RadioDoc Review

Volume 9 • Issue 2 • 2024

2024-10-25

Not just dubbing: ways to work with more than one language in audio documentaries and podcasts

Abigail Wincott, Falmouth University abigail.wincott@falmouth.ac.uk

ORCID ID https://orcid.org/0000-0002-3162-644X

Sven Preger, Neue Zürcher Zeitung

ORCID ID https://orcid.org/0000-0003-3637-3207

Abstract

In radio journalism, 'foreign language' interviewee voices have traditionally been almost entirely covered by a voiced translation of their words. In this article we examine experimentation with some different approaches to multilingual content in recent German, French and English-language audio documentaries that challenges assumptions that dubbing over other languages is the only option for factual audio. We explore the practical constraints influencing these choices, as well as the powerful social, cultural and institutional values that shape usage: concerning the status of different languages, the paradigm of monolingual media, and the quest for authenticity and direct access to reality (the illusion of non-mediation). We use this analysis to suggest some descriptors for these varied techniques and suggest that they offer not only more creative scope for producers, but also new ways for to let more voices be heard, and better reflect the rich, multilingual reality of our world.

Keywords

translation, dubbing, radio, podcasting, journalism, documentary, languages

Recommended Citation

Wincott, A. & Preger, S. (2024) "Not just dubbing: ways to work with more than one language in audio documentaries and podcasts", *RadioDoc Review 9*(2). https://doi.org/10.14453/rdr.1497

The monolingual radio myth in the 21st century

The world is often divided into countries that like their films and TV dubbed (Germany, Brazil) and those (USA, UK) who prefer subtitles (GWI, 2022). But as radio and podcasts can't do subtitles without adding video (Trelles, 2019), dubbing appears to be the only option available to producers of audio content. However, we have found German- French- and English-language audio documentary producers have been experimenting with several different approaches to the issue of multilingual contributions, and we suggest that there are in fact choices available in the treatment of other languages in audio documentary.

As with so much in the form and craft of audio storytelling, there is a gap where specialist terminology might go, and the terms used to discuss the treatment of other languages in TV and film map poorly onto audio production conventions. In this article we explore the range of ways producers have been approaching the inclusion of other languages in German, French and English-language audio documentaries. We then reflect on the implications of these choices and suggest some descriptors for them. We will use this analysis to propose ways audio documentary makers can make space for a wider range of people to be heard, including perhaps a greater diversity of voices from and in the often multilingual countries of the Global South.

The question of whether and how to dub contributions to audio documentary arises because audio content is generally conceived of as monolingual, within a monolingual media offer – that it is, for example, English language content for an English-speaking audience of an English language radio station. Producers of that content nonetheless often want to include interviewees who do not speak English, or do not speak it well enough to be interviewed in English. They might also want to incorporate archive audio that was produced in another language. But that audio must be treated in such a way as to conform to the monolingual framing of the programme.

This understanding of media as monolingual should not be taken for granted and can be situated in a particular historical and political context. Many of the nations where documentary features developed over the 20th century – the UK, Germany, Denmark, Australia, the USA – are regarded as monolingual, defined at least in part by the myth of a single national language (Yildiz, 2013) that thus became the single language of education and of the mass media in those nations.

The rise of the paradigm of monolingual nationhood in 18th century Europe never completely eclipsed multilingualism even in these countries, and language diversity has become more visible in recent years (Yildiz, 2013). In many parts of the world, particularly where national boundaries were drawn by colonial powers rather than negotiated through arguments about cultural identity, linguistic diversity is vast. According to Ethnologue, in Brazil, for example, more than 200 languages are spoken, in Nigeria, more than 500 (Eberhard, Simons and Fennig, 2023). All over the world migration produces pockets of linguistic diversity too, such as a community of German speakers in Brazil (Schulze, 2020) and the Turkish spoken in Germany (Yilmiz, 2013). The expansion of empires and post-colonial influence have fed the growth of vast global languages such as English, Spanish, French and Arabic, which are the official or de facto national languages of several countries and widely spoken in many others. Therefore the idea of a discrete monolingual media culture with its own monolingual audience is a myth, but it remains one that underpins mainstream radio and podcasting, so that podcast apps and web browsers are set to a single user language, for example, and will recommend podcasts from that language. Even where the multilingual nature of nations is recognised, the monolingual media paradigm is largely respected. For example, in Germany, public service media offer Turkish-language content to the Turkish-speaking minority especially in former industrial regions like the Ruhr area. So when pursuing international stories, consulting experts from abroad, or even following stories within their own country's borders, documentary makers will often wish to include content in other languages, yet must make it conform to the monolingual paradigm.

The most common technique used, according to these norms, is to play in the beginning of an excerpt in an 'other' language then fade it down while playing in a translation in the main language of the programme, though the original 'other' language can still usually be heard faintly and may be heard for a few moments after the translation has ended. This is a technique widely used in public radio by the BBC in the UK as well as by Deutschlandradio in Germany or SRF in Switzerland. However, we recently began to notice examples of radio and podcast documentaries that take different approaches to including more than one language. In a German podcast series about cheating scandals in international competitive chess, *Scambit: Schach, Hype und Millionen* (ACB Stories/WDR, 2023), where several interviews are conducted in English, the producers sometimes leave in exchanges entirely in English, sometimes the presenter then

includes a brief German summary in their subsequent narration of the story. At the start of the BBC's *Intrigue: Burning Sun* serial (Hadjimatheou and Puri, 2023), which investigated sexual assault scandals in the world of K-Pop, the (English-speaking) presenter announces that all the interviews were originally recorded in Korean and will be re-voiced in English by actors. And the CBC serial *Someone Knows Something* (Ridgen 2017) was dubbed in its entirety for a French-speaking audience, including the narration and dialogue of its journalist-protagonist David Ridgen.

These examples raise some interesting questions for us. Firstly, what are the concerns that producers need to balance in choosing a technique? Secondly, what is the range of techniques currently in use for including more than one language? And thirdly, what are the implications of different or changing practices for the inclusion of different languages?

Before we look at some of the techniques, we will briefly note the challenges or considerations that these programme-makers are working with. Firstly there is time. News audio packages are very short, and might be played into a longer programme of other content and subject to cuts by the producers of that programme. A longer multi-part serial still has to keep up the pace. Time is a limiting factor in a time-based medium. Secondly, there is comprehensibility, does it seem likely that the audience will understand what is being said? And of course, does the journalist or producer understand the content, where an interviewee or archive clip is in another language? Sometimes the interviewer and interviewee might need to speak a mutual language to each other, which they do not speak at home. For example a German reporter and a Swahili-speaking interviewee might be able to conduct an interview in English. Speaking what is, for them a second or third language, they may have some difficulty expressing themselves, and the translation could be used to sum up more effectively the content of what they say – addressing both consideration 1 and 2. This raises of course another complication inherent in all translation, and which we do not have room to explore in depth, namely the extent to which translation always involves interpretation, and creates new statements, or ‘language-to-language transformations’ (Baumann et al. 2011), which are never the same as the originals, but are intended as an equivalent (Delisle 1999). The judgement about what is equivalent is influenced by individual and social factors as well as journalistic organisational routines and values (Kalantari 2022). So judgements about appropriate fidelity or accuracy of translation are a third consideration for journalists and documentary-

makers. And finally, producers need to bear in mind the expectations of the radio station or podcast strand and its audience, with regards to the mode of delivery of the translation. If they are used to hearing dubbed voices, this method will be accepted, and alternative treatments may give rise to confusion or be rejected, even before broadcast or publication. So what are the options available, and how do they help programme makers work with these considerations? There is some academic writing on the question of these modes of delivery in audio-visual storytelling, but it generally concerns drama, and almost always TV and film, not radio and podcasting. In the next section we briefly review the categorisations of translation delivery techniques in this literature, before considering the options in audio-only journalism and documentary.

‘Played by actors, but their words are real’: approaches to audio documentary in a multilingual world

Writers focus most commonly on the respective merits of lip-sync dubbing and subtitles (Johnson 2020; Mehdizadkhano and Khoshaligheh 2021), though Koolstra, Peeters and Spinhof (2002) also identify three more techniques: off-screen narration (replacing off-screen narration that was provided in the original language), ‘voice over’ in news programmes; and ‘intertitles’ (where the screen is given over to text, between shot sequences). Lip-sync dubbing is used in drama to disguise the act of translation. Off-screen narration can be equated with the narration recorded in studio and used in factual audio and TV programmes to introduce speakers and events, provide facts or context or to progress the story. This voice may belong to a named presenter or reporter or sometimes an actor performs a more ‘neutral’ voice-of-god style narration, from outside the storyspace.

The ‘voice over dubbing’ mentioned by Koolstra, Peeters and Spinhof is commonly used in TV news, where the interviewee’s voice is heard for a few seconds and then faded down, at which point the voice of the person reading a translation of their words is brought in, usually with the original voice heard faintly in the background, and it might also be faded up again at the end, once the translation has finished. In this sense used by Koolstra, Peeters and Spinhof, voice over is very different to lip-sync dubbing used in drama, because there is no pretence that the translated voice belongs to the character we’re hearing from.

We know this is not the original interviewee or archive subject, whose voice remains in evidence, and who is heard first.

As we noted in the introduction, ‘voice-over dubbing’ has traditionally been the form used in audio storytelling. In the absence of visual support, subtitles/intertitles are unavailable, and there can be no lip-sync dubbing without images. In *Burning Sun* though, actors voice the English translation of Korean interviewees’ words, and the original Korean voice is entirely absent. The actors who speak their words have Korean accents and perform the role with natural-sounding intonation and emotion, so that it becomes difficult to tell who is an actor and who is a ‘real’ interviewee. This feels problematic in journalism and documentary because the real person no longer feels present. Furthermore they are represented by an actor. Whether this representation is adequate cannot be judged by listeners. The use of these techniques disturbs the audience’s ability to evaluate factual programmes by using the key criterion of truth or authenticity(Hill, 2007, 3). Although until the late 1950s, it was common for actors to revoice the comments of original interviewees (Lewis 2021), it is a fundamental expectation of modern media literacy that hearing original recordings permits us to witness at first hand a true account of reality. Though style and entertainment value also matter greatly, truth takes precedence. The deadened acoustic of a studio voice-over and the polished delivery of the actors (or the original interviewees performing scripted versions of themselves [McDonald 2020]) might even sound too much like fiction to the modern ear, blurring genre boundaries inadvertently.

The English-language CBC documentary series ‘*Someone Knows Something’* (series 3, Ridgen, 2017) about Ku Klux Klan murders has been made into a French version, also by CBC (Ridgen, 2022) for French-speaking audiences in Canada and elsewhere (it was on the Belgian RTBF podcast app Auvio as well as the CBC Ohdio app). The programme uses on-location recordings, interviews, atmos and archive. The personal journey of the investigating journalist, David Ridgen, structures the narrative – a common feature of true crime podcasts. As the first episode opens, he is visiting a site connected with the murders, with his son, to lay flowers. In the original English version he narrates in the present tense, in the first person. There is atmos of a riverside. Birds call. We hear (non-diegetic) music and what must be the voice of David Ridgen (at 00’11”): ‘It’s July 2017 and I’m standing on a muddy embankment…’ This close-miked, studio-recorded narrative voice is intercut with a recording of a conversation that is, we may

presume from our knowledge of radiophonic storytelling grammar, made at that moment on the riverbank, in 2017. In this recording, the voice of David Ridgen can be heard, but more distant from the microphone this time, in an outdoor acoustic, projecting further as he asks (00’23”) ‘what kind of swallows are those?’ Another male voice responds off-mic, barely audible ‘I think they might be bank swallows’. David Ridgen’s voice replies ‘really?’ Studio-narrating David Ridgen picks up his monologue then: ‘I’m here with my 17-year-old son, Owen, who’s been to Mississippi with me many times over the years.’ This is a fairly complex structure that, in only a few seconds hops between narrative past and present, and between spaces of the riverbank and studio. Sometimes a speaker addresses us as listeners, sometimes another protagonist. It is possible for listeners to interpret and navigate this complexity because of familiarity with the form, and how to interpret its subtle cues such as atmos, distance from microphone and tone of voice.

The French-language version follows this structure exactly, but all English-language elements are replaced with French-language versions recorded in studio. First comes the atmos with birds, then the music, and then the narrating male voice (at 00’35”): ‘Nous sommes en juillet 2017 et je me tiens sur un talus boueux…’ His tone of voice is rather sombre. The microphone positioning and the acoustic match that of the English original, in the case of the narrating voice. However, when the exchange between David Ridgen and his son Owen follows in French, the French-speaking voices are also recorded in studio, close miked, spoken by two actors – the one speaking Owen’s lines closer than the one speaking the lines of David Ridgen. The studio recorded quality surely encourages us to remember that, though the voices say they are there on that riverbank, in fact they are not. They ‘act’, in that there is appropriate intonation – questioning, hesitantly answering, querying. But are they perhaps holding back a little, in a way that signals an incomplete commitment to embodying these two people, so that we might read that they are not pretending to be David Ridgen and Owen, but only French-language conduits of their words? The murmur of other voices can be heard beneath this exchange, so low the words cannot be distinguished.

Still the narration continues, in French in the first person (at 00’52”), ‘I’m here with Owen, my 17-year-old son’. The first person narration, the performance of directly translated dialogue, rather than summarising it in the third person, for example, ‘David is with his 17-year-old son Owen’ invite a reading of the text as

'real', but the low original voices left in, the slightly underperformed delivery of lines, and the introduction from the real David Ridgen at the start invite us to remember this is re-enacted. The web text for the series credits the actors as it would in a drama: 'David Ridgen est interprété par Martin Watier et Thomas Moore est interprété par Iannicko N'Doua': played by or interpreted by – something more than just read or delivered, but still, we are reminded to distinguish the originals from these actors.

The podcast is prefaced by David Ridgen, who says, in French (with an English-speaking accent, unlike that of the French-speaking actor who later plays him), how delighted he is with this French version, which will reach a large new audience. This might suggest the producers are anticipating some doubt on the part of listeners about the quality of a translated version, its proximity to the original, which might be more real. Similarly, the presenter and writer of *Burning Sun*, Chloe Hadjimatheou, issues a warning at the beginning of each episode: 'One thing before we start: all the interviews and some of the clips in this podcast were in Korean. So the voices you'll hear are actors. But their words are real.' Despite their assurances, dubbed audio documentaries and even dubbed interview clips – though frequently used – are often disliked by programme-makers.

In series of articles on translation in radio and podcasting by Transom, Luis Trelles (2019) and Ann Hepperman (2014) reject the conventional voice-over translation as rude or disrespectful to the dubbed interviewee, reflecting perhaps the meaning it would have in face-to-face social interactions if one person spoke over another. If the person reading the translation uses a flatter intonation, to minimize the risk of being interpreted as an interviewee, this risks being construed as disinterest and even disrespect. Hepperman writes:

> Nothing irritates me more than what I call "The Duck and Cover." You know what I am talking about. This is the point in a radio story where you hear a person speaking her native language for two seconds, only to have her voice instantly replaced by an intern who was handed a script and told, "Here, read this. You're a young woman from Senegal. Go." In reality, I'm sure the execution is not so thoughtless, but "The Duck and Cover" just sounds so dehumanizing to me.

In these Transom reflections, dubbing is also framed as a failure of flow. Trelles says he hates voice overs for 'sounding clunky and hopelessly out of date, (like listening to a BBC radio piece from the 50s)' Ironically, the BBC put a lot of effort into their scripting of interviews and dramatised revoicing of original recordings at

this time, in part so they would flow smoothly (see, eg. McDonald 2020) but for Trelles, what feels dated is the awkwardness, a lack of smoothness or seamlessness, and Trelles prefers 'letting the natural flow of the conversation run its course'. Here then, the potential meanings of dubbing as disrupting social interactions meet expectations of audio listening as smooth and pleasurable, where the mediation itself does not intrude. Eleanor McDowell, in her (2017) essay on translation voices as a creative tool, quotes Danish producer Rikke Houd, who says she thinks of a story as a universe to itself 'surrounded by a thin, almost invisible membrane, a bubble' that she fears breaking with 'a clumsy, routine translation'. Such clumsiness:

> has ripped me out of many wonderful stories. One moment I am in a place somewhere, in a situation, let's say with a goat shepherd in Afghanistan. …. But then — bang! … some totally unrelated voice recorded in a studio back home pops into the story…. The magic just disappears.

The magic here is surely the illusion of non-mediation, which permits a sense that an audio story carries us to another time, another space, or another person's reality (Wincott, 2023). This ideal of listening as being transported goes back to the earliest development of audio documentary (Madsen, 2013) and has continued to shape judgements about the aesthetics of audio storytelling (Lacey, 2013). Nonetheless, many producers of serious audio work relish techniques that draw attention to fragmentary, difficult, or awkward transitions and the material reality of their own production, like brutalist architects putting the pipes on the outside of buildings. Indeed, McDowell admires Rikke Houd's documentary which uses 'an inventive audio language which cracks and fragments — English, Persian and Norwegian jaggedly interrupting each other.'

Perhaps then it is also a distaste for non-diegetic voices, a distaste that is not limited to translations. Studio-recorded narration has been the subject of debate since the early days of documentary. In documentary film in particular, explanatory speech can be seen as interfering with the artistry of the image, and aspiring filmmakers are encouraged to show not tell (Kozloff, 1988, p. 13). It is harder to do this in sound, but the European documentary-feature tradition is characterised by a frequent adoption of observational mode, and the use of little to no narration, a discipline contrasted at times unfavourably with the North American storytelling style, which Siobhan McHugh (2016) has called '"talkier" and less crafted' (See also the conversation between Alan Hall and Rob Rosenthal for *Sound School* podcast [Rosenthal, 2023]). Studio-recorded

translating voices are not narrators, but they prioritise the informational content of speech over other aspects of sound. These might include the geographical, age or class information in the accent, the emotional state or health of the speaker, that might be heard in the voice, and if recorded on location, the other sounds of their environment. These other aspects of sound are not only contributors to a sense of the recording as more 'real', they also offer the listener other kinds of information to interpret, contributing to a richness and open-endedness of meaning of the work observed in the opening of *Someone Knows Something,* above.

It would be a mistake to see the studio-recorded dubbing voice as monolithic though and there are choices in technique. The voice is often recorded closer to the microphone than the original, as in *Quelqu'un Sait Quelque Chose*. The closer the voice to the microphone, the less ambient sound we capture, including the sound of footsteps, clothes or jewellery, of movement around the recording space, and the fewer acoustic clues as to the space in which this voice speaks, such as how large, how high, whether outdoors or in, the materials the walls and floor are made of. This absence of sonic cues as to the speaker's presence in a real and material space has the effect of disembodying the voice and removing it from 'real' space of the recorded scene, creating a distinction between the original space and the translated space.

But disembodiment in storytelling has other possible meanings. It is associated with centrality and authority, while lower status or more peripheral characters often have more markers of embodiment attributed to them (Punday 2003). Which markers of embodiment should be reproduced in the dubbing of an interviewee's voice, perhaps making audible dimensions such as socioeconomic background? If they use slang or dialect words that locate them in a geographical or age community, if they are speaking a second or third language haltingly, if they make grammatical mistakes or stumble over their words, and the dubbing voice reproduces these, that could be understood to shame or ridicule the interviewee. All the more so, perhaps, if they also hold back a little in their performance, to signal we should not confuse them with the original. At the same time, these qualities of speech can have meaning. They reflect the identity and experiences or the emotional state of the speaker, to lose them is to lose something of the content, as we have observed. 'Cleaning up' their speech also suggests judgement that shames the original speaker and recalls the debates around scripting and re-performing of regional and working class voices in the

early years of the BBC (Lewis 2021). There is a risk of a flattening of the rich storyworld of a documentary feature, if all voices speak in studio, close to the microphone. The questions of closeness to the microphone and of the availability of location cues like ambient noise and reflected sound are also important in what it suggests about closeness and distance to others. Høier (2012), writing about the placement of TV sports reporters, suggests placement of the voice can be spatial – it situates it within the space of the storyworld, for example on the ski slopes near the competitive skiiers – and also be subjective, implying one identifies more or less or aligns oneself more or less with story actors (eg the skiiers). So when the French-language versions of David Ridgen and his son speak into their studio microphones, and thus each close to the listener's ear, information about their relationship to each other and their environment is lost or overwritten.

Choices in spatial and subjective placement also imply a certain relationship between the speaking voice and the listening audience. If the dubbing voice is close to the microphone, it is close to the listener: an intimate form of address, which when the listener wears headphones may even appear to be inside their head (Wincott, Martin and Richards 2021), while voices more distant from the microphone appear more distant from the listener. *Congo: a River Journey* (BBC World Service, 2018) is produced in binaural[1] so its spatial distinctions are perhaps even more pronounced. Presenter Alastair Leithead talks to fellow Congo river boat passengers, like Chacoupiwe Honoriné (24'28), whose voice appears, barely distinguished from the background sounds of the busy vessel, while the translating voice, recorded in post-production, cuts in to speak to us from front and centre, close to the microphone. This has the effect of making the interviewee appear to recede even further. In the same programme, many interviewees speak English, such as the daughter of former President, Mobuto Sese Seko, Chief Asambia Kapwata Fifi as she and Leithead explore the ruins of the Presidential palace. We hear (19'25) two pairs of footsteps crunching on the rubble, the sound of leaves being pushed aside as they walk around what remains of the family home, and she begins sharing her memories of its heyday, As she is challenged by Leithead to explain her father's brutal reputation (21'25) their voices reverberate in the empty rooms, both marked by an equivalent level

[1] Binaural stereo is a format that, when listening on headphones gives a 3D or spatialised effect. https://www.bbc.co.uk/rd/projects/binaural-broadcasting

of embodiment, both a similar distance from the microphone (and therefore the listener), the interview remains a conversation to which the audience listens in, almost in observational mode. The conversation with Chacoupiwe Honoriné must have been interpreted on the scene too, and if that contemporaneous translation had been preserved, it could have changed the spatial relationship between listener and non English-speaking contributors.

In practice, however, even when the translation is recorded contemporaneously on-location, during the mix, the voice of the interviewee will still be dipped and the translation dubbed over the top. Leaving the whole bilingual exchange in the cut would take longer and as we have observed above, in news in particular, there is very little time to spend. Even in long-form work, long sections of speech in a second language might feel uncomfortable to the producer editing it and they may presume it to be unintelligible to the listener, in the monolingual media paradigm. Therefore leaving the bilingual exchange in in full is desirable using the criterion of truth or authenticity, maintains information from other aspects of sound than words, but is problematic according to three of the constraints we noted in the introduction: time, comprehensibility and cultural expectations.

And yet, the producers of German series *Voice Versa* (from Deutschlandfunk Kultur and the Goethe-Institut 2022-), have taken the decision not to dub non-German contributions and the programmes feature interviewees speaking Arabic, English, Bengali and other languages, as well as German, which according to the podcast are ‘treated equally’ / ‘gleichberechtigt auf Augenhöhe.’ This term ‘auf Augenhöhe’ literally means at eye height, with its implications of looking someone in the eye, showing a mix of closeness and respect, giving someone a fair hearing, even where there are differences of culture or values, and has become a favoured phrase of politicians in recent years. But what does it mean to employ a translation style that treats everyone ‘auf Augenhöhe’?

The series is a kind of magazine format, each episode featuring discreet stories, tied together by host Dominik Djialeu. It directly challenges the assumptions of monolingual German media, contrasting these in episode 1 to the globalised, multilingual spaces of modern Germany. In episode 1, Djialeu (1’56 – 2’26) says:

> Wenn ich das Radio einschalte ist von dieser sprachlichen Vielfalt allerdings oft nicht mehr viel übrig. Bevor wir Luft holen, werden Sprachen die fremd klingen overvoiced und direkt ins Deutsch übersetzt. Superschade denn so gehen superviele Emotionen einfach *lost in translation*. Stellt euch mal vor, es gibt einfach kein Overvoice. Stellt euch vor, wir behandeln Deutsch und, beispielsweise Türkisch, völlig gleichberechtigt auf Augenhöhe. Das machen wir mit diesem Podcast.

> When I switch on the radio, there's often not much left of this linguistic diversity. Before we can catch our breath, languages that sound foreign are dubbed over and directly translated into German. It's a real shame because loads of emotions get 'lost in translation'. Imagine there was simply no dubbing. Imagine we just treat German and, for example, Turkish, completely equally *auf Augenhöhe*. That's what we're doing with this podcast.

The programme goes deeper, to explore the politics of monolingualism that are grounded in colonisation. For example a two-part feature across episodes 1 and 2, centres on the shame and anger that writer Jacinta Nandi and journalist Mithu Sanyal feel at not speaking Bengali, the first language of both their fathers. Their families' decisions to raise them in a monolingual household (English in the case of Nandi and German in for Sanyal) are linked to the parents' and grandparents' shame of being marked out by speaking the language of the colonised, and their desire to fit in in Europe. In turn, their daughters are occasionally shamed by others' disapproval that they do not know Bengali. Colonial politics of language are introduced and complicated in a section on the history of East Bengal, separated from West Bengal at partition and given to the newly formed Pakistan. Pakistan imposed Urdu, and the Bengali language became a rallying point for an independence struggle that resulted in the creation of Bangladesh. All of this history colours the meaning of the language for the women.

Sanyal narrates the feature (in German), but we also listen in to conversations between the women, and between each woman and her father. Nandi often breaks into English to express herself, and the conversation with her father is entirely in English, as her father too has lost his Bengali. This offers a chance for the programme to practise its commitment to 'equal treatment' of languages. When Nandi relates a humorous saying (ep 1 20'12) about the character of Bengalis in English 'one Bengali is a poet, two Bengalis is a literary magazine, three Bengalis are a political party,' Sanyal (in narrating mode) responds 'Sogar ich kenne die Redewendung' (Even I know the saying), which she translates word for word into German.

Sometimes Nandi's shorter bursts of English – 'Divide and Rule!' (ep 1, 19'40) – are simply cut into the mix, unexplained and untranslated. Though often the German narration supports comprehension and understanding of the significance of this English speech, priming the listener with background and related statements, there is often no direct translation. For example the story of Jacinta's father (ep 2, 4'15) begins with German narration by Nandi that explains the two women had assumed their fathers perhaps didn't value them enough as daughters to speak to them in Bengali, and were surprised to discover many men had the same experience, including Jacinta's father. There follows an exchange in English between Nandy and her father, as he recalls how, after they moved to England when he was small, his parents stopped speaking Bengali to him, though they spoke it to each other. He says he can understand it but because he doesn't use it, he can't think of the words if he tries to speak. This is followed by a narration from Sanyal that paraphrases some of what he says in German in the third person, but not all of it, and adds more content about Nandi's father learning it was better not to speak Bengali on the street or to take Bengali food to school, and adds a comment not heard in the English 'Also begann Raj sein Bengali erst zu verstecken und dann zu vergessen' (So Raj began first to hide and then to forget his Bengali' ep 2, 5'36). The German narration is therefore not a translation of the interview, but a parallel contribution that nonetheless relates a complete story – a more complete story, in fact – even if the English is not understood.

This series is all about language and multilingualism, but we have noticed the same technique in use in other German programmes, such as the investigative serials *Scambit: Schach, Hype und Millionen* (ACB Stories/WDR, 2023) about cheating in competitive chess*, Slahi: 14 Jahre Guantanamo'* (NDR Info, 2021, about a Guantanamo detainee) and *Call me Günther* (Deutschlandradio, 2023, the same organisation as *Voice Versa*), about investment fraud operating out of Kosovo. In *Scambit,* experts are interviewed in English and the exchanges not dubbed. Here too, the narration helps to support understanding, translate key content, though, like in Voice Versa, not always word for word. It also adds additional information, not present in the English. In *Slahi: 14 Jahre Guantanamo*, episode 4, an interview in French is partly left intact, and some parts are dubbed into German. In *Call me Günther,* sections of English-language interviews are left in undubbed. However, not all languages are treated in the same way. In *Call me* Günther, interviewees speaking Albanian are always dubbed.

This might be to do with geographical and class expectations about the audience: anticipated to be German, or from other German-speaking European areas; able to understand some spoken English (which is taught in school and is a useful lingua franca for professionals and businesspeople who communicate internationally). Albanian is assumed to be outside their understanding. Even in the streaming age, when everyone, everywhere can listen to everything, this is a reasonable working assumption for a German-language broadcast or podcast commission. However language status is not only practical, but social and cultural and speaking particular languages is linked to issues of power and belonging, something articulated so well in *Voice Versa*. We suggest that a German presenter will be proud of speaking English, perhaps even feel some pressure to demonstrate they can interview in English. Maybe not dubbing over the English flatters the German-speaking listener. Many English words pepper German-medium podcasts – even the titles of *Scambit* (Hype) and *Call me Günther,* and the English-language pun in Voice Versa (a play on Vice Versa). Albanian does not have this status, while French, another high-status world language, seems to, to a lesser degree. We have not seen these techniques used in English-medium programmes, reflecting we assume an assumption that the main programme language is itself of a status that cannot be bettered, and that listeners are not expected to speak any particular other language, as English is also the dominant global lingua franca.

Conclusion

When they decide how to manage content in more than one language, programme-makers work within the constraints we set out in the introduction. The time taken up by the technique matters, comprehensibility for the audience is of course key, but fidelity of translation matters too, and that might include the meaning that listeners can read in distance, tone of voice and other non-verbal aspects of communication. Style or format expectations are also borne in mind. The wider range of techniques in use that prompted us to write this article might have a lot to do with a loosening of this last constraint in the era of podcasting. We speculate that podcasting has brought a new willingness to experiment with form, as well as an expansion of the time available to tell stories in long form and serial podcasts. Something has changed too perhaps in our expectations about the national and linguistic belonging of audio content. In the internet age, a documentary or serial that might have once had an audience of a few thousand

on one FM radio broadcast, in a single nation, might now gradually accumulate a far larger audience internationally. Audio documentary recommendations, festival line-ups and prize shortlists are shared on social media and indirectly different styles and approaches are likely to spread across the old national broadcast borders, as we saw perhaps most remarkably with the influence of *Serial*, and the way the tone and structure of American public radio now echoes in other places and other languages.

New norms or habits for managing the inclusion of more than one language may become fixed in future, but at the moment we see a much more fluid situation than existed in the radio age. Some of the output being produced in Germany could almost be considered multi-lingual, rather than translated, especially in longform documentary and serials, where the constraint of time is the least strict. However time and pacing will always matter. *Voice Versa* is a slow, rather intellectual feature centred on the politics of language, but *Scambit* for example is produced by public service content producers funk, whose content for 19 to 24-year-olds is known for its fast pace and self-consciously irreverent style. Its in-depth reporting on long-running rows about cheating and allegations of cheating in chess has to convey a lot of information and maintain narrative coherence over several episodes, yet both manage to incorporate multi-lingual listening without compromising their pace and tone. To do this they use a range of techniques that do not have a terminology to describe them, as far as we are aware. We identify these as follows and suggest terms in English and in German – the languages of the authors.

Some new terms for translation practices in audio documentary

We are mindful of the fact that these suggested terms might not find traction among radio and podcast producers but offer them as a typology that distinguishes some of the options available through a description of their different characteristics and uses, opening them up for discussion.

- Play the other language in full, then translate it pretty much word for word in full. We might call that **side-by-side translation** in English or perhaps **full translation** – in German, **Voll-Übersetzung**.
- The narrator doesn’t appear to translate directly at all, and instead offers commentary or notes that support understanding before and after the interview clip in an ‘other’ language. This treatment sounds like a normal

narrator's link, but actually contains a condensed translation of the factual information from the clip and adds extra context or interpretation. This could be called **narrative commentary / nacherzählter Kommentar** or **narrative translation / nacherzählte Übersetzung.**

- Very short snippets of the 'other' language are inter-cut or interwoven with the main language ('Divide and Rule!') and remain untranslated. They may seem decorative, but crucially the meaning of the words in the 'other' language also needs to have real relevance and meaning, for those who can understand them. This technique is a feature of a storytelling or narrating style, in contrast perhaps to a straight-forward presenting or reporter style, where the presenter simply presents interviewees to the audience. We suggest here **interwoven** or **dynamic translation / dynamische Übersetzung** The interweaving of the two languages allows for the maintenance of pacing or narrative rhythm, by only using brief clips, and not pausing to translate. However its narrative brevity belies the painstaking planning involved, as Sven found when he used this technique in the science series *Quarks Storys* (WDR, 2019 – 2023), where it took time and practice to design this interweaving of another language into the narration while maintaining the rhythm of narrative speech.
- When an entire series is translated as with *Quelqu'un sait quelque chose / Someone knows something* or most of its content is translated, as in *Intrigue: Burning Sun*, particular actors consistently voice the roles of individual interviewees, reporters or presenters. As we have noted, these are far from the colleague in the office, grabbed to provide a last-minute voice-over. They perform, conveying emotional tone of the original speech, which is largely or completely absent. We argue therefore that this deserves a term that distinguishes it from the more common 'voice-over dubbing,' used in news and documentary. Interpretation, the word used in the credits for *Quelqu'un Sait Quelque Chose* is already in used to describe simultaneous translation by interpreters, so something like **performance dubbing** or **acted dubbing** might describe this approach. In German we suggest '**schauspielerisches Overvoice**' or '**nachgespieltes Overvoice**'.

Choices about when and how to incorporate more than one language in audio storytelling are not only shaped by practical constraints, but powerful cultural forces. Decisions are being taken within the dominant monolingual media paradigm, so all-pervasive it becomes hard to see it, and even harder to imagine an alternative, despite the multi-lingual complexity of so many parts of the world. Choices about when and how to incorporate other languages tell us something about assumptions regarding audiences and perhaps how they wish to be addressed – the languages they speak, their level of education, their cosmopolitan-ness, and it tells us even more about which languages have cosmopolitan prestige and which are held in less regard as merely 'foreign' or 'immigrant'. In *Call Me Günther*, the country of Kosovo is associated with large-scale, organised scamming of German citizens and its language, Albanian, is a language of recent immigrants in Germany. Albanian is voice-over dubbed, but French and English are treated with side-by-side translation and narrative commentary.

Programme-makers' sense of what works and what does not are also often influenced by established values of authenticity and immediacy in non-fiction radio and podcasting. This includes the question of which non-diegetic programme elements we are used to in any time, medium and genre. Revoiced interviews were accepted in the mid-twentieth century (though not without their critics [Lewis 2021]), and the complete replacement of an original quote with its translation remains standard in print journalism (Kalantari 2022). Yet early cultural expectations of radio – that it is a portal to reality (Madsen 2013), more immediate and more live than print – have perhaps influenced the longstanding use of 'duck and cover', which is both economical with time and preserves evidence of the original, and a hesitancy about complete replacement – Intrigue: Burning Sun remains unusual in this respect.

This isn't simply a matter of taste either. Full voice-over dubbing and acted dubbing create a greater distance between listener and interviewee. As their own real voice is masked, we are unable to judge the fidelity of the dubbed account of their words. The smoothness of professional actors' performances in particular turns our attention away from aspects we might otherwise pay attention to: the sounds the mouth makes when someone is old and frail, or nervous; the pauses, the melody, rhythm, pace and so on, that can communicate something to us about the speaker and help us evaluate their words. Documentary and on-the-scene reportage are valued as media forms partly for this extra-linguistic information,

for showing as well as telling. They are often forms that ask the listener to interpret and engage, and the habitual use of voice-over dubbing has failed to live up to this expectation. Even fast-paced 'storytelling' style documentary making, which tends towards the provision of plentiful information, and avoids semantic openness, can still weave in original voices using narrative commentary and interwoven translation. Therefore, we argue, news could conceivably do so too.

We have not considered the role of AI dubbing, which might in the future be used widely in audio journalism and documentary-making. At the time of writing, automated dubbing still cannot reproduce emotional inflection in the voice (Langer, 2023) if it could reproduce the non-verbal aspects of communication discussed above, the questions of closeness, fidelity and authenticity will have to be revisited. Another nascent technical development is object-based audio, being developed and tentatively rolled out by some European broadcasters.[2] It makes it possible to transmit different language versions of the same content at the same time, and for the user to choose their listening language. French international broadcaster, RFI is experimenting with this for their French-teaching output, for example. This technology might allow listeners to slide between dubbed and original voices in documentary.

Returning to the present, we note that in the two global, colonial languages we have looked at, French and English, producers are not experimenting beyond the two forms of dubbing we have described – basic voice-over and performance translation, and their experimentation seems to be limited to expanding the amount of story time they are willing to apply this to. German programme-making in contrast offers some alternative techniques that open up new possibilities. If reporters and documentary makers in the large media organisations of the Global North were to deploy these in their international reporting, large audiences could hear voices that have conventionally been dubbed over. But it could be especially interesting in the most linguistically diverse places of the Global South like Nigeria or Brazil – enabling storytelling at a national level that better represents the rich linguistic diversity of human culture.

---

[2] RFI, BBC, Radio France and BR for example collaborated on the EU-funded research programme, Orpheus www.orpheus-audio.eu/index.html. Object-based audio also has applications for accessibility, as users can set the speech to background sound ratios to suit hearing impairment.

Author details

Abigail Wincott worked as a BBC radio and podcast producer for 15 years and is currently Associate Professor of Audio Journalism at Falmouth University in the UK. Her research explores all forms of factual audio storytelling including news, documentary and heritage media, with a special interest in what happens when new media technology or techniques meet existing professional cultures and contexts.

Sven Preger has worked for 20 years as a documentary maker, podcast producer and story consultant. He specialises in complex narrative audio forms. He has written the textbook "Storytelling in Radio and Podcast". Sven is head of the podcast department at NZZ in Zurich.

References

ACB Stories/WDR (2023) ‘Scambit: Schach, Hype und Millionen’, for funk. Available at https://www.ardaudiothek.de/sendung/scambit-schach-hype-und-millionen/12493605/

Baumann, G., Gillespie, M. & Sreberny, M.A. (2011) Transcultural journalism and the politics of translation: Interrogating the BBC World Service. *Journalism* 12(2), 135–142.

BBC World Service. (2018). *Congo: A River Journey*. Produced by Becky Lipscombe https://www.bbc.co.uk/programmes/p06w66lm

Delisle J., Lee-Jahnke H. & Cormier M.C. (1999) *Terminologie De La Traduction. Translation. Terminology. Terminología De La Traducción. Terminologie Der Übersetzung.* Amsterdam: John Benjamins

Deutschlandfunk Kultur/Goethe-Institut (2022-) *Voice* Versa episode 1 https://www.hoerspielundfeature.de/zwei-sprachen-eine-story-voice-versa-folge-1-wer-hat-die-100.html

Deutschlandradio (2023) *„Call me Günther“: Milliardenschwere Betrugsmasche Cybertrading.* https://www.deutschlandradio.de/podcastserie-call-me-guenther-100.html

Eberhard, D. M., Simons, G. F. & Fennig, C. D. (Eds.). 2023. ‘Nigeria’ and ‘Brazil’ In *Ethnologue: Languages of the World*. Twenty-sixth edition. Dallas, USA: SIL International. Online version: http://www.ethnologue.com.Nigeria Ethnologue https://www.ethnologue.com/country/NG/

GWI. (2022) Share of users who watch content in other languages and prefer subtitling or dubbing as of November 2021, by country [Graph]. Statista. Retrieved October 26, 2023, from https://www.statista.com/statistics/1289864/subtitles-dubbing-audience-preference-by-country/

Hadjimatheou, C. & Puri, K. (2023) *Intrigue: Burning Sun*. BBC. https://www.bbc.co.uk/programmes/m001n9kv

Hepperman, A. (2014) ‘Thoughts on Translation: Ann Heppermann’. Transom https://transom.org/2014/thoughts-on-translation-ann-heppermann/

Hill, A. (2007) *Restyling Factual TV: Audiences and News, Documentary and Reality Genres.* London: Routledge.

Høier, S. (2012). The relevance of point of audition in television sound: Rethinking a problematic term. *Journal of Sonic Studies*, 3. Available at https://www.researchcatalogue.net/view/252390/252391

Johnson, D. (2020). ‘The Fix Is In: Dubbing as Transcultural and Transmedia Adaptation’. *Journal of film and video 72*(3-4), 3–15.

Kalantari, E. (2022). Journalistic Translation: A Gate at Which Journalism Studies and Translation Studies Meet. *Journalism* 23(7), 1411–29.

Koolstra, C.M, Peeters, A.L & Spinhof, H. (2002). ‘The Pros and Cons of Dubbing and Subtitling’. *European Journal of Communication 17*(3), 325–54.

Kozloff, S. (1988). *Invisible Storytellers : Voice-over Narration in American Fiction Film*. Berkeley, USA: University of California Press.

Lacey, K. (2013). *Listening publics: The politics and experience of listening in the media* Age. Polity.

Langer, B. (2023) Can you trust AI video translation? Our linguists put it to the test, 27 September. The Global Marketer https://veracontent.com/mix/ai-video-translation-tool/

Lewis, P. M. (2021) “A Claim to Be Heard”: Voices of Ordinary People in BBC Radio Features: Les Voix des gens ordinaires à la radio de la BBC – expériences 1930-2000. *Revue francaise de civilisation britannique*, Xxvi (1). doi.org/10.4000/rfcb.7681

McDonald, K. (2020) Scripting the Radio Interview: Performing Desert Island Discs. *Radio Journal* 18(2), 175–92

Madsen, V. (2013) ‘Your ears are a portal to another world’: The new radio documentary imagination and the digital domain. In J. Loviglio & M. Hilmes (Eds.), *Radio’s new wave: Global sound in the digital* era (126–144). Routledge.

McDowall, E. (2017) Thoughts On Translation: Eleanor McDowall. Transom https://transom.org/2017/thoughts-translation-eleanor-mcdowall/

McHugh, S. (2016) ‘How Podcasting Is Changing the Audio Storytelling Genre’. *Radio Journal 14*(1), 65–82.

Mehdizadkhani, M. & Khoshsaligheh, M. (2021). ‘Insertion or Voice-Off in Rendition of Graphic Codes: An Experiment in Persian Dubbing’. *Visual communication 20*(1), 81–99.

NDR Info (2021) *Slahi: 14 Jahre Guantanamo*. Written and produced by Bastian Berbner und Ole Pflüger. https://www.ndr.de/nachrichten/info/podcast5156.html

Punday, D. (2003) *Narrative Bodies: Towards a corporeal narratology*. New York; Basingstoke: Palgrave Macmillan.

Ridgen, D. (2017) *Someone Knows Something*: Season 3 CBC.

Ridgen, D. (2022) *Quelqu’un Sait Quelque Chose*. CBC. https://ici.radio-canada.ca/ohdio/balados/9730/quelquun-sait-quelque-chose

Rosenthal, R. (2023) ‘Lean In and Listen Like It’s Music’. *Sound School* podcast. https://transom.org/2023/lean-in-and-listen-like-its-music/

Schulze, F. (2020). *Vom Sklavenhandel zum Auswanderungsland – Brasilien und die Migration.* Bundeszentrale für Politische Bildung. https://www.bpb.de/themen/migration-integration/laenderprofile/suedamerika/317598/vom-sklavenhandel-zum-auswanderungsland-brasilien-und-die-migration/

Trelles, L. (2019) ‘Thoughts on Translation: Luis Trelles’ Transom. https://transom.org/2019/thoughts-translation-luis-trelles/

WDR (2019-2023) *Quarks Storys* https://www.quarks.de/storyquarks/

Wincott, A. (2023) ‘Un bon voyage sonore: Avoiding listener discomfort in immersive audio documentary’. *Journal of Radio and Audio Media*. https://www.tandfonline.com/doi/full/10.1080/19376529.2023.2225062

Wincott, A., Martin, J. & Richards, I. (2021) Telling stories in soundspace: placement, embodiment and authority in immersive audio journalism. *Radio Journal: International Studies in Broadcast & Audio Media 19*(2). https://doi.org/10.1386/rjao_00048_1

Yildiz, Y. (2013). *Beyond the Mother Tongue : The Postmonolingual Condition*. New York, USA: Fordham University Press.