

**Listen! The human voice as a neglected psycho-social research topic and resource<sup>1</sup>**

**ANNE KARPf<sup>2</sup>**

**Abstract**

There is an excellent fit between the speaking voice and the psycho-social approach, in that the voice connects inner and outer worlds while simultaneously challenging such a division. It remains, however, relatively neglected, both as a psycho-social research resource and as a topic for the psycho-social researcher. This article argues that, while researchers are developing increasingly sophisticated ways of harnessing visual research methods, the oral dimension remains marginalised, with voice almost invariably collapsed into speech. Despite the methodological challenges created by using the voice as a psycho-social research tool, attention to the paralinguistic has the potential to enrich research and deepen our psycho-social understanding of human behaviour.

The speaking voice is the nodal point of human communication: notwithstanding the growth of social media and virtual bonds, it remains the prime instrument through which we establish relations with others, acting as connective tissue among both strangers and intimates. Given this centrality, the neglect of the speaking voice as a psycho-social research topic and resource is curious, especially in the context of the enormous interest in conversation and speech over the past few decades. Yet where voice is referred to at all it is often its metaphorical meanings that are signalled: the voice as signifier for political presence and power, a synonym for enfranchisement. Literary studies have also appropriated it, along with other terms from the auditory lexicon such as tone and register, to signal narrative viewpoint or perspective. Voice

---

<sup>1</sup> This article is based on a paper given at the Association of Psycho-Social Research conference at the University of West of England, Bristol, June 2016, and Karpf's PhD submission, *The persistence of the oral: on the enduring importance of the human voice* (2016)

<sup>2</sup> Contact: A.Karpf@londonmet.ac.uk

is often also elided with dialect or accent, especially in the UK, or it is used as a proxy for speech and language.

Again and again major thinkers sidle up to the voice, seem about to grasp it - and then turn on their heels. Bourdieu (2014), for example, came tantalisingly close to describing how social and cultural factors are 'implicated' in the voice but never quite spelt it out. Yet his notion of 'doxa', the relationship between habitus and 'the field to which it is attuned', conceives of the body as a 'living memory pad' (Bourdieu, 2014, p.68), producing 'body automatisms' that are below the level of consciousness, including bodily expressions of emotion such as laughter, along with 'deep-rooted linguistic and muscular patterns of behaviour' (Bourdieu, 2014, p.69). The human voice is located precisely at the nexus of the linguistic and the muscular.

Or there is Ahmed (2004), who makes many illuminating points about signs and how they work on and in relation to bodies through their 'stickiness'. Hate speech becomes sticky, she argues, through repetition, e.g. of the word 'Paki', which accumulates a history to which it is bound. She suggests, further, that naming something as disgusting in a speech act is performative; thereby, the sign itself becomes 'sticky' and fetishistic. Yet, remarkably, the fact that we learn to recognise these signs in part through intonation is relegated to a footnote, even though it is surely the oral register through which it is shouted in hate speech that is the most 'sticky' aspect of the word 'Paki', and disgust is performed at least as much through the vocalisations that accompany words such as 'That's disgusting!' or, more succinctly, 'Yuk!' Strangest of all, Wetherell (2012), in her pioneering exploration of the circulation of affect, alights on vocal states in some of her case-histories, especially in her analysis of the conflict between a pair of teenage girls, but never dwells on their vocality. She argues for the multimodality of situated affective practices, including bodily actions and storytelling, that together create 'an integrated

and organic unfolding and weaving' (p.89). Is there not space here for the modality of voice? The purr, the gabble and the whine, are they not constitutive of meaning, a symphony of signs that draw on cultural repertoires of shared understandings that they also help construct? This article suggests that the speaking voice is a rich research topic for psycho-social researchers.

It is also a potentially valuable research tool, even though as a resource too it has been neglected. Over the past decades visual methods, from photographs and documentary film to video diaries, have been increasingly embraced by qualitative researchers and yet attention to the oral and aural has scarcely developed. An innovative research method, the 'visual matrix', currently being developed, for instance, explores ways of using 'imagery, affect and visualisation... to inquire into phenomena that research participants may find difficult to put into words' (Visual Matrix Workshop, 2015). This has been deployed interestingly to help gauge reaction to local public art (Froggett, 2014). Yet although the final report of the project refers to the 'emotional tone, pace and vibrancy' (p.61) of participants' reactions, it is (understandably perhaps, given its remit) mostly mute about the vocal dimension. Plummer (2001) devotes three lines to sound archives in his catalogue of 'documents for life' compared with a page-and-a-half to visual data.

### **Reasons for the marginalisation of the voice**

A number of different factors are responsible for the discursive absence of the speaking voice. While ideas around embodiment have provided a supportive milieu in which to develop thinking about the voice, they do not provide a perfect fit. Just as the concept of 'nonverbal behaviour' has proved to be a valuable conceptual home for research into the voice yet also one with limitations (since, pace Bourdieu, and leaving aside bodily sounds such as sighs and sobs,

grunts and cries<sup>3</sup>, it is hard to conceive of voice without words), so too does embodiment take us only so far. For the paradox faced by both researchers, interviewees and indeed speakers themselves is that the voice is the product of the solid materiality of the body and yet, as soon as it has been produced, becomes fleeting, temporary and insubstantial - 'materiality at its most intangible' (Dolar, 2006, p.59). 'That which is only audible', wrote Simmel, an early theorist of the senses, 'is already past in the moment of its present and provides no 'property'' (Simmel, 1907). The voice, as Dolar puts it, 'is like a bodily missile which separates itself from the body.... the voice is *plus-de-corps*: both the surplus of the body, a bodily excess, and the no-more-body, the end of the corporeal' (Dolar, 2006, pp. 70-71). In addition, the increasing dissemination of voice through electronic media, along with the emergence of the synthetic voice, problematises the identification of the voice with the body.

The auditory field, at least in modern Western cultures, is thus harder to analyse than the visual field and gesture has proved easier to describe than voice. (Interestingly, spatial metaphors are often deployed to elucidate the paradoxical nature of the voice. As with Dolar, above, so LaBelle conceptualises the voice 'as something expelled from the mouth, but which never leaves me behind... The voice does not move away from my body - the voice stretches me; it drags me along' (LaBelle, 2014, p.5.) This may help explain the difficulty of researchers in retaining a focus on the audible voice. Indeed, a great deal of the interest in the embodied voice arose just at the point where, thanks to new communications technologies, it was becoming disembodied.

Perhaps we should be looking elsewhere for a sensitivity to vocal communication, for instance in the proliferation since the early 1990s of scholarly literature in the field of sound studies.

---

<sup>3</sup> Which have recently acquired their own biographers, e.g Connor, 2014; LaBelle, 2014; Gillie, 2010

One of the earliest and most influential texts was Murray Schafer's exploration of the 'soundscape', an elegy for and a celebration of a lost, rich sonic world, first published in 1977. Murray Schafer argued that the soundscape was changing, with humans now inhabiting an acoustic environment different from any known before and one in which they suffered from the effects of noise pollution (Murray Schafer, 1994). In response Schafer originated the concept of 'acoustic ecology', a deep appreciation and understanding of the relationship between human, animals and the sonic environment (Cummings, 2001).

The studies that followed explored sonic culture in all its variety, amounting to 'a cultural phenomenology of mediated aural practices' (Droumeva and Andrisani, 2011). Coming at the same time as a boom in interest in both oral and aural cultural forms - from the audio book and podcast to the soundwalk and sound art - it might be surmised that we are living in a golden age of interest in sound.

It is striking, however, that the human speaking voice occupies a relatively insignificant position in much of the research in sound studies and mediated aurality. This may be because their interest lies in aurality in its widest definition. It might also signal an attempt to de-privilege the human voice as the pre-eminent sound of sonic culture and set it instead as one element in the acoustic environment - a kind of rebalancing away from human hubris that is integral to the acoustic ecology project, and an implicit repudiation of the logo-centrism of speech. So, while sound studies have played an important role in drawing attention to the importance of sound as a medium and away from the notion of it as somehow unmediated (Bruhn Jensen, 2006), research that identifies itself as part of sound studies nevertheless focuses mostly on the sounds that human beings receive, rather than those that we make.

Interestingly, there is no analogous discipline of voice studies. Insofar as this exists (and the term is only just coming into usage), it is organised more around either physiological, neurological or biomedical matters, such as voice quality and voice perception (Kreiman and Sidtis, 2013), or the vocal arts (voice coaching, the voice in theatrical performance). What is needed is a counter-tendency to the dominant approaches of both sound studies and voice studies: to place the human voice as instrument squarely at the centre of analysis and debate, but to view it through the prism of culture, and in its applied role in situated speaking practices, rather than as part of a theoretical examination of the metaphysics of the acoustic field or the ontology of the auditory. Instead, to give voice to voice.

### **The hegemony of the eye**

The privileging of sight has been traced back to Platonic philosophy. Logos, after all, originally included the spoken word. Cavarero (2005) claims that the 'devocalization of logos' was inaugurated by Plato.<sup>4</sup> Greek philosophy 'refuses to concede to the vocal any value that would be independent of the semantic' (p.35). By reducing the voice to the acoustic signifier of a semantic signified, 'the sense of the voice is entirely bound up with the role of vocalizing concepts, so that whatever is left over is an insignificant remain, and excess that is disturbingly close to animality' (p.34). Another reading dates the hegemony of the visual back to the Enlightenment (itself a visual metaphor), which is held to have accelerated the visualisation of knowledge.

Still today metaphors of sight provide us with a lexicon of terms denoting understanding (I see, perspective, perceptive, focus, insight, observation, point-of-view), thereby impoverishing our

---

<sup>4</sup> Ong (1982), however, argues that Plato's relationship to orality was ambiguous.

sense of the sonic and conjoining knowledge with vision. (Jay, 1994, managed to pack 21 visual metaphors into his opening paragraph alone.) The existence of an extensive lexicon of sight need not, of itself, inevitably impoverish our sense of the sonic: it is possible to imagine a parallel rich lexicon of sound and voice existing alongside it. In practice, however, ocularcentrism has been aggrandising and had a colonising effect on both sonic conceptualisation and language. Ihde (2007) claims that the 'reduction to vision' led to a 'reduction of vision': 'Not only are sounds, in the metaphysical tradition, secondary, but the inattention to the sounding of things has led to the gradual loss of understanding whole ranges of phenomena that are there to be noted' (Ihde, 2007, p.13).

This ocularcentrist reduction may help to explain the discursive absence of the voice but it has also been contested. Schmidt (2003) has challenged the 'sprawling discourse' about hearing's modern diminution, arguing that the Enlightenment advanced not only optics but also acoustics, while Sterne (2003) has given a subtle account of the continuing, but altered, status of the patient's voice in 19th century medicine. Jay (1994) demonstrates persuasively how twentieth century French thinkers challenged ocularcentrism and denigrated sight, with the invention of the camera contributing to the undermining of its privileged status and a 'frenzy of the anti-visual' (p.187). Together they suggest that visualism, perhaps, has been overstated.

More disturbingly, there has been a slippage, from the idea of ocularcentrism as a dominant discourse that marginalises the aural, to a view of it as an accurate description of the modern sensorium. This slippage seems to propose a historical shift not just in our understanding of the senses but in their actual use: as though sight were not only *conceived of* as being superior but over time has *become* our pre-eminent sense. There is a difference, however, between the ranking of importance of senses - the discourses around vision and audition - and the ways in

which they are exercised in daily life. What was lost through ocularcentrism and the shift from oral cultures to literate ones was not, perhaps, so much the importance of the voice but a recognition of its importance.

## **Derrida**

On the face of it (another visual metaphor - they are hard to escape), Derrida's concept of phonocentrism (1997) appears to be the obverse of the claim that the voice is marginalised. Derrida contends, on the contrary, that phonocentrism (which he links with logocentrism) privileges voice, championing speech over writing. Derrida's argument takes the form of a dialogue with, and critique of, Husserl, Saussure and Levi-Strauss. He challenges Levi-Strauss's paeon to the orality of the Nambikwara 'who did not know violence before writing; nor hierarchization' (Derrida, 1997, p.135) and Rousseau's hymn to intonation, to 'tone, stress and memory' in the pre-linguistic (p.247), to

'tone of voices that cannot fail to be heard, penetrate to the very depths of the heart, carrying there the emotions they may wring from us, forcing us in spite of ourselves to feel what we hear. We conclude that while visible signs can render a more exact imitation, sounds more effectively arouse interest' (p. 240)

which, both explicitly and by implication, disparage writing in comparison.

Derrida is having none of this vocal pre-lapsarianism. He contests the notion of speech as uncontaminated by writing (Naas, 2011) - speech, as it were, with direct access to 'the intimacy of self-presence' (Derrida, 1997, p.17), 'deluded with believing itself completely alive' (p.39).

In fact, Derrida's argument in *Of Grammatology* in several ways supports the claim that the voice has been marginalised. Firstly, although Rousseau may have been talking about the embodied voice, Derrida was mostly referring to spoken language. In this sense *Of Grammatology* could be considered another example of the elision of language and voice and of the reduction of voice to speech. In the speech wars between Derrida and Saussure, the voice was mostly a bystander.<sup>5</sup>

Secondly, insofar as Derrida actually had the voice in mind, he was trying to avoid its idealisation as a source of authenticity. Derrida mocked the idea of the purity of speech as something unmediated, with a privileged 'proximity to mind' (Claude Evans, 1991, p.150). All signs, whether oral or written, are mediated; none can be ahistorical, and yet many writers who have recognised the importance of the speaking voice have also polarised sight and sound, creating an opposition between them and, at worst, adjudicating between them - a taking of sides. In precisely the way that Derrida anticipated, many of them idealise the voice.

Vision, claims van Leeuwen (1999), evidently never having looked at a Rothko painting, 'cannot go beyond the surface of things' (p.196), arguing that it isolates and individuates. While Berendt (1988) makes telling observations about the differences between the ear and the eye in the manner of Ong, he is also prone to a crude denigration of 'eye culture', which he links to patriarchy. Similarly, he is fascinating on the heightened role played by sound in Eastern spiritual traditions such as Sufism but then proceeds to make exaggerated claims on the ear's behalf ('the ear establishes a more 'correct' relationship between ourselves and others', (p.28);

---

<sup>5</sup> Cavarero (2005), in a dense but brilliant appendix on Derrida, suggests that Derrida simultaneously opens up the voice and imprisons it, and that the voice in Derrida is a voice of thought, 'totally insonorous' (p.220).

‘eye people taking part in discussions display strikingly more aggressiveness than ear-orientated-participants’, (p.30)).

Most prominently, McLuhan's idealisation of the ear connected it to a kind of noble savage primitivism (‘It is quite obvious that most civilised people are crude and numb in their perceptions, compared with the hyperesthesia or oral and auditory cultures. For the eye has none of the delicacy of the ear’ (McLuhan, 2011, pp. 30-31)). The notion of a pure vocality somehow beyond culture, or indeed that of a single 'true' voice, is essentialist in the extreme. It is surely time to call a truce between the propagandists of the ear and the eye: we need both and human societies at their best demonstrate the capacity to integrate them.

### **Discursive absence and tacit knowledge**

The discursive absence of the voice gives rise to a significant methodological problem. Karpf (2006) conducted 50 open-ended, semi-structured interviews in the UK and USA in which interviewees were asked about their feelings about both their own voice and that of family members, friends and colleagues, in order to construct a sense of their vocal world. Yet, given the discursive absence already discussed, how could they achieve this? If contemporary Western cultures have no shared public language through which to discuss the voice, what could the interviews elicit save confirmation of this thesis? Macfarlane (2015) maintains that ‘language deficit leads to attention deficit’. Charting the impoverishment of language to describe place, he suggests that ‘As we deplete our ability to denote and figure particular aspects of our places, so our competence for understanding and imagining possible relationships with non-human nature is correspondingly depleted’ (p.3).

The human voice is such a taken-for-granted, inescapable aspect of communication that it needed to be in some sense problematised before most of those interviewed by Karpf (2016) felt able to expatiate about it. A number of informants asked for cues or prompts in order to help them begin to formulate answers to the interview questions: they spoke more fluently when asked to comment about the use of their own voice or those of other people on specific occasions or in particular relationships or periods of time so that the voice became situated, actualised and embodied rather than an abstract instrument. Interviews in such a context need to be conducted with particular care so as to ensure that the interviewer's language and assumptions are imported or imposed on interviewees to the least possible extent. Ihde (2007) notes how a student beginning a phenomenology of sound comes upon 'phenomena he has not previously noted and does not yet have or call to mind the words with which to describe such experiences' (p.86). Karpf (2006) reported on one informant who declared that she had never given the subject of voice a moment's thought and then proceeded to regale the interviewer with a lengthy disquisition on the various effects that different voices had upon her. There was a sense that some informants were speaking strongly-held views about voices out loud for the first time.

Polanyi (1967), who coined the phrase 'tacit knowledge', argued that 'we can know more than we can tell' (p.4). Tacit knowledge thus resembles what Christopher Bollas called the 'unthought known' (Bollas, 1987). Polanyi conceptualised tacit knowledge as a matter of attention. In tacit knowledge, he suggested, we attend *away from* one dimension, the proximal, and attend *towards* something else, the distal. Such knowledge is frequently bodily. 'Every time we make sense of the world we rely on our tacit knowledge of the impacts that the world makes on our body and of the response of our body to those impacts' (Polanyi, 1964, p.20). The task of speaking about their voices required in interviewees a gestalt switch, an attention

away from a concentration on language and *towards* prosody and the paralinguistic features of voice, and that they mobilise intuitive and bodily knowledge rather than purely cognitive skills.

### **Triple hermeneutic?**

Giddens (1986) has spoken of the 'double hermeneutic' (p.284) involved in sociological study: researchers interpret fields already constituted as meaningful by social actors, applying yet another frame of meaning - interpreting the interpretations of others (Bruhn Jensen, 2002). Geertz (1993) put it more brutally: 'what we call our data are really our own constructions of other people's constructions of what they and their compatriots are up to' (p.9). How much more complex is it, then, to interview people *about* their voice *through* their voice. Analysing such interviews draws on not only the researcher's interpretive framework but also the researcher's repertoire of beliefs, prejudices and preferences about the voice - in addition to their own voice and voice quality. (The impact of the researcher's voice on the interview and interviewee - and not just their accent - remains a seriously under-researched subject.)

In this sense conducting oral interviews about the voice might be called a 'triple hermeneutic': it has the potential to create self-consciousness as informants become aware that they are required to use, in their responses, the very instrument that they are describing. In some cases, Karpf (2006) found, this can render an interviewee, temporarily, literally mute, as though the expression of any views *about* the voice might simultaneously expose their own voice to critical scrutiny. In such cases humour often proved to be an effective technique for unblocking the interview.

## **Voice and language**

Another recurring problem in researching the voice is that it is invariably accompanied by words: when we hear voice we almost always hear language too. As has been suggested (e.g. Lawrence, 1992), this results all too often in the voice being elided with, or reduced to, speech, with the vocal element downgraded or considered subsidiary to language. To access 'meaning-as-sound', as Ihde (2007) puts it, 'there is a need to take note of the near and far reaches of sounded significance that remain 'outside' language-as-word' (p.151).

Voice research must also confront the elusiveness of vocal communication. In contrast with visual media, the vocal cannot be fixed or perceived in linear mode through, for example, the freeze frame. The auditory field is much less directional than the visual: its omnidirectionality and instantaneity can make it appear a slippery thing. Ihde (2007) noted the 'temporal edge' and 'trailing off' of sound, arguing that 'sound reveals time' (p.102).

Given these difficulties, it is not surprising that many researchers have turned to positivist epistemologies and methodologies in the hope of generating 'hard', sometimes quantitative, data. The best-known example of such voice research was conducted in the US by Albert Mehrabian, who eventually formulated a numerical equation about the voice's contribution to the communication of meaning (38%), as compared to that of words and facial expressions. Karpf (2006) exposed the extraordinary methodological poverty of Mehrabian's research, along with earlier positivist 'voice and personality' studies which attempted to create typologies of pitch and personality types or of pitch and emotions.

Such approaches, in their attempt to emulate the natural sciences, remove communication from its social setting or indeed lived experience and almost always succeed in essentialising the voice, as well as flattening out what Barthes called the ‘voluptuousness of... sound-signifiers’ (Barthes, 1977, p.182).

Karpf (2014a) demonstrated the consistent marginalisation of the oral among oral historians: she suggested that the rush to transcribe recorded interviews was based on a distrust of the modality and a desire to find sanctuary in the supposed facticity of the printed text.

Another methodological difficulty is to find words for a modality that conveys them but is wordless. As Ihde noted, ‘A book is read and its words are seen rather than heard. There are vast differences between hearing voices and reading words’ (2007, p.xx) in translating what he calls the voice’s ‘all-at-onceness’ into a linear mode. Is it even possible to peel away words from the voice and then find words for what can appear as mere residue? Might the attempt to develop a language in which to speak about the voice itself represent a diminishment of the vocal, a de facto declaration that the auditory is not eloquent enough of and by itself and requires the legitimisation of re-translation into words? Can only language valorise? Developing a language to describe and analyse the voice while retaining its affective vitality is indeed a challenge, although voice researchers can learn from the music therapy literature, in its analysis of the ‘pause’ and ‘contour’, for example, or in finding words for a modality that lies beyond them.<sup>6</sup>

Another problem lies in the ear. We access the voice through our listening of it: audition is all. How then can we separate our listening capacities from our sound-making ones? Karpf (2006)

---

<sup>6</sup> The author wishes to thank an anonymous reviewer for this suggestion.

suggested that we hear the genders differently; clearly class and ethnicity - our own and that of the speaker - also inflect the ways in which a voice is heard and understood. Voices cannot exist outside of being heard; the vocal apparatus and hearing organs are twin, reciprocal, symbiotic processes. There is no pure, unmediated space in which the voice can somehow lie suspended, beyond audition; hearing is mediation, and so voices are mediated purely by the act of being heard. Attributing qualities to voices thus must inevitably be a subjective experience, even while it is culturally-shaped, and thus requires a degree of reflexivity in the researcher,<sup>7</sup> as well as posing questions of validity.

This article may, by now, have persuaded the psycho-social researcher to avoid the voice at all costs, given the myriad difficulties attending them. Interesting voice research, however, does not merely recognise the situated, socially-constituted nature of speech and voice but embraces its very temporality. Eschewing the hopeless pursuit of positivist 'certainty', psycho-social researchers work with rather than against the fleeting, flickering nature of the voice - the fact that our own voices differ not only according to whom we speak but even at different times of the day - and its performative features. As Fónagy (1983) acknowledged, 'Il est impossible.. dire [la même phrase] à deux reprises exactement de la même façon' (p.10). ('It is impossible to say the same sentence twice in exactly the same way', AK translation). Karpf (2006), far from attempting to 'factor out' such 'inconsistencies', homed in on them as its research topic in a study representing a rare, if imperfect, sustained attempt to elicit from non-professionals accounts of their attitudes towards their own voice, and of their social and individual competencies in interpreting the voices of others, both subjects almost entirely neglected in the realms of quantitative research on the voice.

---

<sup>7</sup> Karpf (2013b) discusses researchers' personal investment in sensitive research areas.

### **Why psycho-social studies are particularly well-suited to voice research**

Psycho-social studies is not in itself a unified field, using rather a plurality of approaches and a cluster of methodologies. What they have in common is ‘the challenge... to invent new ways of thinking the social and the psychological together rather than separately and hence to recognise the extent to which they are distinct expressions of a unified process’ (Stenner and Taylor, 2008, p.423). Frosh (2010) describes it as a Möbius strip, in which inside and outside flow together as one. This makes it an ideal way of analysing the voice, an instrument that is propelled outwards from the body's interior, at once the product of the social and the psychological and which simultaneously challenges such a division. As LaBelle suggests, ‘the mouth is precisely what puts into question the separation of interior and exterior, as distinct and stable’ (LaBelle, 2014, p.2).

Stenner and Taylor insist on the 'transdisciplinary' nature of psycho-social studies, rather than its multi-disciplinarity or inter-disciplinarity. A multi-disciplinary approach views issues from a variety of vantage-points rooted in different disciplines, while an inter-disciplinary stance brings the concepts and methods of one discipline into another. Transdisciplinarity, however, ‘deals with that which *escapes* disciplinary knowledge.... creating new spaces of knowledge and practice.... a hybrid social space that neither psychology nor sociology adequately comes to terms with’ (Stenner and Taylor, 2008, p.431). Since the voice itself could be said to occupy a hybrid social space, psycho-social studies provide a fitting conceptual framework in which to analyse it. Redman urges researchers ‘to attend to what it is in the phenomenon that is difficult to think, is being avoided, escapes or remains unknowable’ (Redman, 2014a, p.16). So many elements of voice have largely escaped traditional analysis - indeed, often escape language itself, both in the sense that they concern the nonverbal dimension of speech but also

because, outside phonetics, neuroscience and speech therapy, there is no cultural discourse for the extra-linguistic or paralinguistic aspects of voice.

Discourse analysis is perhaps the exception: although it privileges as its object of study grammar, vocabulary and the sentence, ethnographies of speaking and spoken language discourse generally at least allude to prosodic cues - pitch, stress, rhythm, volume and voice quality - even if they often stress their linguistic function or role in turn-taking. Nevertheless, some outstanding work, such as that of Labov and Fanshel (1977), has proved exceptionally sensitive to voice as an extra-linguistic system of meaning, and to the significance of changes of pitch, volume and even breathing. They devoted, for example, an entire volume to analysing 15 minutes of conversation between an anorexic patient and therapist in one session of psychotherapy. In addition to examining the linguistic structure of the interaction, they also paid close attention to the paralinguistic features of the patient's speech - its hesitations, silences, occasional volubility, intonation contours, pitch levels. Indeed, Labov and Fanshel recognised that the therapist's identification and understanding of the patient's intonation and vocal states such as her whine played a central role in the session.

The voice, of course, is the major channel in the psychoanalytic encounter. For the analyst, the patient's voice is a potent resource. Winnicott observed that, in his work with both adults and children, playing 'manifests itself... in the choice of words, in the inflections of the voice' (Winnicott, 1974, p.48). Bollas (1987) displays an acute sensitivity to his patients' voices, noting when one lowers her voice to attack herself and another raises her voice to express her sense of guilt. In the latter he also identifies 'a mature voice, which up till then I had simply not heard' (p.229). On the couch, Vasse (2010) suggests, it is often through a modification of voice that the analyst realises that some desire in the analysand has been touched and that they

have moved beyond discourse. Theodor Reik, a student of Freud's, argues what Polanyi and Bollas would later suggest, that 'the analyst, like his patient, knows things without knowing that he knows them. The voice that speaks in him, speaks low, but he who listens with a third ear hears also what is expressed almost noiselessly, what is said *pianissimo*' (Reik, 1948, p.145).

### **Suggestive psycho-social work on the voice**

Despite the multiple challenges faced by a psycho-social approach to voice research, some of which are outlined above, much fruitful work already exists. Stern (1985) brilliantly identified what he termed 'vitality affects' operating between parent and infant, through which fleeting surges of feeling crescendo and fade away. While many of these are gestural and facial, some are expressed vocally, and many are multi-sensorial. In so many ways the attunement or misattunement between carer and infant that Stern so expertly documents is played out through vocal channels. Stern notes the elusiveness of vitality affects, lying outside 'our existing lexicon or taxonomy of affects' (p.54). 30 years later this remains largely true, and points to some of the methodological problems in researching them.

Gee (1991) analysed the prosodic features of speech. Prosodic phrases provide emphasis in a sentence; pitch disruptions (called pitch glides) help a speaker identify the salient and important information that they wish to impart: 'How a text is actually said is crucial to the structure we assign it in terms of idea units, focuses, and lines.' Deploying discourse analysis along with psychological sensitivity to scrutinise an account by a woman in her 20s with schizophrenia, whose speech was characterised by doctors as 'disturbed', Gee brilliantly reconfigured it as an

example of human narrative sense-making. In this way he reunited voice and language, restoring a unity he found in the woman's account.

Karpf (2006) applied Freud's concept of 'the return of the repressed' to interviewees' negative attitudes towards their voices, exploring ways in which aspects of the self that have been disavowed may leak out through the voice. Karpf (2013a) sought to make sense of the role of the voices of regular radio presenters through object relations theory, by applying the ideas of Bion, Bick and Winnicott, and by bringing together the notions of 'imagined communities' and 'containment'. Indeed, the special issue of the *Radio Journal* in which it was published was probably the first to explore 'the reparative, emotional work of radio' as a significant internal object of mind (Bainbridge and Yates, 2013, p.7).

Pre-eminent among voice researchers is Maiello. Her 2003 paper monitored the contrasting rhythmic interactions between a South African mother and baby and their Western counterparts. In an earlier award-winning paper, Maiello speculated about the role of prenatal trauma in causing later autism, suggesting that its origins may have lain in a psycho-physical retreat from the auditory experience of the mother's voice (Maiello, 2001). Especially suggestive if inevitably speculative, Maiello's hypothesis of the existence of pre-natal sound memories connected with the mother's voice gives rise to the possibility that the foetus is capable of some form of introjection (the unconscious incorporation of other's attitudes and beliefs), and that the introjected elements 'have at least partly sound-qualities deriving from the child's perception of the mother's voice' (Maiello, 1997, p. 158). Maiello calls this, one of the earliest objects, a 'sound-object', and argues that it might entail some rudimentary differentiation between 'me' and 'not-me'. Vocality, in this characterisation, does not so much represent feelings as embody them, and Maiello's case is made the more persuasive by her

Careful observation of infants. By getting a voice of its own, she concludes, the baby can reproduce that part of its pre-natal world that escaped its control during its first, mute part of its prenatal life, ie the maternal voice. This is not so far from Lacan but has been tested out empirically through sensitive infant observation.

The approach and findings of these different studies together suggest four main ways in which the psycho-social researcher can make use of the speaking voice. The first is as an interview aid in research on a wide variety of topics seemingly unconnected with the voice: by close listening to non-verbal tonal cues in the voice of the interviewee, the sensitive qualitative researcher can be guided towards areas that invite further investigation and away from ‘no entry’ ones to which the informant might be highly resistant. There is also a subjectivity and a politics of listening. Norkunas (2013) asks: ‘What can be heard? The listener negotiates what she can hear, must hear, hopes to know and cannot bear to know... Empathetic listeners are ever sensitive to the nuances of trauma in the life story: long silences, detachment, a change in voice or body language.’ Such nuanced listening can attune the researcher to a kind of vocal excess, to passages in the recording of interviews that are paralinguistically charged.

The second way that the voice can be deployed by the psycho-social researcher, therefore, is interpretively after the interview, to deepen their reading and understanding of it. Karpf (2014a), for example, contrasted the transcript and recording of an interview in ‘Belonging: voices of London’s refugees’, part of the Refugees Communities History Project, with Mercedes Rojas about her Chilean husband who had ‘disappeared’.<sup>8</sup> She suggested that the transcript read somewhat generically, as if Rojas were describing an experience that might have

---

<sup>8</sup> <http://collections.museumoflondon.org.uk/Online/object.aspx?objectID=object-748929&start=0&rows=1>, accessed 9.04.14

taken place under any number of Latin American dictatorships. Speaking in a language that was not her mother tongue, Rojas also used the word ‘infringing’ where she probably meant ‘inflicting’: transcripts inevitably draw attention to such slippages because we expect them to conform to the norms of writing rather than speech and notice when they do not.

When she listened to the recording of the interview with Roja, however, what emerged vividly for Karpf was a continuing and very personal human tragedy; coloured by Rojas's soft voice, accent and slow pace, it was transformed into an anguished, highly *embodied* individual attempt to make sense of senseless acts. Moreover, unlike the reader, the listener became aware of an ambiguity conveying a double agony: the torture Rojas referred to was that inflicted upon her husband - but also upon herself. The oral thus animated and particularised a historical trauma.

Thirdly, the voice itself constitutes a fruitful psycho-social research topic, as Karpf 2006 found. Her interviewees expatiated on their sense of identity, their relationship with parents and partners, their transition into motherhood – all through the medium of the voice.

Finally, voice can be adduced to research projects on a wide variety of different subjects as an additional dimension, enabling the researcher to capture aspects of personal and social experience that might otherwise elude description or be too freighted with expectations. An exploration of gendered occupational roles, for instance, might ask participants to identify the characteristics of commanding voices in the workplace and the strategies used by marginalised workers in making a vocal intervention. An investigation into the experience of ageing could include questions about how the respondent feels about their changing voice. The discursive absence of the voice might, in such cases, turn out to be an asset.

## Ways of hearing

Inhering in the voice are all the constituents of self and culture. Butler (1990) maintained that ‘the substantive effect of gender is performatively produced’ (p.24) and that ‘gender is always a doing’ (p.25), to which we might add: and always a speaking. Gender's acts and gestures, she suggests, produce on the surface of the body what appears to be its internal core or substance: the tacit collective agreement to perform polar genders are ‘part of the strategy that conceals gender's performative character’ (p.141). Karpf (2006, 2015) suggests that voice is one of the main avenues through which such performativity occurs.<sup>9</sup>

And of course, ‘even before I open my mouth to speak, the culture into which I’ve been born has entered and suffused it. My place of birth and the society where I’ve been raised, along with my mother tongue, all help regulate the setting of my jaw, the laxity of my lips, my most comfortable pitch’ (Karpf, 2006, p.182). Intonational differences can create cross-cultural misunderstandings; social class and social status are inscribed in (or, perhaps, *intoned through*) the voice. While accent is commonly understood as a vector for class, public discourse is less sensitive, for example, to the relationship between volume and entitlement. Karpf (2006) discussed the loud voice as a mark of dominance and, in other settings, as a defiant gesture by marginalised people congregating in public space. Indeed, pre-war German Jewish refugees were urged to desist from speaking German loudly in public places. By talking at their normal volume in a foreign language, immigrants are understood to be laying claim to public space and resisting power that is exercised vocally and aurally. As Tonkiss (2003) remarked, ‘The

---

<sup>9</sup> Butler (1990) recognises that the effect of gender is produced through repeated bodily gestures, movements and styles but falls short of explicitly including the voice among them. The documentary film *Do I Sound Gay?* (Do I Sound Gay? 2014) explores the performativity of hyper-masculinity through the voice and seeks to both destigmatise what has come to be heard as a gay inflection but also problematise the whole notion of straight and gay voices.

immigrant... is *audible*, and indeed those forms of race thinking that cannot bring themselves to speak of skin often are happy to talk of language' (p.305). Perhaps, too, those forms of race thinking happy to talk about language cannot bring themselves to talk of voice, in every sense of the word: they wish to silence or un-voice the Other. Indeed Mendieta (2014) goes further and suggests that race is a 'sonic stigmata' and racism is phonocratic: 'the racist hears before he sees' (p.109).

About social class and voice, however, there is a remarkable paucity of literature. Where it is not preoccupied with accent, it is largely sociolinguistic in character, focusing on the social origins of 'elaborate' or 'restricted' codes, the pronunciation of particular consonants, dialects, regional speech styles or phonological patterning (Foulkes and Docherty, 2006), with scant attention paid to non-linguistic elements - social class and, say, voice quality, or volume, or tempo, let alone the ways in which voice might articulate the intersection of class and gender or ethnicity. Any of these might generate interesting research findings, as would the whole area of the voice in collective, public settings such as auditoria, political meetings, religious gatherings or sports matches, where orality has always been a powerful means of communication - what Bruhn Jensen (2006) terms 'reverberation' (pp.24-25).

It is a particularly exciting time to be researching the voice. Promising lines of inquiry could develop from the growing interest in the internal voice (e.g. Fernyhough, 2017). Technology is also changing our experience of the voice in intriguing new ways: today, some people some of the time will be more exposed to disembodied, electronically-mediated or digital voices than embodied ones. What then of the voice and body, or Barthes's 'grain of the voice' (Barthes, 1977)? This begins to seem like a serious over-simplification. Even if it evokes a phantasmatic body, the electronic voice has a presence of its own. 'Each technology brings its own

materiality, albeit in ways that are very different from the materiality of the voice eulogised by Roland Barthes' (Neumark et al., 2010 p.3). To some, a vocal interchange with a synthetic voice can never constitute 'real' human communication, but it is more fruitful, perhaps, to investigate the irresistible tendency to anthropomorphise such voices and the ways in which they become naturalised over time. LaBelle (2014) is surely right when he suggests that shifts in technology bring with them 'new configurations of embodiment, and, in addition, resituate how voicing comes to make incarnate a sense of self' (p.147).

## **Conclusion**

The speaking voice has been consistently ignored, marginalised or reduced to its semantic 'cargo', and its role as a research resource undervalued. At the same time, we need to resist idealising or essentialising it, or polarising the eye and ear. Modifying the 'ocularcentrist' argument, what is proposed here instead that it is our auditory lexicon that is impoverished rather than modern vocality itself. (Contrast this with the Tzeltal speakers of Tenejapa, Mexico, who possess over 400 different words to describe vocal states (Stross, 1989).)

Research about the voice and its uses calls not for positivist projects to create a universal typology of meaning for different vocal pitches but instead for a kind of 'situated hearing', contingent upon speaker and listener and the social, gendered and cultural settings in which both are positioned, and capable of echoing the multiple 'soundings' and 're-soundings' produced by the human voice and multiple decodings matched by the human ear.

By engaging with this richest of instruments, the speaking voice, the psycho-social researcher gains access to symphonies of meaning.

## References

- Ahmed, S., (2004). *The Cultural Politics of Emotion*. Edinburgh: Edinburgh University Press.
- Barthes, R., (1977). *Image Music Text*. Translated from the French by Stephen Heath. London: Fontana Press.
- Berendt, J.E., (1988). *The Third Ear*. New York: Henry Holt.
- Bollas, C., (1987). *The shadow of the object*. London: Free Association Books.
- Bourdieu, P., (2014). *The Logic of Practice*. Translated from the French by R. Nice, 1992. Cambridge: Polity Press.
- Bruhn Jensen, K., (2006). Sounding the Media. *Nordicom Review*, 27 (2), pp. 7-33.
- Butler, J., 1990). *Gender Trouble*. New York: Routledge.
- Cavarero, A., (2005). *For More Than One Voice: Towards a Philosophy of Vocal Expression*. Translated from the Italian by Paul Kottman. Stanford, Calif: Stanford University Press.
- Claude Evans, J., (1991). *Strategies of Deconstruction: Derrida and the Myth of the Voice*. Minneapolis, Minn: University of Minneapolis Press.
- Connor, S., (2014). *Beyond Words: Sobs, Hums, Stutters and other Vocalizations*. London: Reaktion Books.
- Derrida, J., (1997). *Of Grammatology*. Translated from the French by Gayatri Chakravorty Spivak. Baltimore, Md: The Johns Hopkins University Press.
- Do I Sound Gay?*, (2014). [Film]. Directed by David Thorpe. USA: IFC Films/Sundance Selects.
- Dolar, M., (2006). *A Voice and Nothing More*. Cambridge, Mass.: The MIT Press
- Droumeva, M., and Andrisani, V., (2011). *Towards a Cultural Phenomenology of Mediated Aural Practices*. Available at: <<http://www.interferencejournal.com/wp-content/uploads/2011/09/Interference-Journal-Issue-1-Toward-a-Cultural-Phenomenology-of-Mediated-Aural-Practices.pdf>> [Accessed: 1 May 2015].
- Fernyhough, C., (2017). *The Voices Within*. London: Profile.
- Fónagy, I., (1983). *La vive voix*. Paris: Payot.
- Foulkes, P. and Docherty, G., (2006). The social life of phonetics and phonology. *Journal of Phonetics*, 34.
- Froggett, L. et al., (2014). *Public Art and Local Civil Engagement*. London: Arts and Humanities Research Council. Available at: < <http://clouk.uclan.ac.uk/10961/> .[ Accessed: 2 July (2015).]

- Frosh, S., (2010). *Psychoanalysis Outside the Clinic: Interventions in Psychosocial Studies*. Basingstoke, Hants: Palgrave Macmillan.
- Gee, P., (1991). A linguistic approach to narrative. *Journal of Narrative and Life History*, 1.1, pp.15-39.
- Geertz, C., (1993). *The Interpretation of Cultures*. London: Fontana Press.
- Giddens, A., (1986). *The Constitution of Society*. Los Angeles, Calif.: University of California Press.
- Gillie, C., (2010). Sanglot. In: C. Gillie, ed. 2010. In: *De l'autre côté de la voix*. Paris: Éditions Solilang.
- Hollway, W., and Jefferson, T., (2000). *Doing Qualitative Research Differently*. London: Sage.
- Hollway, W., and Jefferson, T., (2005). Panic and perjury: a psychosocial exploration of agency. *British Journal of Social Psychology*, 44, pp. 147-163.
- Ihde, D., (2007). *Listening and Voice: Phenomenologies of Sound*. 2nd ed. Albany, NY: State University of New York Press .
- Jay, M., (1994). *Downcast Eyes: The Denigration of Vision in Twentieth-Century French Thought*. Berkeley, Calif: University of California Press.
- Karpf, A., (2006). *The Human Voice*. London: Bloomsbury.
- Karpf, A., (2013a). The Sound of Home? Some Thoughts on How the Radio Voice Anchors, Contains and Sometimes Pierces. *The Radio Journal*, 11 (1), pp. 59-73.
- Karpf, A., (2013b). Chain of Testimony: The Holocaust Researcher as Surrogate Witness. In: N.Chare and D.Williams (eds.), *Representing Auschwitz: At the Margins of Testimony*. Basingstoke, Hants.: Palgrave Macmillan.
- Karpf, A., (2014a). The human voice and the texture of experience. *Oral History*, 42 (2), pp. 50-55.
- Karpf, A., (2014b). Constructing and Addressing 'the Ordinary Devoted Mother': Donald Winnicott's BBC Broadcasts, 1943-62. *History Workshop Journal*, 78 (1), pp. 82-106.
- Karpf, A., (2016). The persistence of the oral: on the enduring importance of the human voice. PhD thesis.
- Kreiman, J., and Sidtis, D., (2013). *Foundations of Voice Studies: An Interdisciplinary Approach to Voice Production and Perception*. Hoboken, NJ: Wiley-Blackwell.
- LaBelle, B., (2014). *Lexicon of the Mouth*. New York: Bloomsbury Academic.
- Labov, W. and Fanshel, D., (1977). *Therapeutic Discourse: Psychotherapy as Conversation*. New York: Academic Press.

- Lawrence, A., (1992). Women's Voices in Third World Cinema. In: Rick Altman, 1992. *Sound Theory, Sound Practice*. New York: Routledge.
- Macfarlane, R., (2015). From aquabob to zawn. *The Guardian review*, 28 Feb. p.4.
- Maiello, S., (1997). Sound-aspects in mother-infant observation. In: S. Reid, (ed.), *Developments in Infant Observation: The Tavistock Model*. London: Routledge.
- Maiello, S., (2001). Prenatal Trauma and Autism. In: *Journal of Child Psychotherapy*, 27(2), pp. 107-124.
- Maiello, S., (2003). The Rhythmical Dimension of the Mother-infant - Transcultural Considerations. In: *Journal of Child and Adolescent Mental Health*, 15(2).
- McLuhan, M., (2011). *The Gutenberg Galaxy: The Making of Typographic Man*. Toronto: University of Toronto Press.
- Mendieta, E., (2014). The Sound of Race: The Prosody of Affect. *Radical Philosophy Review*, 17 (1).
- Murray Schafer, R., (1994). *Soundscape: Our Sonic Environment and the Tuning of the World*. Rochester, Vt: Destiny Books.
- Naas, M., (2011). The Violence of the Letter: From Levi-Straus to Rousseau. In: S. Gaston and I. Maclachlan, eds. *Reading Derrida's On Grammatology*. London: Continuum.
- Neumark, N., (2010). Introduction. In: N. Neumark et al., eds., 2010. *Voice: Vocal Aesthetics in Digital Arts and Media*. Cambridge, Mass: The MIT Press.
- Norkunas, M., (2013). The Vulnerable Listener. In: A. Sheftel and S. Zembrzycki, eds., *Oral History Off the Record: Towards an Ethnography of Practice*. London: Palgrave Macmillan, 2013.
- Ong, W., (1982). *Orality and Literacy*. London: Methuen.
- Plummer, K., (2001). *Documents of Life 2*. London: Sage.
- Polanyi, M., (1964). The Structure of Tacit Knowing. Lecture delivered on 17.02.64 at Duke University. Available at: < <http://www.polanyisociety.org/Duke/Duke2.pdf> > [Accessed: 29 June 2015].
- Polanyi, M., (1967). *The Tacit Dimension*. London: Routledge and Kegan Paul.
- Redman, P., (2014a). What are psychosocial studies anyway? And do we need them? Paper given at the *School of Applied Social Sciences*, Brighton University, 14 January, 2014.
- Redman, P., (2014b). What is the psycho-social anyway? Paper given at the *Association for Psychosocial Studies Conference*, University of Central Lancashire, 16-17 December 2014.

- Reik, T., (1949). *Listening with the Third Ear*. New York: Farrar, Straus and Company.
- Schmidt, L.E., (2003). Hearing Loss. In: M. Bull and L. Back, eds., 2003. *The Auditory Culture Reader*. Oxford: Berg.
- Simmel, G., (1907). Sociology of the Senses. In: D. Frisby and M. Featherstone (eds.), 2000. *Simmel on Culture: Selected Writings*. London: Sage.
- Stenner, P., and Taylor, D., (2008). Psychosocial welfare: Reflections on an emerging field. *Critical Social Policy*, 28(4), pp.415-437.
- Stern, D., (1985). *The Interpersonal World of the Infant*. New York: Basic Books.
- Sterne, J., (2003). Medicine's Acoustic Culture. In: M. Bull and L. Back, eds., 2003. *The Auditory Culture Reader*. Oxford: Berg.
- Stross, B., (1989). Speaking of Speaking: Tenejapan Tzeltal Metalinguistics. In: R. Bauman and J. Sherzer, eds., 1989. *Explorations in the Ethnography of Speaking*. Cambridge: Cambridge University Press.
- Tonkiss, F., (2003). Aural Postcards: Sound, Memory and the City. In: M. Bull and L. Back, eds., 2003. *The Auditory Culture Reader*. Oxford: Berg.
- van Leeuwen, T., (1999). *Speech, Music, Sound*. Basingstoke, Hants: Macmillan.
- Vasse, D., (2010). *L'arbre de la voix*. Montrouge: Bayard.
- Visual Matrix Workshop, (2015). *Visual Matrix Workshop: Imagery, Affect and Visualisation as a Psychosocial Research Method*. London: Association of Psychosocial Studies/Birkbeck Institute for Social Research. <<http://www.psychosocial-studies-association.org/visual-matrix-workshop/>> [Accessed: 2 July 2015].
- Wetherell, M., (2012). *Affect and Emotion*. London: Sage.
- Winnicott, D.W., (1974). *Playing and Reality*. London: Penguin.