Transcribing (Multilingual) Voices: From Fieldwork to Publication

Erika Kalocsányiová and Malika Shatnawi

Corresponding author: e.kalocsanyiova@gre.ac.uk

Abstract:

Several authors have argued in the past that transcribing is a political act: it involves judgements about how to represent *voice* in writing, *what level of detail to choose*, in *which* language(s) and *for whom*. This chapter provides an in-depth account of the processes and politics of multilingual transcribing. In reflexively analysing different transcript formats, it casts light on the complexities, challenges and opportunities that emerge in making collaborative decisions about the translation of speech to a written medium, as well as into other languages. The transcripts for the analysis have been pulled from a larger qualitative study on forced migrants' linguistic integration. Special focus is thus given to transcripts that capture the sometimes unorthodox resources and mixed language practices of migrants, and to their reception in an interdisciplinary framework.

Transcribing (Multilingual) Voices: From Fieldwork to Publication

The old adage about 'sink or swim' often reflects early career researchers' experiences of transcribing (Bird 2005). It certainly demands, among others, a lot of time spent immersed in someone else's emotions and stories. In Tilley's observation (2003), this involvement and connection to the speakers contrasts sharply with views of transcription as technical, objective or value free. Transcribing involves judgements about *what level of detail to choose*, how to *represent voice*, in *which language(s)* and *for whom*. When undertaking multilingual research, 'the decisions inherent in interpretation and written representation – from the most intriguing (was that literal or ironic?) to the most mundane (colon or no colon?) – can quickly multiply' (Vakser 2017:233).

This chapter provides an in-depth account of the processes and politics of multilingual transcribing. The transcripts for the analysis have been pulled from a larger qualitative study on forced migrants' linguistic integration in Luxembourg. They broadly fall into two categories: records of multilingual classroom interactions and interview data. The research, which was carried out by Kalocsányiová at the University of Luxembourg, examined the language learning trajectories and communicative resourcefulness of newly-arrived migrants from war-torn countries in the Middle East (Kalocsányiová 2019). Much of this research depended on third parties to document the participants' communicative resources and emerging mixed language practices. The second author (Shatnawi) was one of the language experts brought onto the project to assist with these tasks. As an external researcher, she was involved in the transcription, analysis and translation of ethnographic interviews and classroom interactions. She also offered alternative interpretations of the data and suggested new lines of enquiry.

One of the main rationales for researching multilingually is to achieve a more democratic and inclusive research praxis, or as Resch and Enzenhofer (2018: 133) put it, 'to make diverse – and often vulnerable – groups and their perceptions, experiences, and needs visible and to include their experiences into the discourse'. This chapter addresses a key methodological challenge of this process: transcription. It discusses the role of collaborative transcribing in transforming (multilingual) speech to a written medium, as well as into other languages. It also examines representational choices for interview excerpts with adult (migrant) speakers of English.

The Politics of Transcribing and Translating Transcripts

Several authors have argued in the past that transcription, that is, the process of representing spoken language in an (ortho)graphic medium is a political act (Bucholtz, 2000; Green *et al*., 1997; Jaffe, 2006; Ochs, 1979; Roberts, 1997; Ross, 2010). Ochs (1979) was perhaps the first researcher to address the politics of transcription explicitly. Drawing on examples from her work on language development in children, she exposed the theoretical and cultural underpinnings of transcription. She also showed that transcripts can (mis)attribute social and interactional roles to the speakers represented. Nearly 20 years later, Roberts (1997) and Green *et al*. (1997) refuted altogether the idea of neutral transcripts and re-defined transcribing as a situated act which reflects theories and disciplinary conventions as much as transcribers' experiences, biases and beliefs about language: 'as transcribers fix the fleeting moment of words as marks on the page, they call up the social roles and relations constituted in language and rely on their own social evaluations of speech in deciding how to write' (Roberts, 1997: 168). In her contribution to the topic, Roberts (1997) focused, among other things, on the representation of learner varieties of English through both standard orthography and respelling (also called 'eye dialect'). Her main arguments are highly relevant to our discussion because each person involved in the study spoke English and/or French (the main languages of data collection and research dissemination) as additional languages. We have transformed these arguments into a series of questions to guide our work: How can study participants, and especially forced migrants, convey their identities through the filter of transcription? How can their voices be heard in the way they wish them to be heard? How do different transcription choices (e.g. adoption of standard spelling even when a person is using a nonstandard/learner variety) affect our readers' views of the text and of the people represented within it?

Writing from a social constructionist perspective, Bucholtz (2000) argued that:

> All transcripts take sides, enabling certain interpretations, advancing particular interests, favoring specific speakers […] Embedded in the details of transcription are indications of purpose, audience, and the position of the transcriber toward the text. Transcripts thus testify to the circumstances of their creation and intended use. As long as we seek a transcription practice that is independent of its own history rather than looking closely at how transcripts operate politically, we will perpetuate the erroneous belief that an objective transcription is possible. (Bucholtz 2000: 1440)

In the multilingual research context of the present work, this understanding of transcription demanded an acknowledgement of the researcher's language-communicative capabilities and a critical engagement with the ramifications of linguistically (in)competent field research (Phipps, 2013a).

Historically, issues related to linguistic incompetence have received little attention from the academic community, although some scholars (e.g. Andrews, 2013; Ganassin & Holmes, 2020; Gibb *et al*., 2019; Gibb & Danero Iglesias, 2017) have recently examined how their own (lack of) knowledge of different languages affected their research activities. But why do others take refuge in silence? According to Tonkin (1984: 178), being unable to learn a language and admitting to serious difficulties in communication is often perceived by ethnographers as a 'failure to measure up to a publicly required occupational definition'. After all, if we cannot learn 'something as public as the language [how could we ever claim to be able] to understand the innermost meanings that people attach to things and events, or to discover the hidden mechanisms that make society function, or the secrets hidden from outsiders and casual observers?' (Borchgrevink 2003: 96). This fear of being found wanting is 'hardly conducive to frankness about language capacities and degrees of understanding' (Borchgrevink 2003: 99).

The experience of researching in second or additional languages had a strong bearing on the research questions the initial study set out to answer, the methods and processes employed in its conduct, and on decisions we made together about transcribing, analysing and disseminating the data multilingually. Arabic – the research participants' strongest language in which most peer-to-peer exchanges took place – is a case in point. As an experienced language learner, Kalocsányiová did not regard learning Arabic as a viable avenue for data gathering. Starting from scratch, a year-long learning effort in Luxembourg would have amounted to little more than basic phrases and expressions. Certainly not sufficient for getting on the inside or establishing real communication (Borchgrevink, 2003)! (Linguistic) ethnographers like to believe that their learning efforts 'symbolize a commitment, a respect and appreciation for the cultural heritage of the people they study' (Duranti, 1997: 111). However, we believe it is equally important to be realistic and acknowledge that most of us will remain less-than-fluent in a new language learned for fieldwork purposes (see also Gibb & Danero Iglesias, 2017; Tremlett, 2009). At the same time, silencing the languages spoken by (forced) migrants risks neglecting their lived experience and possibly misrepresenting their realities. It also forecloses important analytical possibilities as we will show in the next section.

Finally, no contemporary discussion of the politics and processes of (multilingual) transcribing could be complete without addressing the question of translation. Since an important part of mainstream research is currently conducted in (super)diverse societies (but published in English), researchers need to find ways 'to carry meanings across linguistic, discoursal and cultural boundaries' (Holliday, as cited in

Holmes *et al*. 2013: 293). The view of translators has long shifted from that of *conduits of communication* (Freed, 1988) to *mediators* of culture and cross-cultural communicative functions (Bedeker & Feinauer, 2006). Yet, in most migration and social research contexts their contributions are expected to be 'neutral and invisible' (Shklarov, 2007: 532) or are treated as 'needing control of some sort' (Resch & Enzenhofer, 2018: 132). This chapter takes an opposite approach and shows how transcribers and translators are the first to tap into research participants' thoughts and experiences, and thus become active producers of data, meaning and knowledge (see also Kalocsányiová & Shatnawi, 2020).

Although translation and transcription are integral (if not inevitable) parts of multilingual research, the subject has received little scholarly attention until fairly recently (e.g. Nikander, 2008; Nurjannah *et al*., 2014; Sutrisno *et al*., 2014). Some of the first publications (e.g. Nikander 2008) focused on the *physical* presentation of translation in print, noting that: 'Publication policy […] often gets to dictate how data translations are produced and presented. Each choice carries not only practical but also ideological implications of language primacy' (Ibid. 227). Others (e.g. Nurjannah *et al*., 2014) discussed the *timing* of translation in the research process, and/or produced a set of recommendations on translation in general (Resch & Enzenhofer, 2018) and for various qualitative approaches (Nurjannah *et al*., 2014; Sutrisno *et al*., 2014). In the next sections, we discuss the timing and the physical presentation of our translated data, but we also go beyond most past research efforts to consider the translation of dynamic multilingual exchanges.

Speaking and being heard through translation are also part of power relations (cf. Resch & Enzenhofer, 2018). *Who is translating whom*, *when* and *for what purposes* are all political decisions. Using a translation, rather than the original as the start point for analysis can have far reaching consequences for validity, transparency and reflexivity in qualitative inquiries, and particularly in linguistically inspired approaches (e.g. discourse analysis). Just like a researcher-transcriber, translators leave their mark on research: extensive works have shown in the past that people hear interviews and other multi-speaker scenarios through their own cultural-linguistic filters (e.g. Oliver *et al*., 2005), and as a result, each person's contribution represents 'a key moment of choice and the exercise of power in the research process' (Ross 2010, np.). There is not and cannot be a *single* correct translation. Bucholtz (2007: 801), for instance, raised interesting questions about the 'unhappy choice between a colloquial translation style, which may imply that the speakers are "just like us", despite significant cultural and other differences, and a formal translation style, which may position the speakers as "not like us at all", but rather as foreign, stiff and oldfashioned'. This is one of the many challenges that have to be tackled when researching multilingually. In the process, multilingual researchers are likely to face an ethical double bind, that is, the 'impossibility of fully rendering another's voice or meaning, and yet the necessity of making the attempt' (Apter, 2005: 89). Finally, the choice of language medium for publication is among, if not, the most political decision researchers from multilingual contexts face (see also Robinson-Pant, 2017). The remainder of this chapter explains how we tackled these challenges during our joint transcript construction.

Researching multilingually: background, processes, and examples

The transcripts discussed in this chapter are part of the data collected for Kalocsányiová's doctoral research at the University of Luxembourg. Taking a qualitative approach informed by linguistic ethnography (Copland & Creese, 2015; Rampton *et al*., 2015), she investigated the language learning and integration experiences of five men who, fleeing war and violence, sought refuge in the Grand Duchy of Luxembourg. The fieldwork of this project was carried out between 2016 and 2018. In the first phase, French, English, and German language courses were observed to examine if and how the research participants' language resources were deployed, overlooked, or silenced in classroom-based training. In the second phase, the research centred on the learning trajectories of the same five participants across time, societal influences, and language ideologies.

We have chosen to focus on Kalocsányiová's work as it entailed research *in* and *into* multiple languages, including French, Arabic, German, and English. Multilingualism in all its hybrid forms and manifestations was both a research object and a methodological element to consider in her project. In what follows, we first examine the written representations of multilingual classroom interactions from recording, through translation to publication. We then discuss interview extracts that capture the project participants' voices in English. The parts that pertain to Kalocsányiová's field research experience will be reported in first person singular.

Transcribing multilingual classroom discourse

Prior to relocating to Luxembourg I operated daily in four to five languages (mainly English, Spanish, Hungarian, and Slovak); and yet I experienced the above-mentioned project from a position of considerable 'lack, limitation, wound and partiality' (Phipps, 2013a: 336). The fieldwork during the first eleven months took place chiefly in French language classrooms; for me this meant collecting data in a language in which I was *less-than-fluent* at that point. These days, researchers often find themselves working in communities where multiple languages are in contact and many of us feel that our field language competence is insufficient (e.g. Tonkin, 1984; Moore, 2009; Tremlett, 2009). It is also hard to avoid the impression that field researchers have somewhat of a double standard when it comes to language:

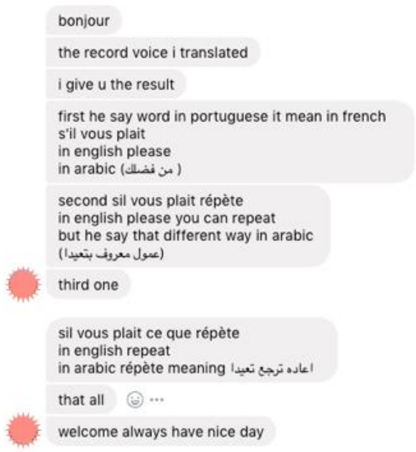
> While it was generally considered adequate for an [ethnographer] to investigate non-European culture without developing sufficient speaking competence in the relevant language, the same low standard did not apply to the study of western cultures. That is, whereas one wouldn't expect a fieldworker working on Kiyansi culture (Democratic Republic of Congo) to "virtuously" master Kiyansi […] one would accuse of lack of professionalism any fieldworker who has only an approximate command of French. (Vigouroux 2007: 76)

Vigouroux's reflections resonated strongly with me. It was nevertheless the awareness of being less-than-fluent that led me to reflect on transcription as a methodological process. After the first cycle of classroom ethnographies, I started working on the transcripts wanting to record and represent everything as 'precisely' and 'faithfully' as possible. Transcribing in multiple languages – some of which I had only a surface knowledge of – felt unorthodox initially, and it induced a certain sense of guilt and discomfort with the data.

Being a non-speaker of Arabic restricted what I could do (particularly during a 4-year doctorate), and I had to tailor my research to this limitation in several ways. In developing strategies for multilingual data collection and analysis, I followed some of Tremlett's (2009) suggestions. For instance, I made extensive use of audio recording to capture interviews, talk-in-interaction and classroom exchanges on tape so that these could later be transcribed, translated and commented on both by me and others. It was the recorded data that enabled me to overcome the anxieties I felt in being less-than-fluent/lacking the language, as it allowed for re-living the classroom exchanges at my own pace, and as many times as necessary. Without a systemic audio recording the language dimension of the fieldwork would have been a tough nut to crack: 'although not infallible (and of course with its own representational issues in the processes of production, translation and analysis), [recorded data] at least could be put on view, re-played, re-checked and displayed more confidently […]' (Tremlett 2009: 81).

Still, producing (bi)multilingual transcripts with even a low-level of interactional detail was a time-consuming venture unfeasible for all classroom data (approximately 70 hours of recording). Hence, the ambition of a verbatim multilingual transcript with all exchanges represented as uttered was traded off against a thick description pieced together from timestamped field notes, extensive descriptive summaries, and dozens of rough-draft transcripts. At this stage, informal translations were often solicited from the research participants for instance during breaks or after classroom observations. Working closely with research participants to decode and couch the language in writing has a long tradition in ethnography-inspired research (Vigouroux, 2007). Most of the exchanges about translation, transcripts and meanings are not preserved in the audio files; however, there were instances in which texts were produced and shared through messaging applications.

Extract 1.



Extract 2.

R [Pouvez-vous ::]

يعني من فضلك por favor بالبرتغالي يعني
in Portuguese por favor *which means please*

M s'il vous plait

R repetir إنّك تعود إنّك ترجع تعيدها مرّة ثانية إعمل معروف أنّك تعيدها
repetir *means you need to say it again, say it again make me a favour and say it another time*

R مرّة ثانية
another time
[…]
من فضلك ترجع تعيدها مرة ثانية pouvez-vous (--) parler
could you please repeat it again

P repetir

R انك ترجع تعيدها مرة ثانية
you should repeat it again
من فضلك انك ترجع تعيدها مرة ثانية
repeat it again please

The rough-draft transcript (Extract 1) was produced by a research participant – Ahmad (pseudonym) – who had on occasion volunteered for small translation tasks both within and outside the project. The same excerpt was later re-worked and published as an exemplification of translanguaging and meaning-making processes in the context of a French classroom (Extract 2, Kalocsányiová, 2017: 488). Ahmad's initial help in the field was invaluable; however, he was not versed in transcription techniques required for more layered representations. These were done in collaboration with Malika Shatnawi (second author). In the course of this joint work, we quickly realised that collaborative (multilingual) transcribing constituted an important area of inquiry in its own right.

Being in the field allowed me to create a first layer of transcription signalling the speakers, contextualisation cues, possible intended meanings, and any chunk of talk uttered in a language other than Arabic at the research sites. Afterwards, drawing on her linguistic and cultural expertise Shatnawi completed the Arabic-speaking parts and proposed a translation into English. Throughout the process, we engaged in a continuing dialogue, audio-recording our online work sessions to better capture interpretative practices and reflections about multilingual data representation. Embedding Arabic script (which is written right-to-left) in European languages written left-to-right was one of the obvious difficulties to tackle. Other questions that arose included: *How do we ensure readability*? *Shall we label languages as distinct codes*? *Do we record pronunciation particulars*? *How do we represent translingual practices in translation (if at all)*? The need for translation combined with the use of multiple scripts amplified the issues of transcript formatting and layout.

Extract 3.

peer-to-peer: participants (M and R)		whole-class: teacher (T) and student (F)	
		T	gèlera
R	gèlera gèl- ((to the class))		
		T	alors le gel ((writing on the whiteboard))
M	gel تلج *ice*		
R	جيل يعني تلج (0.1) *gel is snow* (0.1)		
		T	l'adjectif c'est gelé (--)
M	gelé ((to the class))		
			et le verbe (0.1) il gèle
R	il gèle ((to the class))		
M	مصقع- صقيع *frosted- frost*	T	ou il gè-le-ra
R	هي (0.1) كتبتها صح -كتبتها صح *this (0.1) if I write it correctly- if I write it correctly*		
		F	(gel) ça fait très froid
		T	moins de zéro (0.1) degré (-) moins un
M	بتطلع اكتر من عشر *it takes more than one tenth*		
R	هل هاي صحيحة *is this correct?*	T	moins deux
M	لا في هااي(0.1) هاي هي *no there's* (0.1) *there: is* (xxx) اه ههhe ah ((laugh))	T	moins trois (-)
		T	c'est le gel
		F	ça fait très froid ?
		T	oui […]

In Extract 3, the data are presented in two parallel columns to keep better track of the speakers and their turns. The learners/participants' contributions (M and R) are given primacy and visibility through their placement on the left-hand side (Ochs, 1979) (the primary audience to whom this example was presented included speakers of French, German and English, i.e. left-to-right readers). For better readability, we decided to limit the prosodic and paralinguistic cues to signalling pauses, overlapping speech, and sound lengthening. However, even this level of detail often proved chaotic across languages, translations and two writing systems. As shown in Extracts 2 and 3, we constructed the transcripts in the source language(s). Para-linguistic annotations and transcriber comments were added in English (the common language between us). Prior to dissemination, translations were introduced only for words and turns in Arabic (with the original preserved) to support a collaborative in-depth analysis. The first draft translations from Arabic into English were generally more detailed than shown in Extract 3, supplying the first analyst (me) with information on the structure and semantics of the original language. This was essential for identifying moments of interlinguistic transfer and peer-to-peer translation, which both allowed for insights into the communicative resourcefulness of the research participants (one of the major themes of the study).

In many published works, researchers carried out translations from audio recordings directly into English (e.g. Andrews, 2013), or prioritised a coded English translation over a multilingual transcript, with or without the source language version attached as a supplement to the published article (e.g. Kirsch 2017). Such representational choices may sometimes be necessary or even desirable due to time constraints or word limit restrictions, but Nikander (2008: 227) warns us that ‘hiding the original from the reader’s view [can] violate the validity [of the research] through the transparency and access principle’.

One of the study’s aims was to document the language learning trajectories of newly arrived migrants, and thus the question arose whether and how to record pronunciational particulars in writing. From the perspective of language acquisition and interlanguage research, there is an analytical advantage to using nonstandard orthography or phonetic transcription in International Phonetic Alphabet (IPA) to represent second/additional language learners’ linguistic performances. Taking this approach, however, created new complexities. First, as co-transcribers our perceptions of sounds and voices differed. Bucholtz (2007) explained this phenomenon as placing speech sounds into one’s internalised phonemic categories: ‘what counts as salient depends on the “listeners” own dialectal [and we would add language] backgrounds and phonological expectations as well as the practices of “professional hearing” into which they have been socialized’ (ibid. 798). Second, the likelihood of phonemic false evaluations (Buckingham & Yule, 1987) and subsequent misrepresentations increases when transcribing a multilingual corpus. Third, although we had both completed graduate-level courses on phonetics and IPA transcription as part of our studies, we fall short of the expertise necessary for developing a systematic and rigorous phonetic representation across multiple languages and levels of fluency. This forced us to abandon some of the research goals and re-focus our attention to areas for which our transcripts were appropriate (e.g. a qualitative inquiry into the affordances of multilingual learning situations in classroom-based language training).

Although we opted for transcribing orthographically (i.e. employing the standard spelling system of each source language), there were instances that called for the inclusion of some pronunciation details, as illustrated in Extracts 4 and 5.

Extract 4.

M (research participant); T (French teacher)

M	[…] Paris Saint Germain gagné quatre zero /*zirō/
T	quatre nul
M	autour game l'autour jeu
T	le match de retour?
M	le match de retour Barcelona a gagné six /siz/ un haha comment cinema (/*sinəmə/) ((chuckles)) (0.2) oui (-) action /ækʃən/ (--) action /aksjɔ̃/

Extract 5.

M (research participant); T (French teacher)

M	[…] Paris Saint Germain gagné quatre **zero**
T	quatre nul
M	autour **game** l'autour jeu
T	le match de retour?
M	le match de retour Barcelona a gagné six un haha comment **cinema** ((chuckles)) (0.2) oui (-) **action** (--) action

Recording pronunciational particulars served to differentiate word pairs that have identical or almost identical spelling across two or more languages, as in action (/ækʃən vs. aksjɔ̃/) or cinema (/sinəmə vs. sinema/). The failure to record these pronunciation details would have rendered some of the learner's (M) translanguaging shifts invisible in the transcript. Extract 4 was a working draft that provided a springboard for further inquiry into M's emerging multilingualism and creative language uses. An alternative representation is proposed in Extract 5, where elements that do not conventionally belong to the French lexicon are bolded. This second version was used as a discussion starter in an undergraduate course for future school teachers, where a familiarity with phonetic transcription systems could not be assumed.

The choice of demarcating languages through colour-coding or font variation (like in Extract 5) extends beyond questions of audience, readability and styling. It is rooted in some fundamental onto-epistemological differences between translanguaging and code-switching (code-mixing), which have both been widely used in educational contexts. Our work has largely drawn on the concepts of 'translanguaging' and 'communicative repertoires': these see 'multilinguals as possessing a unitary linguistic system […] that is not compartmentalized into boundaries corresponding to those of the named languages' (García & Otheguy, 2020: 25). Fluid language uses emerged in our data as a result of the researcher and participants' restricted competence and their deployment of 'those linguistic, mental, interactional as well as intercultural competences which are creatively activated when interlocutors listen to linguistic actions in their "passive" language[s]' (ten Thije *et al*., 2012: 249). At other times, the choice to engage in complex exchanges mobilising entire repertoires was an expression of the research participants' emerging identities as multilingual language users. Consequently, our default transcription choice was to not label these resources as belonging to separate languages wherever possible. This is not to say, however, that demarcating distinct languages was not useful, or even necessary in some of our materials (e.g. in Extract 5).

Many transcription choices were in fact altered over time to accommodate disciplinary practices, editorial requests, and potential audiences (see also Mondada 2007). One of the details that we added/subtracted according to the (assumed) audience was translation. For instance, in the article containing Extract 2 (Kalocsányiová, 2017) only the Arabic utterances were initially translated into English. It was during the revision process that equivalents were added for the other languages as well, principally, to make the analysis more accessible to an English-speaking readership. This task also posed the difficult question of representing translingual practices in translation:

> بالبرتغالي يعني por favor يعني من فضلك
> *in Portuguese* por favor *which means please*

After long discussions, we decided to translate the matrix (dominant) language of the text (in the above case, Arabic) into English, while the embedded elements ('por favor' above) were transferred in an unchanged form. Importantly, our translations presented in Extracts 2 and 3 should be read as approximations. The fragmented nature of spoken language makes it very difficult to come up with translations that reproduce 'the original in all its meanderings': a translation which 'mimics a very fragmented and hesitant original can be perceived as exaggerated, as not just translating but also evaluating what has been said (Bot & Wadensjö, 2004: 360). Our examples, of course, are not intended to provide an oversimplified model for translating dynamic multilingual exchanges. Keeping or translating source language elements will always depend on at least two factors: the assumed audience's linguistic profile and the analytical argument one is trying to make. Developing Jaffe's argument (2006: 180), we also believe that as researchers we should be realistic about 'our ability to make readers *see* [in English] what we *hear* [in other languages]'. Consequently, we argue that original data should be given primacy in research wherever possible.

Finally, if there is one aspect that this work once again highlights, it is that collaboration with study participants, transcribers, and translators brings researchers closer to sensitivity, transparency and reflexivity in qualitative inquiry (Jeppesen & Hansen, 2011). Collaborative research processes also generated richer data and more accurate reporting (see also Kalocsányiová & Shatnawi, forthcoming).

Transcribing interview data

Most work discussed up to this point has been concerned with collaborative transcribing and the technical features of multilingual transcription. The question that underpins this section is equally, if not more challenging for researchers working with marginalised and migrant populations: How can forced migrants, especially those speaking non-native language varieties, convey messages about their lives and identities through the filter of transcription? 'Nativeness', 'competence' or even 'language' have come to be recognised as problematic in applied linguistics (e.g. Gibb & Danero Iglesias, 2017; North & Piccardo, 2016; Phipps, 2013a); however, the ubiquity of *languaging* and *linguistic fluidity* is yet to be registered on the radar of scholars in other disciplines.

In social sciences and migration research, one's mother tongue is almost always positioned as the most desirable interviewing language. But what if research participants' expectations or wishes are at odds with this common practice? Tsuda (2010: 261) reminds us of the importance of respecting people's language choices: 'Everyone is entitled to the right to use the language(s) s/he chooses to speak and this right should be honoured in *all forms of communication*' [emphasis added]. Guided by this view, the participants in this study were given the choice of doing the interviews in their language of preference. Most of them opted for English. Arguably, they felt a sense of pride and accomplishment in speaking a language other than their mother tongue.

Already two decades ago, Kvale (1996: 172–173) argued that a 'verbatim transcribed oral language may appear as incoherent and confused speech, even indicating a lower level of intellectual functioning' and cautioned that its publication might 'involve an unethical stigmatization of specific persons or groups of persons'. These effects may well be magnified if a person uses a nonstandard/learner variety of English, especially in publications aimed for a broader audience across various disciplines (e.g. Kalocsányiová 2020). On this point, we concur with Jaffe (2006: 181): everyday habits of reading may lead audiences 'to read stigma in analyses that we intend to have the opposite effect'. The data extracts analysed in continuation (such as Extract 6 below) illustrate some of the difficulties associated with transcribing interviews conducted with forced migrants.

Extract 6.

P yeh↑ I did not found a place (0.2) in the beginning ↓ (0.3) I found a global (resort)
was giving me bon ehm: for the first time ↓ (0.1) my social agent she gave me
bon for French (0.1)
R [ehm]
P [I] >don't have I WANT< bon for Luxembourgish (0.2)
she said NO you should start with French (0.3) she:: put me in a >very
embarrassing– I said< I NEED– okay <I need a bon to study German> (0.2)
she said that's °impossible° ((mimicking the voice)) (0.2) I talked to the social–
also contact[ed] her again then she gave me a bon (01) but [I:]
R [how] did she explain
why you SHOULD learn French and not Luxembourgish or German?
What was the=
P the way that he- she she pushed me to (it) she did not give me (0.1) a bon↓
R did she give you any explanation w[hy– why↑]
P [said that you >you don't have a PAPER< (0.2)
and we cannot give you a bon for Luxembourgish to study (0.4) because you
don't have to re– >may↑be the:y< eh:m they give you a REJECtion (0.1)
and you will go back to your country >you– we don't know< (0.1)

Extract 6 offers a detailed visual representation of the research participant's (P) communicative repertoire, voice and habitual form of expression in English. Similarly, it captures the interactional dimension of qualitative interviewing, by showing, for example, the researcher (R) asking for clarification and expansion. It also documents interruptions, which signal the strength of feeling regarding the issue at hand (see Greatbatch, 1992). At the same time, however, it also puts *on display* and *intensifies* the participant's disfluencies (Jaffe 2006). These indicate far more uncertainty when transmitted in text than in the original recorded sound. Furthermore, the level of detail addressed above could hardly ever be reported outside linguistically-oriented inquiries for a number of reasons. First, a verbatim reproduction of the entire sequence would simply take up too much space considering the world count guidelines of most scientific journals. Second, the extract bears little resemblance to typical academic prose and might look 'like unintelligible gibberish to the untrained eye' (Bot & Wadensjö, 2004: 361). It certainly draws attention to itself by being 'manifestly un-text-like' (Ross, 2010: np), and would need to be read out loud in order to understand that it made perfect sense to the people talking.

These considerations prompted us to explore alternative theories and ways of transcribing. Kvale (1996) encourages researches to think of transcripts as interpretative constructs. As a guiding principle, he proposes to imagine how interviewees themselves would have wanted to formulate their statements in writing. Although this suggestion has its limitations, it does push for the criticism of ‘the naïve assumption that talk is easily converted into written text’ (Tilley, 2003: 760). Roberts (1997: 169) also argues for experimental ways in which speakers’ voices can be evoked and contextualised:

> [transcribers] have a duty to find ways – perhaps experimental ways – for evoking some of the struggles, some of the emotional qualities of an encounter in which minority ethnic speakers are striving to be themselves in another language […] these encounters are so frequently asymmetrical and stressful, transcribers have the additional duty of accounting for these effects on the talk they transcribe.

Seeing transcripts as constructs created for particular audiences allows for challenging the widespread acceptance of verbatim transcription as the sole desirable element in research methodologies (Bezemer & Mavers, 2011). Furthermore, as Phipps (2013b: 21) explains, ‘[t]he flattened, coded tones of transcripts, with their numbered rows are divorced from the highly storied, narratively and performatively rich contexts of intercultural communication’. Privileging exact words over the lived experience of the interview as a ‘moment of engagement, a site of participation in the life [of others]’ (Hockey & Forsey, 2012: 75) can, in our experience, misrepresent a person or situation.

Before we proceed, a few words are in place about the research process. After each interview, Kalocsányiová took notes of major ideas, experiences and anecdotes reported by the research participants. These notes were written in English (her ‘third’ language), though, logically, they contained fragments of other languages as well. As a next step, she reviewed the audio recordings while the reflections were still fresh and gradually amended the initial notes until a descriptive representation of significant interactions was created. This served as a springboard for subsequent detailed *verbatim* transcriptions (e.g. Extract 6), *covert* transcripts (Ross 2010) (e.g. Extract 7), as well as for *re-constructions* of participants’ voices through stories (Extract 8). Roughly similar approaches to data reporting are discussed in the literature under headings such as ‘reflective journalizing’ (Halcomb & Davidson, 2006) ‘reflexive interviewing and reporting’ (Loubere, 2017) and ‘construction of journalistic narratives’ (Jeppesen & Hansen, 2011).

Extract 7.

> However, as the interviews unfolded it became obvious that the “choice” to learn French was to a great extent imposed upon them: “the social agent gave me a *bon* for French but I asked for Luxembourgish and she said no, you should start with French [….] I said okay, I want a *bon* for German but she said it was not possible.” (August 17, 2016) The social worker’s conduct could be explained by the widely held belief that French facilitates economic integration better than any other language in the local labor market. (Kalocsányiová, 2020: 225)

The quote in Extract 7 (which is an edited version of a sequence from Extract 6) can be considered an example of covert transcription: a format that ‘blends in seamlessly to material which was “born” textual’

(Ross, 2010: np). Covert transcripts are common, if not universal in social science research outside language-related fields. The extract above took its final shape after negotiations with the publisher's editors who asked for removal of interview noise, punctuated sentences, and compliance with standard writing conventions and grammar rules. The subtraction and amendment of these details was meant to create a recipient-oriented presentation (Mondada, 2007) and avoid the stigma attached to 'incoherent and confused speech' (Kvale, 1996) and language uses considered foreign or spoken by minorities.

Extract 9 exemplifies another representational choice aimed at avoiding stigmatisation and myths of deviance (Kvale, 1996):

Extract 9.

> During our second meeting (September 17, 2016), Ahmad told an anecdote which exemplifies the initial confusion he experienced. In his imagination, Luxembourg was a German-speaking country: "I didn't know anything, I just thought it was like Germany." A couple of hours after his arrival, he and his travel companions overheard a conversation in (what they believed was) French at the refugee center. Driven by curiosity, Ahmad asked around among the other residents at the center, who gave him his first bits of information about Luxembourg's language environment. Once he corroborated that "French was everywhere", he asked in bewilderment, "What comes next?" (Kalocsányiová, 2020: 224)

Extract 9, which, prior to its publication, was presented to audiences from migration studies, sociolinguistics and (language) education backgrounds, entailed a profound reorganisation of data and meaning, based on a 'dialogue between speaker and listener(s), investigator and transcript, and text and reader' (Riessman, 2008: 139). Though this final format provides little information about the participant's communicative repertoire, idiolect or the structural features of his talk, it might well be the best *translation* of his intended meanings into an academic output. As Vakser (2017) has argued, in representing research participants we must be highly attuned to what is said, how it is conveyed (and we would add) how it may be perceived across different contexts and disciplines.

Conclusions

This chapter has examined examples of multilingual transcription and representational choices for interviews conducted with forced migrants. It has built on data pulled from a larger qualitative study on forced migrants' linguistic integration. In reflexively analysing different transcript formats (co-)constructed between researcher and transcriber/translator at different points of the research process, we demonstrated the value and importance of continual reflection on the part of qualitative researchers about language (in)competence, (in)visibility, voice and stigma. Overall, our examples have illustrated the consequences particular theoretical and methodological approaches may have for transcribing practice and data reporting. The role of language-communicative capabilities, ethical considerations, and disciplinary boundaries have also been addressed and exemplified through our experience of (multilingual) transcribing.

The chapter has outlined different proposals for evoking the voices of forced migrants who ‘are striving to be themselves in another language’ (Roberts (1997: 169). We invite other researchers, including those in other disciplines, and especially those working with vulnerable and marginalised populations to reflect more deeply about their transcription choices, and account for these in their research. An interdisciplinary perspective is certainly necessary to explore more fully the processes and politics of this highly complex research component. Our findings are expected to open up new avenues for theoretical inquiry and foster a more reflective practice of transcribing.

Despite its limited scope, this chapter also adds to the rapidly expanding scholarship on the theory, methods and ethics of researching multilingually. It provides a detailed account and examples of how to engage with multilingual research materials and outputs collaboratively. There is wide agreement in the research community that a greater involvement of transcribers and translators in researcher processes produces richer descriptions of the phenomena under investigation. In this chapter, we have shown that collaborative transcribing also encourages reflexivity, enables a critical engagement with data, and contributes to a more transparent reporting. Hopefully, our contribution will be useful for early career researchers dealing with similar issues in their own work, and also established researchers who are challenged by the prospect of researching multilingually. Finally, we encourage doctoral researchers to resist monolingual institutional norms and to engage in discussions with their supervisors and peers about language choices for transcription, analysis, and research dissemination.

References

Andrews, J (2013) “It’s a very difficult question isn’t it?” Researcher, interpreter and research participant negotiating meanings in an education research interview. *Applied Linguistics* 23 (3), 316–328.

Apter, E. (2005) PART TWO Introduction. In S. Bermann and M. Wood (eds) *Nation, Language, and the Ethics of Translation* (pp. 89-92). Princeton: Princeton University Press.

Bedeker, L. and Feinauer, I. (2006) The translator as cultural mediator. *Southern African Linguistics and Applied Language Studies* 24 (2), 133–141.

Bezemer, J. and Mavers, D. (2011) Multimodal transcription as academic practice: a social semiotic perspective. *International Journal of Social Research Methodology* 14 (3), 191–206.

Bird, C. M. (2005) How I stopped dreading and learned to love transcription. *Qualitative Inquiry* 11 (2), 226–248.

Borchgrevink, A. (2003) Silencing language: of anthropologist and interpreters. *Ethnography* 4 (1), 95–121.

Bot, H. and Wadensjö, C. (2004) The presence of a third party: a dialogical view on interpreter-assisted treatment. In J. P. Wilson and B. Droždek (eds) *Broken Spirits: The Treatment of Traumatized Asylum Seekers, Refugees, and War and Torture Victims* (pp. 355–378). New York: Brunner-Routledge.

Bucholtz, M. (2000) The politics of transcription. *Journal of Pragmatics* 32, 1439–1465.

Bucholtz, M. (2007) Variation in transcription, *Discourse Studies* 9 (6), 784–808.

Buckingham, H. and Yule, G. (1987) Phonemic false evaluation: theoretical and clinical aspects. *Clinical Linguistics and Phonetics* 1, 113–125.

Copland, F. and Creese, A. (2015) *Linguistic Ethnography: Collecting, Analysing and Presenting Data*. London: SAGE Publications.

Duranti, A. (1997) *Linguistic Anthropology*. Cambridge: Cambridge University Press.

Freed, A. O. (1988) Interviewing through an interpreter. *Social Work* 33, 315–319.

García, O. and Otheguy, R. (2020). Plurilingualism and translanguaging: Commonalities and divergences. *International Journal of Bilingual Education and Bilingualism* 23, 17–35.

Ganassin, S. and Holmes, P. (2020) ‘I Was Surprised to See You in a Chinese School’: Researching Multilingually Opportunities and Challenges in Community-Based Research. *Applied Linguistics*. doi: 10.1093/applin/amz043

Gibb, R. and Danero Iglesias, J. (2017) Breaking the silence (again): on language learning and levels of fluency in ethnographic research. *The Sociological Review* 65 (1), 134–149.

Gibb, R. *et al* (2019) *Learning and Using Languages in Ethnographic Research*. Bristol: Multilingual Matters.

Greatbatch, D. (1992) The management of disagreement between news interviewees. In P. Drew and J. Heritage (eds) *Talk at Work* (pp. 268–301). Cambridge: Cambridge University Press.

Green, J. *et al*. (1997) The myth of the objective transcript: Transcribing as a situated act. *TESOL Quarterly* 31 (1), 172 –176.

Halcomb, E. J. and Davidson, P. M. (2006) Is verbatim transcription of interview data always necessary? *Applied Nursing Research* 19 (1), 38–42.

Hockey, J. and Forsey, M. (2012). Ethnography is not participant observation: Reflections on the interview as participatory qualitative research. In J. Skinner (ed.) *The Interview: An Ethnographic Approach* (pp. 69–88). London: Berg.

Holmes, P. *et al* (2013) Researching multilingually: New theoretical and methodological directions. *Applied Linguistics* 23 (3), 285–299.

Jaffe, A. (2006) Transcription in practice: Nonstandard orthography. *Journal of Applied Linguistics* 3 (2), 163–183.

Jeppesen, J. and Hansen, H. P. (2011) Narrative journalism as complementary inquiry. *Qualitative Studies* 2 (2), 98–117.

Kalocsányiová, E. (2017) Towards a repertoire-building approach: multilingualism in language classes for refugees in Luxembourg. *Language and Intercultural Communication* 17 (4), 474–493.

Kalocsányiová, E. (2019) *Towards an Understanding of the Language–Integration Nexus: A Qualitative Study of Forced Migrants' Experiences in Multilingual Luxembourg*. Esch/Alzette: University of Luxembourg [unpublished thesis]

Kalocsányiová, E. (2020) Researching forced migrants' trajectories: encounters with multilingualism. In B. Nienaber and Ch. Wille (eds) *Border Experiences in Europe. Everyday Life – Working Life – Communication – Languages* (pp. 215–234). Baden Baden: Nomos.

Kalocsányiová, E. and Shatnawi, M. (2020) 'He was obliged to seek refuge': an illustrative example of a cross-language interview analysis. *Qualitative Research*.

Kirsch, C. (2017) Young children capitalising on their entire language repertoire for language learning at school. *Language, Culture and Curriculum* 31, 39–55.

Kvale, S. (1996) *InterViews: An Introduction to Qualitative Research Interviewing*. Thousand Oaks: SAGE Publications.

Loubere, N. (2017) Questioning transcription: the case for the systematic and reflexive interviewing and reporting (SRIR) method. *Forum Qualitative Sozialforschung* 18 (2), art. 15.

Mondada, L. (2007) Commentary: transcript variations and the indexicality of transcribing practices. *Discourse Studies* 9 (6), 809-821.

Moore, L. C. (2009) On communicative competence…in the field. *Language and Communication* 29 (3), 244–253.

Nikander, P. (2008) Working with transcripts and translated data. *Qualitative Research in Psychology* 5, 225–231.

North, B. and Piccardo, E. (2016) Developing illustrative descriptors of aspects of mediation for the Common European Framework of Reference (CEFR). *Language Teaching* 49 (3), 455–459.

Nurjannah, I. *et al*. (2014) Conducting a grounded theory study in a language other than English: Procedures for ensuring the integrity of translation. *SAGE Open* 4 (1), 1–10.

Ochs, E. (1979) Transcription as theory. In E. Ochs and B. Schiefflin (eds) *Developmental Pragmatics* (pp. 43–72). New York: Academic.

Oliver, D. *et al*. (2005) Constraints and opportunities with interview transcription: towards reflection in qualitative research. *Social Forces* 84 (2), 1273–1289.

Phipps, A. (2013a) Linguistic incompetence: Giving an account of researching multilingually. *Journal of Applied Linguistics* 23 (3), 329–341.

Phipps, A. (2013b) Intercultural ethics: questions of methods in language and intercultural communication. *Language and Intercultural Communication* 13 (1), 10–26.

Rampton, B. *et al*. (2015) Theory and method in linguistic ethnography. In J. Snell *et al*. (eds) *Linguistic Ethnography: Interdisciplinary Explorations* (pp. 14–50). Basingstoke: Palgrave Macmillan.

Resch, K. and Enzenhofer, E. (2018) Collecting data in other languages: Strategies for cross-language research in multilingual societies. In U. Flick (ed) *The SAGE Handbook of Qualitative Data Collection* (pp. 131–146). London: SAGE.

Riessman, C. K. (2008) *Narrative Methods for the Human Sciences*. Thousand Oaks: SAGE Publications.

Roberts, C. (1997) The politics of transcription. Transcribing talk: Issues of representation. *TESOL Quarterly* 31 (1), 167–172.

Robinson-Pant, A. (2017) Research in multilingual contexts: joining an international community of researchers. In A. Robinson-Pant and A. Wolf (eds) *Researching Across Languages and Cultures* (pp. 117-135). Oxon: Routledge.

Ross, J. (2010) Was that infinity or affinity? Applying insights from translation studies to qualitative research transcription. *Forum Qualitative Sozialforschung* 11 (2), art. 2.

Shklarov, S. (2007) Double vision uncertainty: the bilingual researcher and the ethics of cross language research. *Qualitative Health Research* 17, 529–538.

Sutrisno, A. (2014) Incorporating translation in qualitative studies: two case studies in education. *International Journal of Qualitative Studies in Education* 27 (10), 1337-1353.

ten Thije, J. D. *et al*. (2012) Receptive multilingualism – Introduction. *International Journal of Bilingualism* 16 (3), 245–247.

Tilley, S. A. (2003) 'Challenging' research practices: turning a critical lens on the work of transcription. *Qualitative Inquiry* 9 (5), 750–773.

Tonkin, E. (1984) Language learning. In R. F. Ellen (ed.) *Ethnographic Research: A Guide to General Conduct* (pp. 178–187). London: Academic Press.

Tremlett, A. (2009) Claims of 'knowing' in ethnography: realising anti-essentialism through a critical reflection on language acquisition in fieldwork. *Graduate Journal of Social Science* 6 (3), 63–85.

Tsuda, Y. (2010) Speaking against the hegemony of English: problems, ideologies, solutions. In T. K. Nakayama and R. T. Halualani (eds) *The Handbook of Critical Intercultural Communication* (pp. 248–269). Malden: John Wiley.

Vakser, S. (2017) Multilingual dynamics in the research process: transcribing and interpreting interactional data. In M. Martin-Jones and D. Martin (eds) *Researching Multilingualism: Critical and Ethnographic Perspectives* (pp. 229–244). London: Routledge.

Vigouroux, C. (2007) Trans-scription as a social activity: an ethnographic approach. *Ethnography* 8 (1), 61–97.

Acknowledgements

First and foremost, we thank all research participants who made this study possible. We are also grateful to Natalia Bîlici for her support with transcription. As parts of this publication are based on Kalocsányiová's unpublished PhD manuscript, we also wish to thank Sabine Ehrhart and Ingrid de Saint-Georges (University of Luxembourg) for their comments and advice. Thanks are extended to the editors for their thoughtful comments on previous drafts. Any errors or inconsistencies are our responsibility alone.

Appendix 1: Transcription Conventions

(0.3)	timed pause
(-) (--) …	brief (untimed) silences
[ ]	overlapping speech
((laugh))	paralinguistic features and situational description
=	no gap at all between the two turns
?	rising intonation or question
!	strong emphasis
yes.	falling final intonation
descr↑iption↓	arrows denote a marked rising/falling shift in intonation
word-	false-start or self-correction
a:nd	lengthening of the preceding sound
good	emphasis
NO	large capitals indicate loud volume/strong emphasis
[…]	ellipsis
<yes>	< … > indicates slowed down delivery relative to the surrounding talk
>yes<	>…< indicates speeded up delivery relative to the surrounding talk
(xxx)	stretch of inaudible/unclear talk
/sedei/	slashes indicate phonetic transcription
*	approximate (intended) pronunciation/meaning
'text'	exact words (in descriptive transcripts)
°text°	decreased volume/whisper