological representation /q'/. In the same way, if you select the Maasai Language (Glottocode: masa1300), the Maasai sound inventory proposed by the Stanford Phonology Archive (SPA) or by the UCLA Phonological Segment Inventory Database (UPSID) includes a voiceless alveolar trill [r̥] whose voiceless feature is contested (Karani et al., 2023).

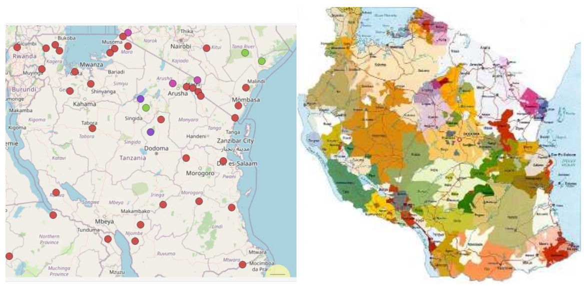

Figure 1: Languages spoken in Tanzania.
On the left: Inventory in PHOIBLE database (Moran, S., & McCloy, 2019).
On the right: the Atlasi ya lugha za Tanzania, (2009)

For all these reasons, it is important to collect empirical data in the form of sound, video and physiological recordings. This is the goal of the projects, "SYSORI" (Sound Systems of Rift Valley languages) and its extension "COSYSORI", funded by CNRS (see §.6 acknowledgements). The objective of this article is to provide some preliminary details.

The Fifth Workshop on Resources for African Indigenous Languages @LREC-COLING-2024 (RAIL), pages 1–9
25 May, 2024.

2. Elements of the methodology

2.1 Field linguistic or fieldwork at home

There are several ways to do field linguistics. If the grammar of the language, and in particular the phonological system, is precisely described, it is possible to use an ecological approach as recommended by Gasquet (2015). As we work on under documented languages, it is not possible for us to describe linguistic phenomena on the fly. We apply an approach based on the corpus.

We practice elicitation with speakers from a defined community, allowing the description of phonetic characteristics of the language studied. Field linguistics certainly requires specific preparation, knowledge of places, history, geography, adaptation to local culture, the use of interpreters, and contact with people. If we refer to Crowley's (2007) description, we practice "dirty feet" linguistics. Nothing to do with the recording of native speakers within universities, defined by (Crowley, 2007) as "fieldwork at home". In the field, we have to face sociolinguistic realities (social hierarchies, tensions, discrimination, differences according to age, sex, etc.). We must also find solutions linked to practical, logistical and technical challenges. We can mention, for example, sometimes it is very difficult to access isolated villages in the countryside. We also think about the absence of electricity or an intermittent electricity supply, thus requiring the use of batteries and solar panels. We can cite the difficulties of obtaining a low-noise environment with the nearby presence of children, livestock, etc. Such practice requires good field experience and enough preparation before embarking on fieldwork missions.

2.2 Ethical aspects

In our fieldwork, an interpreter, in our case a native speaker who is a trained linguist, clearly explains to volunteers what the goal of the research project is. Participants are informed about the way research will be conducted as well as their rights in the project including freedom to withdraw themselves from the project without giving reasons to do so. As our participants do not always know how to read and/or write, this information and the participant's agreement is audio or video recorded. The consent recordings are kept in a special storage location shared by project team members because they are non-anonymized data. All scientific data is then anonymised through the use of codes. Only a local investigator keeps private information concerning the speakers in order to contact them again in the future if necessary for more fieldwork. Only general anonymous information (age, sex, languages they speak, and the level of education) is accessible as methodological data. To compensate for the time spent, all participants receive a small cash allowance, as do the people who coordinate the target speakers.

2.3 How do we design corpus?

For us, the main corpus is oriented towards phonetic description. To obtain relevant and usable data, our method is to create an ad hoc corpus to answer a particular question. Most of the time, it is a list of words which includes the phonemes which we think have phonetic qualities under investigation. For instance, to observe the contrast between modal vs labialised consonants in Iraqw, we proposed a list of words with the modal expressions (konki, dakaát, anága kií/, daktani…) vs labialized (lakwanti, múk tlakw, án aga kwandeékw, tatlkwa…). To study the non-pulmonic consonants, we select words with ejective consonants in Iraqw or implosive consonants in Maasai. In this type of study, it is important to contrast these units with pulmonic ones in order to observe the typicality of non-pulmonic contrasted with pulmonic ones. In all the cases, the lists should be well designed and balanced in terms of occurrences (same amount of data by group if we study a contrast). If possible, the target phoneme should be placed at the initial, median or final position. The left and right contexts of the target expression should be managed. It can vary if we consider that the context can influence the realization of the target phoneme.The way to elicit the corpus can vary, depending on the capabilities of the speakers. Reading the list is the easiest way but it means that the chosen words should be written without ambiguity and above all, the participants should be able to read fluently. This is far from being the case in the field. An alternative is to practice a repetition: a referent speaker pronounces the target and the participant repeats it. We can also reinforce the elicitation by the proposition translated into another language mastered by the participant. For instance, in Tanzania where the majority of people speak Kiswahili, we would start by the word in Kiswahili and then the word in the studied language. The amount of data must be large enough to obtain statistically significant results and above all to have legitimacy to generalize the results. It is also important to record at least 10 speakers (5 male & 5 female) in order to neutralize a possible atypical participant. It can be important to mix also young, medium and old people but most of the time, it is not possible to fulfil all the ideal conditions in the field since each field site has its unique settings and requirements due to its geography, socio-political and cultural aspects. During the recording of data, it is also better if the speakers repeat the utterance or corpus several times in order to see if the phonetic realization is stable or not.

The study of prosodic features needs another type of item. If the studied language has tones where the variation of tone can change the meaning, it is possible to work with isolated minimal pairs: for instance in Maasai, áàdàmú ('he will remember me') vs áádámú ('I remember you'). This can be the first step to study the tonal system of the language. But for a substantial prosodic analysis, it is better to observe longer sequences; it can be sentences, texts or spontaneous speech. An valuable aspect of data collection is also to record spontaneous speech. It is important to collect stories and not only specific target data. The content can be a valuable element of heritage and further analysis of data to answer. In this situation, we try to record oral history using a video camera. The technical design is described further in the subsequent sections.

2.4 How do we record the corpus?

It is now possible to use applications, which can help with data recording. For instance, LIG-Aikuma is a mobile app for speech data collection and language documentation (Blachon et al., 2016). One particular capability is the "respeaking" mode where the operator can record first a referent speaker. After a short processing step, the experimenter can playback this reference for other speakers. This mode can avoid a fastidious repetition for the native referent speaker. It does not replace the physical presence of an interpreter or native speaker but it could allow us to record new data without the physical presence of the referent if we cannot do otherwise. This application has the advantage since it is portable and it can work on a mobile phone. Unfortunately, it does not work on a slightly more sophisticated station equipped with a laptop, a sound card and a good microphone. Indeed, in order to obtain good sound data, we prefer to use good electrostatic microphones connected to an external soundboard. For certain specific corpora, we also use specialized devices that allow us to record physiological data. We will describe this further.

2.5 Annotation and segmentation

Annotating a corpus consists of adding relevant information for its use. This consists of indicating orthographically the content of the oral data but also and above all (in phonetics) of precisely identifying the linguistic boundaries (segmentation). When the corpus is elicitated as we do, the transcription is trivial if the speaker pronounces correctly the list of items. But it is rarely perfect and we are faced with hesitation, errors, autocorrections… The second step is the phonetic segmentation where we need to set the contents and the boundaries of the spoken text. This process is time-consuming because, with no-documented languages, only a manual process is possible. We know that some projects propose some tools to automate this processing. We think for example the BULB project which aims at supporting the documentation of unwritten languages with the help of automatic speech and language processing, in particular automatic speech recognition (Adda et al., 2016). We also are aware of the CLD2025 project (https://anr.fr/Project-ANR-19-CE38-0015) whose goal is to facilitate the task of documenting endangered languages by leveraging the potential of computational methods. But for the moment, the use of these technologies is not yet possible due to the lack of data and knowledge of the languages studied.

2.6 Data sharing

In order to archive the data collected in the field, we asked ourselves the question of sharing data on cloud. In France, the Pangloss Collection is a digital library whose objective is to store and facilitate access to audio recordings in endangered languages of the world (Boyd et al., 2014). This platform is a potential final archive of the data we collect. For the moment, we mainly need a collaborative work platform in order to exchange raw data, enriched data, results or bibliography relating to our fieldwork between European and African partners. Our choice temporarily is the RESANA platform run by the French government that allows partners to perform the above-mentioned activities virtually.

3. Technical considerations

To collect our data, we have different technical designs depending on the purpose.

3.1 Simple audio design

It is very tempting to record speakers with a cell phone. This is the strategy used by the BULB project (Adda et al., 2016) using the application LIG-Aikuma (Blachon et al., 2016). This way is good for documenting a language. However, in a precise phonetic analysis, we need more controlled data.

Our minimal recording installation consists of a microphone connected to a sound card, which is connected to a laptop. We have 2 types of microphones: a professional head-worn condenser microphone (AKG C520) or a microphone on-stand (AKG C1000S). The advantage of the head-worn microphone is that it focuses on the speaker's speech and therefore considerably limits surrounding noise. It also maintains a constant distance which can allow measurements of variations in speech intensity. The downside of this equipment is that it must be attached to the speaker's head or ears, requiring touching the face or moving the hair, which is sometimes tricky. In addition, this type of poorly shielded microphone can be sensitive to electromagnetic interference leading to unwanted

electrical noise. This is not the case with a stand microphone, for example, AKG C 1000 S where its gold sputtered capsule housing makes the microphone extremely rugged against humidity or adverse condition. The position of the microphone is important. In order to avoid "pop" noises, which occur on plosive or fricative consonants, it is preferable to shift the microphone away from the speaker's axis and direct it towards the mouth, which forms an angle of approximately 45°.
The microphones that we use are condenser ones. They all need a phantom power because of condenser technologies. This is the first reason to use a sound card that directly powers the microphones via the cable. The second reason to use an external sound card is that the signal-to-noise ratio is better than recording directly with a laptop. The models of soundcard change quickly but we can give the models used at the moment: Focusrite Scarlet 2i2, Solid Static Logic SSL2, RME Fireface UCX.
It is also possible to use a portable wav/mp3 recorder, for example the famous Zoom H4n. We can use this device in a standalone mode because it integrates microphones, amplifiers, digitizer and storage. It is also possible to connect external microphones like the ones that we described above. This device can power condenser microphones. The drawback of this equipment is that it is not easy to control the recording level, especially if the device is set near the speaker and far from the operator. A possible incorrect level setting will only be discovered at the end of the recordings, which is very regrettable. The second drawback is that the recorded files are named automatically on the local storage and it is sometimes tricky to recover who and what was recorded on the labelled in the following format: File0035, File0078 etc.

3.2 Video Imaging design

As we mentioned in § 2.3, it is also important to record people telling stories. These recordings are rich data on heritage values. In this case, a video recording is preferred as a simple sound track.
The first simple way is to capture the images with a mobile phone or with a video camera where the sound is recorded with the integrated microphone. The problem with that is that most of the time, the camera is far from the speakers, the sound is imprecise, and the level of noise is high, especially if the scene is recorded outside.
In order to obtain good sound data, we use a setting with 2 external microphones. The main one is a Sennheiser MKH 416-P48, a shotgun interference tube microphone designed for film, radio, and television, especially for outside applications. This microphone has excellent directivity and a good sensitivity (25 mV/Pa). In order to reduce external noise, we equip it with a rigid windscreen MZW60-1 and we add above a windmuff MZH60-1. This microphone is set on a stand and the operator can manually orient it in order to target more precisely the speaker during the interview. A second microphone (AKG C1000S) can be connected as another sound source. We generally placed this transducer close to the interviewer (Figure 2). These two condensers microphones are connected to a Zoom H4n recorder which (1) delivers power to the microphones, (2) adjusts the recording level, (3) records the soundtracks in standalone, and (4) delivers good sound signals to the video camera. Indeed, we connect the Zoom H4n output to the external sound input of the video camera. It is important to control finally the sound quality by using a headphone connected to the video camera. Thus, the operator can orient the shotgun microphone and adjust the recording levels. At the end, we can obtain a film where a good soundtrack is directly synchronized with the images without an important post-processing. A copy of the soundtracks is also available on the Zoom H4n recorder.

Figure 2: Video installation for documenting languages (Karatu-Mbulumbulu-Lositete, Tanzania, Maasai Speakers, 2023)

3.3 Physiological data

The most original aspect of our project is the recording of physiological data synchronized to speech signals. These data are necessary to understand how speech sounds are produced and what are the best gestures or features necessary to describe these sounds (Demolin, 2011). A second important aspect is quantification of data which allows researchers to address the fundamental issue of variation of the studied phenomena. In the case of physiological data, we can access the roots of the variations, and not only the surface variations of the speech sound. The use of physiological data is sometimes essential to describe precisely speech sounds.
For example, the study of phonatory characteristics is facilitated by the use of electroglottography, which allows selective and direct observations, unlike the study of the speech signal, which is the result of very complex phonatory and articulatory convolutions (Figure 3). Likewise, the acoustic study of nasality is always a challenge while the physiological

mechanism of velum opening/closing remains a rather simple operation if we have a means to observe it. Measuring the nasal airflow can be a well-adapted solution for that (Figure 4). Finally, the mode of producing some obstruent consonants as plosives (egressive) or implosives (ingressive) is not so easy to detect in the speech signal. The measure of oral airflow or intraoral pressure during speech production can give quantified data about this phenomenon (Figure 5). Intraoral pressure in speech is an important physiological parameter because it highlights some complex mechanisms.

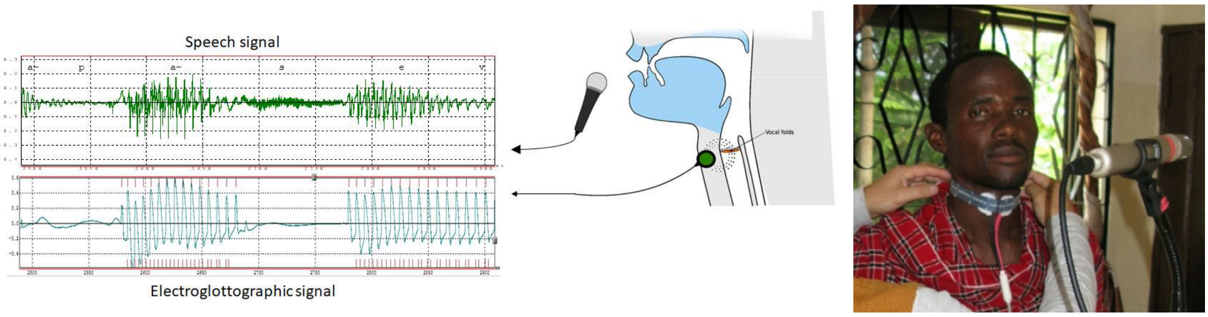

Figure 3: Selective observation of phonatory mechanisms through the electroglottography (Mto Wa Mbu, Tanzania, Maasai Speaker, 2023)

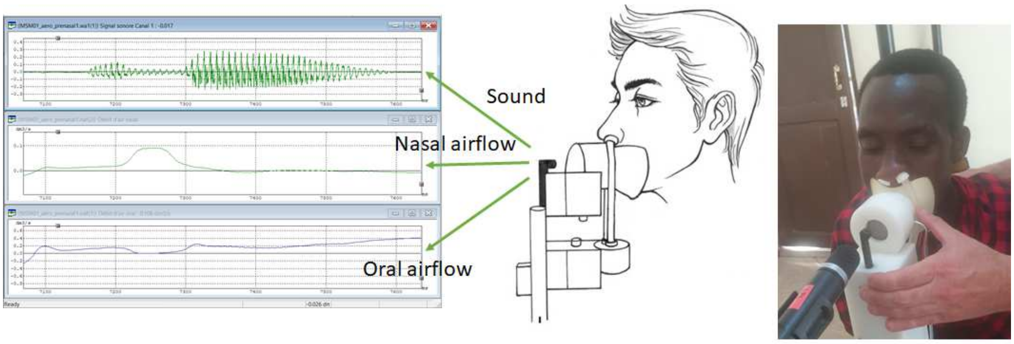

Figure 4: Measuring nasal airflow for selective observation of nasality mechanisms. Speech signal, nasal airflow and oral airflow for the Maasai word “Embaoi” (‘timber’) (Mto Wa Mbu, Tanzania, Maasai Speaker, 2023)

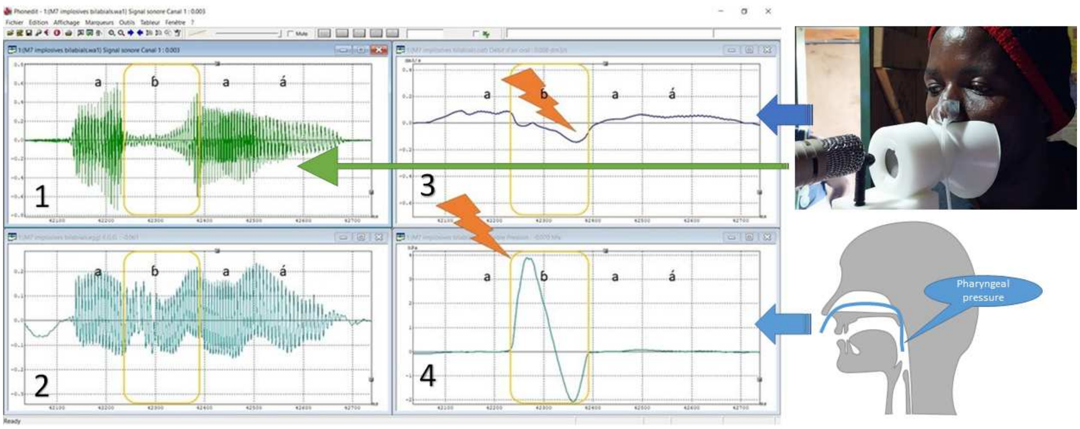

Figure 5: Multiparametric data for selective observation of implosives consonants with a Maasai speaker producing the word “aɓaá” (“to crack”). Speech signal (1), EGG (2), oral airflow (3), pharyngeal pressure (4) (Arusha, Tanzania, Maasai Speaker, 2022-2023)

During the production of an implosive consonant /ɓ/, we can see on curve n°3 a stop of the air output through the mouth. This is the occlusion mechanism. The pressure in the oral cavity increases (curve n°4) because air continues to be pushed by the lungs. Here we observe the classic process of producing a plosive. In the case of an implosive, there is an enlargement of the oral cavity and a lowering of the larynx. As the oral cavity is closed, thanks to the lips, the pressure drops suddenly and becomes negative (curve n° 4). When the lips open again on the vowel, this depression is cancelled when air enters the oral cavity.

Intraoral pressure is also interesting for the production of other consonants as, for example, ejectives in Iraqw (Demolin, Ghio et al., 2021). Getting pharyngeal pressure is helpful when we want to measure intraoral pressure regardless of the place of articulation during an oral occlusion. This involves inserting a catheter into the nasal cavity until reaching the cavum (Figure 5 bottom right). Without a medical doctor, we cannot perform this procedure ourselves. To respect this ethical constraint, we ask participants to do the insertion themselves. This operation is delicate and many participants refuse it, especially women. We always explain clearly why and how to insert the tube into the nose and before they do it we demonstrate how to do it ourselves. If speaker are not ready to do it we respect their decision. Most of the male speakers were ready to try the tube even though the insertion process sometimes failed. There is a significantly less invasive alternative for measuring intraoral pressure via the "airway interrupted method" proposed by Smitheran and Hixon (1981). The principle is to place a pressure probe in the oral cavity passing through the lips. The tube must be short and has to stop just behind the lips in order not to disturb articulation. The pressure can be measured during labial occlusions, but not for other occlusions articulated more posteriorly, which is a compromise in the field. More generally, this type of multiparametric observation help us to understand the mechanisms of speech production beyond the simple acoustic signal. Technically, we collect these data with the EVA2 workstation, developed by LPL-SQLab, Aix-en-Provence, France (Ghio et al., 2004). This equipment allows synchronized measurements of aerodynamic, electroglotto-graphic (EGG) and acoustic data. Two airflow channels and two pressure channels are available to measure oral, nasal airflow and oral pressure if necessary. The EGG signal is provided by an EG2-PCX2 model from Glottal Entreprise.

3.4 Power supply in the field

One of the major problems in using sophisticated equipment for fieldwork is the sometimes non-existent or intermittent electricity supply. This can become a real problem if power cut is frequent and it causes recording devices to malfunction (Figure 6). The solution we adopted is to work permanently with a 12V battery connected to a VICTRON inverter Phoenix 12|500 which provides a stable pure sinus electrical power supply during the recording session. We emphasize the need to use a pure sine inverter (and not a quasi-sinus inverter) because a poorly stabilized power supply can pollute the recorded signals by adding noise pulses linked to electric current chopping (Figure 7).

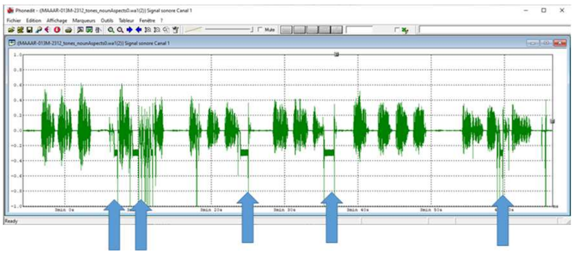

Figure 6: Effects of power supply microcuts on the sound wave (Arusha, Tanzania, 2023)

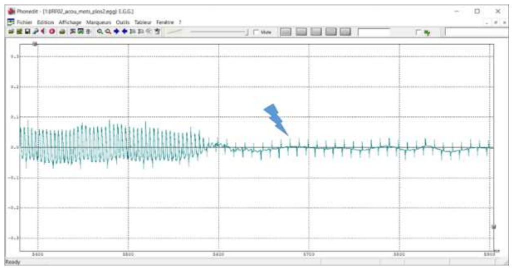

Figure 7 : Effects of a non-pure sinus inverter on the EGG signal (Kwermusl, Tanzania, 2020)

The next important question is to have a solution for charging the batteries. We have sometimes adopted the solution of portable solar panels connected to a charge controller. It is a solution that makes us completely autonomous, which is interesting in isolated areas. The disadvantage is the weight and bulk of the panels when travelling. The other solution is to have several batteries, to charge them at night or during breaks and then use them during recording sessions. Please note that it is forbidden to transport some batteries on passenger aircraft, therefore, it is necessary to find a solution with local providers.

4. Some sound systems of African Rift Valley languages

4.1 Why focus attention on the African Rift?

The Great African Rift includes several linguistic areas belonging to the four major language families of the continent: Afro-Asiatic, Niger-Kongo, Nilo-Saharan and Khoesan (Kießling et al., 2007). The linguistic diversity of this area is the result of migrations and contacts, sometimes very ancient, between populations of hunter-gatherers, pastoralists and farmers. The comparison between languages suggests several migratory phases and contacts that have repeatedly modified the linguistic landscape. The wide geographic range of some language families, such as Cushitic, which extends from Ethiopia to Kenya and Tanzania, is indicative of these ancient population movements. Some languages like Sandawe and Hadza have similar sounds, clicks, to those of the Khoesan family found in Botswana and Namibia.

The sound systems of the Rift languages have particular sound types, some of which are not

common in the world's languages. The case of non-pulmonic consonants such as ejectives, implosives, and clicks is particularly notable. Are these consonants very ancient remnants of elements that are reflected in present-day languages? Or are they the product of mechanisms of innovation and the complexification of sounds and sound systems? What is the link between the non-pulmonic consonants and the physiological mechanisms of swallowing or coughing? Is there a link between clicks and ejective consonants?
The answers to these questions require an accurate observation and description of the phoneme production mechanisms of these languages. In our project, the focus is on several languages from different linguistic families in the area (Figure 8).

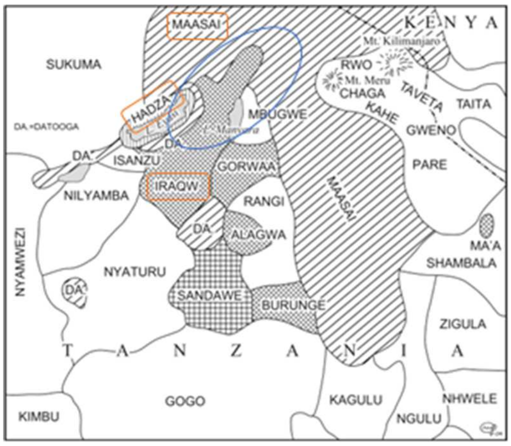

Figure 8: The languages of the Tanzanian Rift Valley Area (Source: Kießling et al., 2007)

4.2 Hadza

Hadza is a language spoken along the shores of Lake Eyasi in Tanzania by around 1,000 people (Figure 8). Traditionally, Hadzabe are full-time hunter-gatherers. Nowadays, most of them are bilingual in Swahili but Hadza language transmission to younger generation is still robust. However, there is no established Standard orthography (http://glottopedia.org).
Greenberg classified Hadza as Khoisan due to its use of click consonants but the Hadza is now considered as an isolate. The debate is delicate because if Hadza is linked to the Khoisan family, we can consider it as a remainder of the extended settlements of the Khoisan people in central Africa. If, on the other hand, as other scholars think, the Khoisan affiliation of the Hadza is not proven, we would admit that other linguistic families had existed in southern Africa, probably in contact with Khoisan, and that they finally disappeared, leaving behind only Hadza as a witness (Philipson, 2017). The precise description of the Hadza Sound System is therefore an important issue.

The consonantal system of Hadza is one of the most complex in the world, with around 60 consonants (Table 1). The inventory is controversial (Miller, 2008; Sands, 2013). It is sometimes difficult to know whether it is actually a distinct phoneme or an allophonic variation of another one.
We conducted a field mission in 2020 based in Mwangeza, a town located in the south of Lake Eyasi (https://mapcarta.com/34359060). This town is not inside the Hadza area and our speakers had to travel to this town for data recording. We recorded 4 female and 4 male speakers. Our corpus includes data on clicks, ejectives and prenasals features. We have also a spontaneous speech for a single speaker telling a story about hunting. The results on clicks are available in Demolin, Harvey et al., (2021).

	Consonants	Labial	Dental	Alveolar	Lateral	Post-alveolar	Velar	Labialized velar	Radical
Clicks	Oral clicks		c ch /ǀ ǀʰ/	q qh /! !ʰ/	x xh /ǁ ǁʰ/				
	Nasal clicks	(mcw) /ʘʷ/	nc /ǀ̃/	nq /!̃/	nx /ǁ̃/				
	Glottalized nasal clicks		cc /ǀˀ/	qq /!ˀ/	xx /ǁˀ/				
Stops	Oral stops	b p (ph) /b p pʰ/		d t (th) /d t tʰ/			g k (kh) /g k kʰ/	gw kw (khw) /gʷ kʷ kʷʰ/	' /ʔ/
	Prenasal stops	(mb mp) /mb mpʰ/		(nd nt) /nd ntʰ/			(ng nk) /ŋg ŋkʰ/	(ngw) /ŋgʷ/	
	Nasal stops	m /m/		n /n/		(ny) /ɲ/	(ng') /ŋ/	(ng'w) /ŋʷ/	
	Ejectives	(bb) /p'/		zz /ʦ'/	dl /ʧ'/	jj /ʈʂ'/	gg /kx'/	ggw /kxʷ'/	
Affricates	Oral affricates			z ts tsh /ʣ ʦ ʦʰ/	tl tlh /ʧ ʧʰ/	j tc tch /ɖʐ ʈʂ ʈʂʰ/			
	Prenasal affricates			(nz ns) /nʣ nʦʰ/		(nj) /nɖʐ/			
Fricatives	Fricatives	f /f/		s /s/	sl /ɬ/	sh /ʂ/			(hh) /ʜ/
Approximants	Sonorants				l /ɾ/	y /j/		w /w/	h /ɦ/

Table 1: Hadza consonants (adapted from Miller, 2008)

4.3 Iraqw

Iraqw is a South Cushitic language spoken in the Manyara and Arusha regions of Tanzania by more than half a million speakers (Kießling et al., 2007). Iraqw people are mainly agriculturalists. Iraqw is one of the southernmost Cushitic languages in the Afro-Asian phylum (Figure 8).
The phonetic nature of the Iraqw sound system (Table 2) is (1) a series of ejective consonants, (2) a long set of unvoiced fricative consonants, (3) a contrast between modal and labialized consonants.

	Labial	Alveolar	Lateral	Palatal	Velar	Uvular	Pharyngeal	Glottal
Voiced Plosive	b	d		ɟ <j>	g gʷ			
Voiceless Plosive	p	t		c <ch>	k kʷ			ʔ ⟨'⟩
Ejective Stop						q' qʷ'		
Ejective Affricate		ts'	tɬ' ⟨tl⟩					1
Voiced Fricative							ʕ ⟨/⟩	
Voiceless Fricative	f	s	ɬ ⟨hl⟩	ʃ ç ⟨sh⟩	x xʷ	2	ħ ⟨hh⟩	h
Nasal	m	n		ɲ ⟨ny⟩	ŋ ŋʷ			
Liquids		r	l					
Approximants	w			j ⟨y⟩	3			

Table 2: Iraqw consonants (adapted from Mous, 1993)

The ejective status of /q'/ is under debate (Demolin, Ghio et al., 2021). This question relating to ejectives is interesting because ejectives are also in the Hadza sound system (Table 1) as well as among the Sandawe, another click language spoken in the south of the area (Figure 8). Knowing that / tɬ'/ is only found in 1% of the world languages (Phoible, 2019), a presence in 3 geographically close languages is probably due to a contact effect which is yet to be studied.
We conducted a first field mission in 2020 based in Kwermusl, a town located in the south of Mbulu district (https://mapcarta.com/N5033497899). This village is the heartland of the Iraqw area and we recorded 5 female and 5 male speakers. Our corpus includes data on ejectives, pharyngeal, glottal articulation, labialized vs modal contrast, fricatives, and vowels. We also had a single speaker reading the translation of the story in Iraqw "the wind and the sun". We added two male speakers in Arusha in June 2022. We completed the corpus with three male and three female speakers recorded in Mto Wa Mbu (https://mapcarta.com/34353418) at the border between the Iraqw area and Maasai land in December 2022. We have some video data to study the difference between labial movements involved in the production of consonants /ŋʷ, kʷ, gʷ, qʷ´, xʷ/ and compare them with the gestures of the bilabial nasal [m] and the labiovelar approximant [w] (Ghio et al., 2021).

4.4 Maasai

Maa is a Nilotic language spoken in Southern Kenya and Northern Tanzania by 1.5 million Maasai people. Maasai steppe covers a large area in both Kenya and Tanzania and several dialects can be distinguished: Samburu, Ilchamus, Ilkeekoɲokie, Purko, Ilwasiŋgiʃu, Arusa, Kisongo, Parakuyo. Traditionaly, Maasai are pastoralists but some sections are also farmers, especially Arusa and recently during the fieldwork in December 2023, we noticed that Kisongo too in Monduli District have started subsistence farming. In its phonological system (Table 3), Maa has a complete set of implosives, a complete system of vowels [+/- ATR], tones, and a fortis/lenis contrast for glides /j/ vs /j:/, /w/ vs /w:/, and also for rhotic /ɾ/ vs /r/. There is also a very elaborate set of interjections, ideophones and animal calls that include sounds not described in the International Phonetic Alphabet (Andrason et al., 2023; Andrason et al., 2021; Karani et al., 2022). The implosive mechanism is not easy to observe and the use of aerophonometry, described in §3.3, is necessary.

		bilabial	alveolar	postalveolar	palatal	labiovelar	velar	glottal
plosive	explosive	[p] [b]	[t] [d]				[k] [g]	ʔ
	implosive	[ɓ]	[ɗ]				[ɠ]	
affricate	explosive			[t͡ʃ] [ɟ]				
	implosive			[ʄ]				
fricative			[s]	[ʃ]				[h]
nasal		[m]	[n]		[ɲ]		[ŋ]	
rhotic	tap		[ɾ]					
	trill		[r]					
lateral			[l]					
approximant /glide	lenis				[j]	[w]		
	fortis				[j:]	[w:]		

Table 3: Maasai consonants (Karani et al., 2023)

We conducted 3 field missions in 2022-2023. The first location was Ilkurot village around Arusha (https://mapcarta.com/N10836698430) where the Arusa dialect is spoken. We recorded 14 male and 8 female speakers for phonetic purposes (implosives, vowels, approximants, tones, ideophones…). Moreover, we video-recorded twice additional speakers telling stories. The second location was Esilalei and Selela villages where Kisongo dialect is spoken (https://mapcarta.com/N946585023). We recorded 8 male and 4 female speakers for phonetic purposes. Likewise, we video-recorded additional speakers telling stories in the villages. In order to study the contact between Maasai and Iraqw people, we conducted a mission in Lositete (https://mapcarta.com/N7512233367) where we recorded video clips of 5 speakers telling stories with the specificity of being Arusa Maasai surrounded by Kisongo Maasai and Iraqw. Some preliminary results on implosives have been shared by Demolin et al. (2022).

5. Conclusion

The necessity of empirical data to quantify observed phenomena is part of the scientific approach. This is the case in linguistics particularly in African Linguistics. The examples given before show that phoneticians have to obtain sets of quantitative data in order to understand patterns of variation. The Rift Valley in Tanzania is a perfect area to observe various and original linguistic features. With the experience we gain in the missions doing fieldwork attests that 'dirty feet' fieldwork is not a walk in the park but certainly feasible and it can be very successful if researcher prepare well. Since some linguistic data are rare and sometimes difficult to collect, data sharing is an important aspect that linguists need to take seriously during data collection and after fieldwork. Our fieldwork missions have been successful to a large extent because one of the collaborators is a trained linguist who is a native speaker of one of the languages under investigation. Hence, we recommend working with local researchers to maximize the chances of succeeding in fieldwork missions.

6. Acknowledgments

This study was supported by grant CNRS "SYSORI" (Dispositif de Soutien aux Collaborations avec l'Afrique subsaharienne , 2021) and by grant CNRS "COSYSORI" (Africa Visiting Fellowships Programme, 2023). We thank the CEP (https://cep.lpl-aix.fr/) for the loan of equipment used in the field. We thank Andrew Harvey and Richard Griscom for their participation in the recordings with the Hadza. We thank Maarten Mous for his participation in the recordings with Iraqw. We thank L Jalby, from OZO company (https://ozo-electric.com/en/) for their advice and products relating to the supply of electricity using solar equipment and batteries.

7. Bibliographical References

Adda, G., Stüker, S., Adda-Decker, M., Ambouroue, O., Besacier, L., Blachon, D., Bonneau-Maynard, H., Godard, P., Hamlaoui, F., Idiatov, D., Kouarata, G.-N., Lamel, L., Makasso, E.-M., Rialland, A., Van De Velde, M., Yvon, F., & Zerbian, S. (2016). Breaking the Unwritten Language Barrier : The BULB Project. *Procedia Computer Science*, 81, 8-14. https://doi.org/10.1016/j.procs.2016.04.023

Andrason, A. & Karani, M. (2023). Emotive interjections in Arusa Maasai. Italian Journal of Linguistics, 35(2), 75-118. https://doi.org/10.26346/1120-2726-212

Andrason, A. and Karani, M. (2021). Conative calls to animals: From Arusa Maasai to a cross-linguistic prototype. *Łódź Papers in Pragmatics*, 17(1), 3–40. https://doi.org/10.1515/lpp-2021-0002

Blachon, D., Gauthier, E., Besacier, L., Kouarata, G.-N., Adda-Decker, M., & Rialland, A. (2016). Parallel Speech Collection for Under-resourced Language Studies Using the Lig-Aikuma Mobile Device App. *Procedia Computer Science*, 81, 61-66. https://doi.org/10.1016/j.procs.2016.04.030

Boyd, M., Mazaudon, M. , Michaud, A., Guillaume, S., François, A, Adamou, E (2014). Documenting and Researching Endangered Languages: The Pangloss Collection. *Language Documentation & Conservation*. 8: 119-135.

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

Demolin, D. (2011). Aerodynamic techniques for phonetic fieldwork. *ICPhS International Congress of Phonetic Sciences*, Aug 2011, Hong Kong, China. ⟨hal-00646103⟩

Demolin, D., Ghio, A., Mous, M. (2021). A phonetic study of Iraqw ejectives consonants. *10th World Congress of African Linguistics (WOCAL)*, Leiden, Netherlands. ⟨hal-03156237⟩

Demolin, D., Harvey, A., Griscom, R., & Ghio, A. (2021). Acoustic features of Hadza clicks. *The Journal of the Acoustical Society of America*, *150*(4_Supplement), A68-A68. https://doi.org/10.1121/10.0007649

Demolin, D., Ghio, A., Karani, M. (2022). Acoustic, aerodynamic and articulatory features of Maasai implosives. *Colloquium on African Languages and linguistics*, Leiden, Netherlands ⟨hal-03767661⟩

Gasquet-Cyrus, M. (2015). I come and go between fieldwork: Thoughts on Fieldwork in Sociolinguistic Theorization. *Langage et société*, 154, 17-32. https://doi.org/10.3917/ls.154.0017

Ghio, A., & Teston, B. (2004). Evaluation of the acoustic and aerodynamic constraints of a pneumotachograph for speech and voice studies. *International Conference on Voice Physiology and Biomechanics*, Univ. Méditerranée, 2004, Marseille, France. pp.55-58. ⟨hal-00142982⟩

Ghio, A., Mous, M., & Demolin, D. (2021). Labialized consonants in Iraqw. 10th World Congress of African Linguistics. ⟨hal-03156276⟩

Güldemann, T. (Éd.). (2018). The Languages and Linguistics of Africa. De Gruyter. https://doi.org/10.1515/9783110421668

Karani, M., & Andrason, A. (2022). Ideophones in Arusa Maasai : Syntax, morphology, and phonetics. *Open Linguistics*, *8*(1), 440-458. https://doi.org/10.1515/opli-2022-0220

Karani M., Andrason A. (2023), A Living Grammar Sketch of Arusa, *Living Tongues Institute for Endangered Languages* https://doi.org/10.5281/ZENODO.10389820

Kießling, R., Mous, M., & Nurse, D. (2007). The Tanzanian Rift Valley area. In B. Heine & D. Nurse (Éds.), A Linguistic Geography of Africa (p. 186-227). Cambridge University Press. https://doi.org/10.1017/CBO9780511486272.007

Maddieson, I. (2002). Phonetics in the Field. *Annual Meeting of the Berkeley Linguistics Society*, 28(1), 411. https://doi.org/10.3765/bls.v28i1.3855

Miller, K. (2008). Hadza grammar notes. *Symposium on Khoisan Languages and Linguistic: In Memory of Professor Anthony Traill*, Riezlern.

Mous, M. (1993). A Grammar of Iraqw. Helmut Buske Verlag.

Phillippson, G. (2017). Langues et histoire dans le Rift. In *Le Rift est-africain* (IRD Éditions, p. 367-376). https://books.openedition.org/irdeditions/1790

Sands, B. (2013). Phonetics and Phonology: Hadza. In R. Vössen (Ed.). The Khoesan languages. London. Routledge.

Smitheran, J. R., & Hixon, T. J. (1981). A clinical method for estimating laryngeal airway resistance during vowel production. *The Journal of Speech and Hearing Disorders*, *46*(2), 138-146. https://doi.org/10.1044/jshd.4602.138

8. Language Resource References

Atlasi ya lugha za Tanzania. (2009). Mradi wa Lugha za Tanzania, Chuo Kikuu cha Dar es Salaam. ISBN: 978-9987-691-26-5

PHOIBLE 2.0 as CLDF dataset (v2.0), Moran, S., & McCloy,(2019). https://doi.org/10.5281/ZENODO.2593234