

10. *Vocal persona*

THE VOICE is mankind's primary musical instrument.¹ Its importance has already been mentioned in conjunction with prosody, with timbre and aural staging, with pitch range and register, and of course with melody. As we'll see in Chapter 13, voice is also at the basis of several musical sign types, including transscansions, language identifiers and paralinguistic anaphones. The purpose of *this* chapter is to suggest ways of denoting perceptions of the nonverbal aspects of voice.

Before going any further I need to clarify two points. One is the meaning of PERSONA, the other an explanation of the mainly vernacular source of ideas presented in this chapter.

Persona

PERSON, without the final *A*, means an individual human being and PERSONALITY 'the distinctive character or qualities of a person'. In Latin, Italian and Spanish PERSONA (*with* the final *A*) just means PERSON but in English PERSONA denotes 'an aspect of the personality as shown to or perceived by others'.² Actors, singers and other types of performer aren't the only ones to present personas² because we all have to assume different roles in different situations at different times of life. Here are sixteen examples from my own life: [1] child in relation to parents; [2] parent in relation to a child; [3] student in relation to teachers and [4] fellow students; [5] teacher in relation to students and [6] colleagues as well as [7] administrators; [8] lover; [9] husband; [10] good friend; [11] reasonably

1. For the centrality of voice, see, for example, the *Kodály Approach* (n.d.) at <http://britishkodalyacademy.org/> [120117]. The term 'vocal persona' originates with Cone (1974). Its various meanings are instructively discussed by Frith (1996: 196-200).
2. PERSONA is conceptually opposed to ANIMA, the individual's *inner* personality. I'll be using the English (or Spanish) plural form PERSONAS rather than the etymologically correct Latin plural *personae* which is rarely used when talking about the phenomenon. All quoted definitions are from *The Concise Oxford Dictionary* (1995).

'angry young man'; [12] even more reasonably (and cheerful) 'angry old man', latterly also 'benevolent but eccentric patriarch'; [13] 'one of the guys'; [14] classical musician; [15] rock musician; [16] solitary writer of academic texts like this.

It's not always easy to adopt the right persona in the right situation, especially if the role expected of us has to change, for example from child to parent or from student to teacher, but there's nothing intrinsically dishonest or schizophrenic about our ability to adapt to the appropriate role in the appropriate situation. On the contrary, it's an essential social skill. That's why VOCAL PERSONA should not, in what follows, be primarily understood as role play in the sense of putting on a vocal front, although that may sometimes be the case, but as *any aspect of personality as shown to or perceived by others through the medium of either prosody or of the singing voice*.

Vernacular sources

The ideas presented in this chapter derive less from the wealth of scholarly writing on voice, much more from having run popular music analysis classes for many years. Insights gained from that experience are supplemented with observations about how voice seems to be described in music reviews, album inlays, in ads for voiceover artists, even in casual conversation.³ All these vernacular sources for the verbal description of voice share a common trait: unlike the poietic terms designating musical structure defined by parameters of pitch, tonality, metre and episodicity, descriptions of voice, like those of timbre, are mainly *aesthetic*. This tendency may well be due to the fact that conventional music studies have yet to establish a systematic and widely accepted poietic terminology for vocal expression. There's simply very little by way of such jargon to intimidate non-musos, many of whom may struggle with the designation of music's tonal aspects and who are much less inhibited about describing timbre and vocal sound.

-
3. Two reasons for prioritising personal experience over scholarly writing in this chapter: [1] A sufficiently authoritative survey of relevant literature would take so much time that I'd never finish this book which is already long enough. [2] I'm more likely to add to the general body of knowledge about voice by focusing on my own experience and knowledge than by summarising those of others.

Another 'democratic' aspect of voice as part of musical analysis is that it's an instrument we *all* use in a musical way — prosodically — every time we *speak*. Most of us are experts at using our voices, not just to utter words but also to present our individual or group identity, and to express emotions, attitudes and behavioural positions (vocal personas). That's why I'll start with the music of the spoken voice, more precisely with my mother, followed closely by Robert De Niro.

'Don't worry about me'

When I was a child my mother would sometimes say 'DON'T WORRY ABOUT ME — I'M FINE' in a very sad voice. I remember the confusion that statement caused me. Did she mean the *words* DON'T WORRY ABOUT ME — I'M FINE or should I pay more attention to the *music* (prosody) in her statement: PLEASE WORRY ABOUT ME — I'M MISERABLE?

The second interpretation was probably nearer the truth than the first, not least because she wasn't always a happy person. She might have been feeling unwell or have just been involved in a domestic disagreement. Another reason for prioritising the 'music' of her statement was that her facial expression, body posture and gestures (in this case a lack of gesture), all aligned with her vocal timbre, volume, intonation, diction and speech rhythm but contradicted the meaning of her words. With a child's understanding of words and reason as privileged modes of symbolic interaction among grown-ups (although I wouldn't have put it that way at the time), I remember opting to take my mother's DON'T WORRY ABOUT ME at lexical face value. That decision once prompted my father to chide me for being insensitive. I didn't know what 'insensitive' meant but it didn't sound good, so I reverted to a more instinctive (or childish?) mode of interpretation, paying more attention to mother's 'music' and less to her words. Unfortunately, reading her statements on the basis of their 'music' (timbre, volume, inflexion, posture, facial expression, etc.) and ignoring her words also turned out to be wrong, because if I responded to her plaintive tone by asking 'What's the matter?' in a sympathetic tone of voice, I risked insulting her pride and hearing her retort: 'I said I was fine. Why do you never listen to what I say?'

It took me many years to realise that I could interpret my mother's [*plaintive voice* →] DON'T WORRY ABOUT ME — I'M FINE [*normal* →] as an integral statement, despite its mixed message. She actually meant:

I'm very sad and I find it hard to put on the brave face of self-control I know that grown-ups should. So, please show me some kindness while respecting the fact that I at least know I'm supposed to put on a brave face, even if I expect you to see through it.

That statement would have taken mother much longer and have demanded an unrealistic amount of reflective self-control. Her 'mixed message' was in that sense more efficient. I was simply slow to learn that you could consider the narrative context, scene, body language, the words *and* the music of my mother's mixed messages *as a whole*. It was a musogenic statement like the clear but complex musical moods mentioned in Chapter 2. I'm referring to those 'pallid verbal approximations' like DESPERATELY TROUBLED IN THE MIDST OF CALM AND BEAUTY, or SICK OF THE WORLD AND FEELING ALIVE BECAUSE OF THAT DISGUST.⁴

The DON'T WORRY ABOUT ME anecdote illustrates three important points about musical meaning, the first two of which have been discussed earlier. This chapter focuses on the third point.

1. Musical meaning is never created by the sounds on their own. They always exist in a syntactic, semantic and socioculturally pragmatic context upon which their semiosis depends.
2. Precision of musical meaning does not equal precision of verbal meaning or that of any other symbolic system. Music's apparent contradictions of *verbal* meaning (pp. 66-68; 167, ff.) should be understood as *musically* coherent.
3. Vocal timbre, pitch, intonation, inflexion, accentuation, diction and volume, plus the speed, metre, rhythm and periodicity of vocal delivery are parameters of expression conveying information about the sociocultural and personal identity (including meta-identity) presented by speakers or singers, as well as about their attitudes, feelings and emotions (i.e. their vocal persona).

4. For more on these *linguistically contradictory* approximations of *unequivocal musical* mood see pp. 66-68. See also the Mendelssohn quote on page 237.

'Are you talking to me?'

The third point just listed is illustrated in the video *Vocal Persona Commutations* ([YT OL7uc6L5nMQ](#)) which uses a twelve-second extract from the film *Taxi Driver* (1976) to highlight central aspects of links between voice and personality. In that twelve-second extract, Travis Bickle, the film's taxi-driving main character played by Robert De Niro, has just exercised his second-amendment right and acquired a gun to bolster his confidence when faced with the miscreants he meets in his job. In the clip he prepares to confront such scumbags by rehearsing the famous line 'ARE YOU TALKING TO ME?' in the mirror. It's worth examining the twelve seconds it takes De Niro to ask the question three times, including pauses, in order to discover which parameters of vocal expression communicate what. It's also worth testing which voices can and cannot be substituted for De Niro's in that famous scene so as to reveal the extent to which vocal persona is dependent on congruence with such factors as gender, ethnicity, age, social position, personality, clothing, opinion and attitude, acoustic distance and setting.

Leaving aside gesture, posture and facial expression for the moment and concentrating solely on the sound of De Niro's voice, minor differences of inflection, intonation, volume and accentuation can be discerned between the three variants of ARE YOU TALKING TO ME? In the first variant his voice is low-key but quite rapid with the quick but substantial rise of pitch normally used in English to pose questions expecting the answer yes or no; but it does sound sudden, as if he had been taken off guard. The second utterance is slightly slower, a little more deliberate and has clearer diction, suggesting that the imaginary low-life interlocutor did not take him seriously the first time. The third utterance is once again quite contained but includes more emphasis on 'me' and a little less on 'talking'. This shift in accentuation underlines personal involvement in the imagined encounter. Apart from these minor variants, it should be noted that De Niro does not raise (the volume of) his voice in anger or frustration, and that his is the normal voice of a young, probably white, North American, English-speaking male. In fact, without the narrative context and without De Niro's body language, there is nothing remarkable about his vocal persona in this scene any more

than there is about Travis himself, even though his lack of charisma may be what makes him narratively interesting.

Given that this relatively normal, neutral and uncharismatic personality has a correspondingly normal, neutral and uncharismatic vocal persona, it ought to be possible to replace his voice with others in order to discover which vocal elements are compatible or incompatible with which other simultaneous aspects of non-verbal communication.

The fact that we're in a noisy kitchen and that Travis is white, unshaven and wearing what appears to be a grey flannel air-force jacket tells us quite a lot. It certainly rules out several of the persona substitutions in the *Vocal Commutations* video. It's obvious that we're not hearing/seeing a child, nor a woman or old man. It isn't anyone African-American or East Asian, nor anyone from the higher echelons of society (unless they're slumming it). Nor can it be a samurai warrior from the sixteenth century or a young executive in Qatar or Saudi Arabia.⁵ The visuals also rule out robots, death-metal monsters, chipmunks or anything else that doesn't look or sound like a Caucasian male, a member of the popular classes, and aged between 25 and 45.⁶ But there's more visual information restricting the vocal commutation possibilities.

Since De Niro is about one metre away from the camera, convincing alternative voiceovers cannot sound too close or too distant. For example, the repugnant intimacy of the lecherous DIRTY OLD MAN voice in the commutation video only works if De Niro's face is in extreme close-up. Obviously, then, one parameter of expression for vocal persona is perceived proximity. Another parameter is acoustic space. The commutation video's MONSTER and EVIL GOD voices, for instance, have been given cavernous reverb incompatible with the size and acoustic properties of the cluttered kitchen we see on screen.

-
5. There are many amusing pastiches of De Niro's ARE YOU TALKING TO ME? These include: [1] an infant imitating his aunt - [viiLqmeArdgeOI](#); [2] De Niro lampooning himself - [viiEpZxUhCE5I8](#); [3] spoken in Arabic by a man in a Saudi thobe who emerges from an airport toilet and brandishes a banana instead of a gun - [viiCK7WCKpeVic](#); [4] *Wrestle Mania 21* commercials - [viiCw39KMcrJBU](#); [5] in French — 'C'est avec moi que tu parles? — from *La Haine* (1995) - [viiokQJPUTQMqA](#).
 6. Monster and robot voiceovers work better if you manipulate the visuals (6:50-7:10).

The first time Travis asks the famous question he is at the far right edge of the screen with his body facing screen left. He turns his head towards us, as if just having heard something coming from the direction of the camera. He looks surprised, his eyebrows are raised and his head tossed back a bit. It's the look of someone literally taken aback. However, there is nothing except the immediate narrative context that rules out the possibility of *pleasant* surprise, which is why the commutation video's first BABY TALK voiceover works well if viewers imagine the camera being the baby's point of view and that the De Niro character is a proud father, surprised and delighted by his infant's contented gurgling as he walks past.

For the second version of the question De Niro has half turned toward the mirror/camera, tossed his head back a bit more and raised his eyebrows higher. Once again, it's mainly the narrative context that rules out a possibly positive interpretation of Travis's body language and which lead us to believe that this more clearly 'taken aback' posture is more likely to express affront and irritation than surprised delight. Even his teeth, visible for a short moment in an unsmiling mouth, suggest confrontation. He also seems to be looking down his nose at his imagined interlocutor, and since his diction and accentuation are slightly more forceful than before, the BABY TALK voiceover of the delighted dad is less convincing here. Furthermore, the despondent, depressed and weak vocal persona substitutions align badly with De Niro's posture, facial expression, accentuation and diction during these three seconds.

The third version is gesturally the clearest. His body is turned a little more towards the camera as he points to his own chest in sync with 'to me'. Again, prior knowledge of the Travis character will likely lead viewers to see his grin as insolent, and his hand gesture as expressing personal affront. However, without such prior knowledge and with the addition of a few sonic correctives to the narrative (gurgling baby, the mother's 'aaah!'), ARE YOU TALKING TO ME?, spoken by a delighted and proud father, aligns quite convincingly with this third variant of the famous question.

Several vocal persona commutations don't work because of problems with lip sync. For example, stereotypical robot voices, as we saw earlier


(pp. 281-282), apply equal durations for each syllable, while depressed and despondent statements are delivered at a slower rate than that of ARE YOU TALKING TO ME? spoken normally. Besides, a depressed voice is usually accompanied by depressed body posture and facial expression — drooping shoulders, head hung low, eyes looking down, no eye contact, etc. Lip-sync problems also demonstrate that whispering and other types of vocal close-up are incompatible not only with the lack of extreme visual close-up in the *Taxi Driver* sequence but also with its speed of delivery. Whispering has to be slower than talking because it must compensate for the absence of voiced consonants and the full transients that identify vowel sounds, while intimate statements delivered forcefully at breakneck speed sound ridiculous.

Poietic, acoustic and aesthetic descriptors

None of the observations just made about ARE YOU TALKING TO ME? in the *Vocal Persona Commutations* clip should come as a surprise.

‘[L]isteners who hear voice samples can infer the speaker’s socio-economic status..., personality traits,... and emotional and mental state... Listeners exposed to voice samples are also capable of estimating the age, height, and weight of speakers with the same degree of accuracy achieved by examining photographs... Independent raters are also capable of matching a speaker’s voice with the person’s photograph over 75% of the time.’ (Hughes *et al.*, 2004: 296)

Indeed, the relationship between an individual voice and its unique personal identity has given rise to the voice print branch of the security industry with its biometric claims about defeating credit card fraud or ensuring ‘that prisoners incarcerated in their homes or out on temporary passes [are] where they were supposed to be’.⁷ Whether or not the sales spiel of voice print marketers has any validity isn’t the point here, although incredulity may be warranted, bearing in mind the technical crudity and socio-linguistic stupidity of most corporate ‘voice recognition’ systems.⁸ The point is that insights about congruence between

7. SearchSecurity.com  searchsecurity.techtarget.com/sDefinition/0,,sid14_gci944937,00.html [2008-02-11]. ‘Voice authentication products’, the site informs us, ‘are available from a number of vendors, including Vocent,... Courion Corp., and VoiceVault’.

individual voice and personal identity are nothing new. Indeed, the very word *person* contains the morpheme *son*, meaning sound, and Latin's *personare* literally means to sound (*sonare*) through (*per*), to sound forth, etc. Moreover, the original meaning of the Latin word *persona* is 'a mask... as worn by actors in Greek and Roman drama'.⁹ Its transferred meanings of performed role, personality, etc. derive from the fact that revealing the true nature of a dramatic character involved projecting the voice of that individual through the mask worn by the actor playing that role. His or her voice had literally to sound (*sonare*) through (*per*) the mask — *vox personans* — out into the auditorium, into the audience's ears and brains.

Links between voice and personality are also clear from numerous on-line searches for terms like VOICE, VOCAL, PERSONA and PERSONALITY. Although descriptive adjectives of voices were, as we shall see, far from uncommon, another frequently recurring type of voice characterisation related, unsurprisingly, voice to personality. Among the more striking examples found of persona descriptors of Anglo-US singing voices were (artists in brackets) HARD-EDGED SEXUAL EXUBERANCE (Chaka Khan), IMPISH CHIRP (Katryna in *The Nields*), [they looked and sang like] BARBIE DOLLS (Wilson Philips), CUDDLY VOCAL PERSONALITY (Beverly Sill), a NERVOUS TEENAGER, FEARFUL OF BEING REJECTED (Buddy Holly), an ANGRY SMURF (Eminem) and THE WESTERN MYTHICAL GIRL/WOMAN, HEARTBROKEN YET RESILIENT AND ENTIRELY FEMININE... [with an] EDGE BETWEEN VULNERABILITY AND WILLFULNESS (Linda Ronstadt).¹⁰

The voice descriptions just listed sound neither serious nor scientific. They're more likely to come across as spuriously subjective, at best as amusing or imaginative. That's an understandable objection but it needs to be moderated in the light of four points made so far: [1] the fact that '[i]ndependent raters are... capable of matching a speaker's voice with the person's photograph over 75% of the time'; [2] the apparent commercial success of voice print companies; [3] the patterns of congruence and incongruence in the *Taxi Driver* commutation clip; [4]

8. For documentary evidence of 'voice recognition' incompetence see Tagg (2008b).

9. *Cassell's Latin-English Dictionary*, London, 1968. See also Lacasse (2000: 42-46).

the etymology of the word *person[a]* itself. Those four points suggest that patterns of linking voice with personality do exist and that such links can be verified intersubjectively in given cultural contexts. We'll return to these links and to their usefulness in discussing the 'meaning' of singing voices, but it's useful to be first aware of other approaches to the issue of describing vocal sound.

The 'musical' properties of vocal sound, spoken or sung, can in general be understood and verbalised using one or more of three main perspectives: [1] the physical techniques of its production (poietic perspective); [2] its measurable physical attributes as sound (acoustic); [3] its perception, interpretation and effects (aesthetic).

The POIETIC PERSPECTIVE focuses by definition on how particular parts of the human body are used to produce particular vocal sounds, e.g. larynx, throat, mouth, jaw, tongue, nose, lungs, diaphragm, shoulders, chest, head. Recurrent concepts are breathing, control, projection and register (chest, mixed, head, falsetto). Now, as we'll see later in this chapter (p. 376 ff.), the ability to reproduce, at least roughly, a vocal sound can help us understand its meaning. That's why some familiarity with the physical implications of the terms just mentioned can be useful in identifying the body posture (shoulders, chest, head, etc.) and facial expression (mouth, jaw, nose, etc.) most conducive to the production of a particular vocal sound. That knowledge in its turn contributes to insights about the emotional state of the person[a] behind the vocal sound in question.

-
10. The first three comments were online at: [1] [\[1\] !\[\]\(815df092dd722ee9268ef8e6d0193e3a_img.jpg\)rollingstone.com/artists/chakakhan/albums/album/243746/review/5945280/chaka](http://www.rollingstone.com/artists/chakakhan/albums/album/243746/review/5945280/chaka); [2] [\[2\] !\[\]\(c72edb9626cad660f3a9f5fb0f22a68c_img.jpg\)furia.com/page.cgi?type=twas&id=twas0196](http://www.furia.com/page.cgi?type=twas&id=twas0196); [3] [\[3\] !\[\]\(0c564128c6342bd2f601e97f4518828a_img.jpg\)whiteperil.com/posts/1093202710.shtm](http://www.whiteperil.com/posts/1093202710.shtm). The Buddy Holly comment is in Bradby & Torode (1984) and the Eminem description comes from one of my students in Liverpool (c.1997). The Ronstadt words were at [\[4\] !\[\]\(5cb79a1c9acdf5d94bce345803852578_img.jpg\)superseventies.com/spronstadt.html](http://www.superseventies.com/spronstadt.html). Here are a few more colourful descriptors of popular music vocal personas culled from the internet: 'Dylan on too much coffee and not enough sleep'; 'forlorn foghorn'; 'about as human as a voice-mail'; 'smooth-sailing love man'; 'flitting from folksy romantic to cute little girl to abrasive spite-monger'; 'jauntily devilish'; 'husky, mordant'; 'down-and-dirty'; 'all dressed up for a late-night smoky jazz club or an upscale blues joint'; 'world-weary-cool-kitty'; 'ultra-snide and confrontational'; 'from playful coquette to vintage jazz diva'; 'vulnerable tough guy'; 'schizo-barmy [and] speedball-bonkers'; 'from the old-dog croon... to the gruff, staccato bark' (source details searchable in [\[5\] !\[\]\(cc23775bf31a648cde5902baa397f9aa_img.jpg\)tagg/articles/VocPersUnsystNotes.txt](http://www.tagg/articles/VocPersUnsystNotes.txt) [120117]).

The ACOUSTIC PERSPECTIVE focuses on the physical properties of vocal sound, i.e. on volume (dynamics, intensity) and timbre (attack, decay, fundamental pitch, overtones, etc.). The number of possible variations in these quantifiable parameters is virtually infinite; their combination forms the physical basis of the enormous variation of sounds that human voices can produce and of how those sounds are perceived. Now, there's no room here to explain even the rudiments of acoustic physics in relation to the human voice and its perception. Readers are instead referred to a wealth of literature dealing with correlations between the measurable physical properties of particular sounds and their perception.¹¹ That said, basic awareness of parameters like fundamental pitch, overtones, intensity, attack and envelope can, by drawing attention to the physical properties of a particular sound, refine procedures of commutation (e.g. changing timbre to check on possible changes of perceived effect) and lead to greater precision of semiotic analysis.¹²

The AESTHETIC PERSPECTIVE is characterised by how sounds are perceived, interpreted, reacted to and used by those who hear them. Since this book is aimed primarily at music's users I'll try, in what comes next, to sort out the various ways in which we seem to verbalise our perception of different voices. Then, after an excursion discussing basic differences between speaking and singing, the chapter will end with suggestions about how categories of vocal persona can be used in the semiotic analysis of music.

Aesthetic descriptors

Between 2005 and 2008, I trawled cyberspace for websites containing various combinations of VOICE, VOCAL or VOICEOVER and including words like QUALITY, TIMBRE, PERSONA, PERSONALITY, ATTITUDE and CHARACTER. In addition to having annoyed students, friends and colleagues by asking them to describe voices to me, I also took an interest in vocal casting, a specialist profession in which verbal descriptions of voice play an essential part. For example:

11. See Sundberg (1987); see also, for example, Bouchard (2010), Lacasse (2000), Lomax (1968), McHugh *et al.* (1997), McPherson (2005), Mossberg (2005), Riding *et al.* (2006).

12. See Chapters 8 and 9, especially pp. 277-283, 305-315, 317-318.

'Seeking voiceover talent who can recreate a female witch voice... [The] project involves an English dub of a Russian animated feature... The witch is very old, around 70.¹³ Also seeking a counsellor voice. High pitched and whiny,... middle-aged.'¹³

Here's a character description circulated by a Hollywood agency looking for computer game voiceover artists.

'X is the comically annoying, shape-shifting spirit of an ancient Druid Priest who serves as a kind of guide to [the hero] throughout the ages, as well as being a bothersome pest. He pops up unexpectedly to give advice, frequently at less than opportune moments, although he basically means well. He has a sarcastic, dry wit and is an irritating, amusing, occasionally caring and sincere presence that [the hero] has little choice but to tolerate throughout time. Since he can become anyone or anything, he exhibits a wide variety of voices and personalities. [This character is] "a sophisticated elder" voice in the range of Sean Connery or Ian McKellan, as Gandalf in *Lord of the Rings*, with comedic undertones. *Vocal Quality*: should be older and wise-sounding, but also with a "Celtic"-type accent.'¹⁴

That neither of these adverts describe voice from the poietic or acoustic perspective is hardly surprising since the jobs aren't for musicologists, singing teachers or acousticians; but the paucity of aesthetic sound-descriptive words does seem a little strange —just HIGH-PITCHED and WHINY for the counsellor and nothing else. Is this type of descriptor less relevant than others when advertising for a voice relating to a specific dramatic personality? To answer that question it's best to have an overview of the basic categories of aesthetic voice description. These categories are based on observations made from: [1] student comments in popular music analysis seminars since 1992; [2] online descriptions of speaking and singing voices; [3] comments from a voice casting agent in direct response to specific questions (p.359).¹⁴ Examples of descriptors from these three sources are shown in Table 10-1 (pp.356-357) where they are grouped into the following three principal categories.

13. voice123.com/lv/3093614.html [2006-04-20] (errors of English corrected).

14. Thanks to Dawn Hershey of Blindlight (blindlight.com, December 2007), for invaluable help with charting voice-descriptive language in the casting profession. Thanks to Peter D Kaye (Santa Monica) for putting me in touch with Blindlight.

[1] *SOUND DESCRIPTORS* denote perceived qualities of sound and are of two types: [1a] DIRECTLY SOUND-DESCRIPTIVE ADJECTIVES AND VERBS; [1b] *GENRE DESCRIPTORS* referring to the musical style and by extension to the genre associated with particular types of voice.

[2] *TRANSMODAL/SYNAESTHETIC METAPHORS* like ROUGH, SMOOTH, VELVETY and GRAVELLY connote sound on the basis of homologies from senses other than hearing. These synaesthetic descriptors are like anaphones¹⁵ in reverse in that they denote mainly kinetic and tactile sensations that are transferred to the perception of sound.

[3] *PERSONA DESCRIPTORS* seem to be the most common type of vocal characterisation. They can be divided into four subcategories.

Subcategory 3a in Table 10-1 (p.357), *NAMED PERSONS WITH DISTINCTIVE VOICES*, is often found in reviews, presumably to give readers an idea of what sort of vocal sound to expect from a recording they have yet to hear. My unjustifiably disparaging remark that Portishead's Beth Gibbons, in *Western Eyes* (1997), sounds like an under-age Billie Holiday belongs to this descriptive subcategory.¹⁶

Subcategory 3b in Table 10-1, *DEMOGRAPHIC DESCRIPTORS*, covers the gender and age, as well as the ethnic, cultural, social and economic background, of the vocal persona in question. These descriptors are very common in characterisations of both singing and speaking voices.

Subcategory 3c, *PSYCHOLOGICAL, PSYCHOSOMATIC AND EMOTIONAL DESCRIPTORS* (p.356), are the most common of all. They qualify or allude to the feelings, attitude and morality, and to the state of mind or body of the vocal persona in question.¹⁷

15. Anaphones are discussed in detail in Chapter 13, pp. 487-514.

16. Tagg & Clarida (2003:456).

17. Ear-Nose-Throat specialists document links between phonation and emotional state. Deary *et al.*, (2003:374) describe how '[v]oice production is subject to and indicative of psychological status'. See also McHugh-Munier *et al.* (1997) on links between coping strategies, personality and voice in female subjects.

Table 10-1. (a) Aesthetic voice description categories with examples¹⁸

1. Sound descriptors	
1a. Directly sound-descriptive adjectives and verbs	high-pitched,* whiny;* squeaky, booming, low-pitched, deep, full-throated, gruff, breathy, husky, guttural, distinct, harsh, indistinct, muffled, plaintive, rasping, roaring, shrill, stammering, loud, declamatory, soft, quiet, monotone, lispy, bird-like, hoarse, throaty. babble, bark, bawl, belch, bellow, bleat, blubber, boom, buzz, cackle, caterwaul, chant, chatter, chuckle, chirp, cluck, complain, cough, croak, croon, cry, declaim, denounce, drone, exclaim, gargle, gasp, giggle, growl, grumble, gurgle, hiccup, hiss, hoot, howl, hum, lament, laugh, lilt, moan, mumble, mutter, praise, preach, proclaim, pronounce, quack, quip, rant, rap, recite, roar, scream, screech, shout, shriek, sigh, snap [at], snarl, snigger, snore, snort, sob, spit, splutter, squawk, squeak, stammer, stutter, twitter, ululate, wail, warble, weep, wheeze, whimper, whine, whinge, whisper, whistle, whoop, yammer, yap, yawn, yell, yelp, yowl
1b. Genre-specific descriptors	e.g. blues shouter, Bollywood vocalist, cantautore, cantor, chansonnier, crooner, death metal growler, dramatic ballad star, fadista, folk singer, gospel artist, Irish tenor, jazz vocalist, lyrical soprano, muezzin, opera diva, payador, rapper, singer-songwriter, troubadour
2. Transmodal descriptors (anaphonic/synaesthetic descriptors)	
abrasive, angular, bouncy, brassy, clean, clear, creamy, effortless, full (-bodied), grainy, gravelly, hollow, laid back, meaty, piercing, rasping, relaxed, robotic, rough, rounded, sandpapery, scratchy, shaky, sharp, smooth, stilted, strained, sweet, textured, thick, thin, velvety, wobbly,	

* Word taken from the voiceover ads cited on page 354.

Subcategory 3d, ARCHETYPAL DESCRIPTORS, combines traits from all the other categories into personality tropes, sometimes in the guise of professions (priests, teachers, etc.), more often as narrative roles (heroes, villains, victims, lovers, parents, sages, witches, wizards, fools, tricksters, etc.). This subcategory has obvious advantages and drawbacks. Consider, for example, the following extract from a review of the 2005 Audio Bullys album *Generation*.

‘[T]he intro welcomes back Simon Franks’ pot-smoking, pill-popping, wife-beating, bottle-lobbing, “yes I do live on a council estate thank you very much”, vocal persona’...¹⁹

18. The aesthetic descriptors in this table are merely *examples* that in no way constitute a reliable scientific sample, let alone an exhaustive listing.

19. Review by Jamil Ahmad — musicomh.com/albums/audio-bullys_1005.htm [070225].

Table 10-1. (b) Aesthetic voice description categories with examples¹⁸

3. Persona descriptors	
3a. Named persons with distinctive voices	e.g. Sean Connery or Ian McKellan;* Clint Eastwood, the CLINT-EASTWOOD-IS-DIRTY HARRY guy, The Smurfs, Donald Duck, R2-D2, Richard Attenborough, Orson Welles, Morgan Freedman, Billie Holiday; Elvis Presley, Adele, Kate Bush, Björk, Maria Callas, Elba Ramalho
3b. Demographic	e.g. female, male; very old, around 70, middle-aged, older; young, child 'Celtic' accent; African American, French, Asian, Southern [US], British, upper class, working class, well spoken, from the country/slums, slang, regional accent
3c. Psychological, psychosomatic & emotional traits	means well*, caring*, sincere*, kind, friendly cute, cuddly, sweet, nice wise, intelligent, controlled, confident, regal arrogant, dramatic, over-the-top, extravert, provocative, ecstatic, orgasmic willful, determined, courageous energetic, flamboyant, bubbly, cheeky, cheery, comical*, coquette, jaunty, playful, keen, eager, sassy, interested interesting, complicated, quirky, annoying,* bothersome,* eccentric, cartoony hip, cool, sophisticated,* sensual, seductive, sexy vulnerable, embarrassed, scared, edgy, nervous, angry, frustrated, irritated, exasperated, bitter dark, mysterious, introvert sad, depressed, heartbroken, miserable, anguished melancholy, bored, bland, nondescript, neutral intimate, subdued, laid-back, relaxed, soft spoken, humble, simple, innocent, childlike angelic, ethereal raw, rude, tough, rugged, gritty, macho, aggressive devious, slimy, sleazy, nasty, evil, petty sardonic, sarcastic,* dry wit,* ironic, acerbic
3d. Professions, roles and archetypes	witch,* counsellor,* Druid Priest,* guide,* elder* little girl, heroine, leading woman, loving mother, devoted wife evil queen, witch, violent bitch, pretty princess, Barbie doll, vamp villain, big boss, gangster, lager lout, hooligan, dirty old man little boy, hero, father figure, leading man, wise old man monster, alien, robot sissy, miser, imp, evil child, suicidal student, nervous teenager, wiseguy, nerd, geek.

Even though POT-SMOKING, PILL-POPPING, WIFE-BEATING and BOTTLE-LOBBING may derive from the duo's lyrics, those epithets also connote the sort of voice many urban UK residents would, in 2005, associate with (male) slob behaviour (uneducated, careless, thoughtless, self-centred), not least because the activities of wife beating and bottle lobbing imply a particular (and particularly impaired) emotional state, as well as specific body postures, breathing patterns, etc.²⁰ Restricting ourselves to words listed in Table 10-1, it's much more likely that the vocal persona in question is loud and booming rather than soft or muffled, brassy

20. This voice has a lot in common with the 'narcissistic-aggressive type' (Benis, 2005).

rather than wobbly, working-class rather than upper-class, arrogant rather than humble, over-the-top rather than subdued, etc., in fact the sort of voice associated with football (soccer) hooligans (typically loud, male and working-class) and lager louts (vocally similar to football hooligans but with bottle lobbing as a likely additional trait).

The advantage of epithets like BOTTLE-LOBBING and LAGER LOUT is that they each encapsulate in a single concept a wealth of behavioural, psycho-social and vocal characteristics. The disadvantage is that descriptors like LAGER LOUT are culturally restrictive: only those familiar with particular aspects of UK popular culture in the post-Thatcher era will grasp the relevant social and vocal implications. As for the final epithet, the 'YES I DO LIVE ON A COUNCIL ESTATE THANK YOU VERY MUCH vocal persona', it would take another chapter to convincingly explain COUNCIL ESTATE and its relevant connotations, yet another to provide a viable socio-linguistic analysis of 'YES I DO LIVE'... and the final 'THANK YOU VERY MUCH'.²¹ In short, while the semantic efficiency of such epithets is undeniable within a restricted socio-cultural sphere, their connotations may well be meaningless to the rest of humanity, unless adequate equivalents can be identified in other cultural contexts.²²

Despite problems of cultural specificity, there is little doubt that aesthetic descriptors are in much wider general use than their poietic or acoustic counterparts and that persona descriptors, especially the demographic, psychological and archetypal subcategories, are particularly popular. This observation was substantiated by Dawn Hershey, a Hollywood professional specialising in vocal casting for video games and animated productions for film and TV. Here are two abbreviated extracts from email correspondence I had with Dawn on the subject.²³

21. Here's a drastically simplified and pallid summary of implied connotations with whose 'logic' I don't necessarily agree. Council estates are areas of low-cost rentable housing in the UK. Living on a council estate implies lower economic and educational status. Proclaiming 'yes I do live on a council state' implies being proud of not aspiring to higher social status. Adding 'thank you very much' to the statement is ironic. The speaker seems 'proud to be ignorant, happy to be a slob', etc.

22. For more on this recurrent problem of cultural specificity in aesthetic denotation, see the 'wet echo' issue on page 216.

23. Thanks to Dawn Hershey of Blindlight (blindlight.com, December 2007).

What problems do [producers] have in describing the type of voice they want?

The biggest problem they have when they first contact me is that they... describe body type, hair color [etc.]... I often need to ask more questions, such as age, accent, vocal quality, personality traits, quirks, and temperament...

How often do you or they refer to voices in terms of character archetypes?...

Almost always. Most frequently requested are LITTLE BOY, LITTLE GIRL, 20S HEROINE, 20S HERO, LEADING MAN, EVIL QUEEN, VILLAIN, MONSTER, ALIEN, SOLDIER, WISE OLD MAN, BIG BOSS, FAT CAT, GANGSTER.

Of course, none of the aesthetic vocal description categories discussed so far are mutually exclusive. For example, a particular kind of WITCH voice (description category 3d) might also be described as HIGH-PITCHED and CACKLING (category 1), SCRATCHY and PIERCING (2), as sounding like an ANGRY and EVIL (3c) EIGHTY-YEAR-OLD (3b) version of the ANNETTE BENNING CHARACTER IN *AMERICAN BEAUTY* (3a). Moreover, many descriptors bridge two or more categories: RASPING, for example, may be most commonly used to qualify sound (category 1), but the act of rasping (using a rasp as a coarse file in the original sense of the word) has as much to do with touch and movement (category 2) as with sound. Similar observations apply to words like SCRATCHY, PIERCING, CLEAN, SHAKY, STRAINED and GRAVELLY. In fact the the idea behind introducing the categories just mentioned isn't to create some sort of watertight taxonomy — a fruitless task in view of music's synaesthetic properties (p.62 ff.) — but to provide insights into the various ways that vocal sound is popularly perceived and described on an everyday basis. The aim of that exercise is in its turn to develop richer and more nuanced descriptions of what a vocal sound can communicate.

As endnote to this section it's worth mentioning the rich store of vocal personas exploited in consumerist propaganda. You only need think of the MOTIVATIONAL FOOTBALL COACH voice hyperventilating about 'all the fantastic bargains' ('Only 99.99!'... 'And that's not all!'... 'Hurry!'... 'Get yours now!', etc.), or of the HARD-BOILED SERIOUS-BUSINESS TOUGH MAN of action film trailers ('CLINT EASTWOOD IS DIRTY HARRY' etc.) to get the idea. Then there's the FEMALE BEST-FRIEND voice telling 'the girls' how to lose weight by buying low-fat cereal brand X, the HUSKY

TONGUE-IN-YOUR-EAR voice seducing you to buy super-silky shampoo Y or to stuff your face with super-smooth creamy chocolate Z. And don't forget the CHEERFUL BUT MATTER-OF-FACT YOUNG MOTHER enthusing about supermarket A or microwave meal brand B. The list could go on forever. The point is that this supply of regrettably recurrent and often regressive vocal stereotypes in commodity fetishism can be a very useful source of vocal persona descriptors, as long as you're sharing your observations about voice, spoken or sung, with others involuntarily exposed to the same sad sort of consumerist culture.²⁴

Vocal costume

'[C]lothing for a particular activity' or 'an actor's clothes for a part' are, according to *The Oxford Concise English Dictionary* (1995), two common meanings of the word COSTUME. With expressions like NATIONAL COSTUME, notions of group identity are added to the concept. In simple terms of perception, someone wearing a swimming costume is probably dressed for swimming (although it may be just a photo shoot), someone wearing the garb of a sixteenth-century Italian nobleman might be acting in Shakespeare's *Romeo and Juliet* (or just going to a fancy dress party), and a man in a tartan kilt and tweed jacket might have intimate ties with the Scottish Highlands (or be a tartanry fake). COSTUME is etymologically related to CUSTOM ('a particular established way of behaving') and semantically to the noun UNIFORM, meaning 'distinctive clothing worn by members of the same body', i.e. another type of costume signalling group identity.

*Vocal costume*²⁵ is a metaphorical expression meaning those aspects of phonation serving the three same sorts of function as literal costumes do: [1] to more easily carry out a particular activity; [2] to assume a role or to act a part; [3] to signal a particular group identity and/or to conform to a given set of cultural norms. Vocal costumes are something people put on like clothes for any or all of the reasons just mentioned:

24. Ads for the same product are often marketed differently for different target groups. See also cultural specificity of LAGER LOUT (p. 358) and WET ECHO (p. 216).

25. It took a year of sporadic reflexion to come up with the term *vocal costume*. *Vocal mould*, *uniform*, *habitus*, *template*, etc. were all dumped for a variety of reasons.

they are used on an everyday basis in both speaking and singing, as, I hope, the next section will illustrate.

Spoken costumes

PHONE VOICES provide a rich resource for studying vocal costumes, most probably because talking on the phone involves a particular type of sensory dislocation. It's one-to-one audio close-up (if the line is good) but without the visual, kinetic and potentially tactile aspects of one-to-one close encounters. A phone call takes place in the intimate acoustic space determined by the minimal distances between earpiece and eardrum, between lips and mouthpiece. Like it or not, we are at sonic kissing distance from our telephonic interlocutor down the road or on another continent. Such sensory dislocation may be less problematic when phoning 'friends and family' but it requires corrective measures if we're on the phone to someone we don't know, maybe talking to a representative for a large corporation or public institution. In these types of telephone encounter vocal costumes can come in handy.

When phones were a novelty in UK homes after World War II, many people of my parents' generation put on a special vocal costume when answering the phone. It was a more posh, more official-sounding voice whose diction, vowel sounds and intonation resembled that of BBC radio announcers or newsreaders of the day. These closely miked but widely broadcast official voices, by occupying the public space of the then contemporary media, seem to have been taken to represent a sort of common ground for close-up speaking with which everyone was familiar. Of course, since this vocal costume was also that of the old British establishment, it was not the most comfortable clothing to wear and was usually dropped when the person at the other end of the line was identified as more 'friends and family' than 'authority'. Moreover, the old-establishment BBC voice later became an anomaly in the wake of socio-economic change leading to the use of other vocal costumes. Technological development played a central role in this process.

As the number of radio channels increased, and as TV and hi-fi recordings became part of both individual and domestic acoustic space, the



repertoire of closely miked but widely disseminated voice types available for use as vocal costumes expanded radically. Consumerist propaganda was not slow to start using particular voice types corresponding to the intersubjectively verifiable and exploitable desires of a particular demographic. Those voice types are often used today in automatic phone 'dialogue' and 'voice recognition' systems. Or, as one EU-funded eCommerce document puts it:²⁶

'Advertisers adopt different strategies depending on the product they are selling and the intended audience. The same is true for creating automated telephone service dialogues.... Two of the [phone answering] personalities ['John' and 'Kate'] were created with the intention that they would portray younger, more streetwise [bank] agents and therefore would appeal to younger users.'

This sort of vocal costume marketing has led to telecommunications catastrophes like 'Simone' (Virgin Mobile USA), 'Claire' (Sprint), 'Julie' (Amtrak) and 'Emily' (Bell Canada). While each pre-programmed vocal persona initially sounds like an attractive, engaging, educated, helpful young woman, she turns out, in the reality of dialogue, to have the brains of a pea and the socio-linguistic skills of a drainpipe. So blind is the faith of corporations in the hocus-pocus of vocal pseudo-personalisation that huge amounts of consumer time and corporate money are wasted by replacing human beings with machines.²⁷ That said, although 'John', 'Kate', 'Simone', 'Claire', 'Julie' and 'Emily' are mere vocal drapes covering dummies in a sonic shop window, vocal costumes can serve some purpose, even inside the field of telephony, as long as no false claims are made about 'interactive dialogue systems'. For example, calling Milan's Radio Taxi 8585 in 2008 triggered a hold message advising you not to lose your place in the phone queue. The recorded voice sounded like that of a coquettish female secretary with a hidden laugh of flirtatious complicity in her tone; or, as a Milanese friend put it:

'It's as if she's saying to male customers "who knows what you and I could get up to while you wait?" ... It's not the voice of a mother — that

26. Spotlight project 1999-10314: 'Mass Market eCommerce Services using Multi-language Natural Spoken Dialogues'  [spotlight.ccir.ed.ac.uk/\[080224\]](http://spotlight.ccir.ed.ac.uk/[080224]).

27. Bell's 'Emily' (2003) cost \$10 million; see  speechtechmag.com/Articles/Editorial-Feature-Its-a-Persona,-Not-a-Personality-36311.aspx and  tagg.org/zmisc/FidoCallTranscr.htm.

would sound too old— or of a wife because that would be no fun. It's closer to the voice of an attractive and well-spoken lover... They assume of course... that most customers are men in need of flattery.'²⁸

Outside the weird world of brand-fixated, market-driven automated telephony, vocal costumes are simply a very real part of everyday life. If you have to address a crowd of people and there's no microphone, or if you have to keep order in a primary school class, or if you have to make your bid heard in a capitalist casino (stock exchange), you'll have to put on a vocal costume to do your job and to avoid causing long-term damage to your larynx. Hopefully, you'll change into a softer, happier, more sing-song costume ('motherese') when you talk to your baby child, into something less lilting when you have to answer important job interview questions, into something more CONTRITE YET COMPETENT when you have to explain why you are late delivering work to your boss, and so on. Or perhaps you're a psychoanalyst dealing with a highly strung patient, in which case you may well be tempted to put on your psychologist's VOCAL VALIUM costume. If you do, your patient will hopefully be less likely to throw a fit and, even if he/she does start kicking and screaming, you can at least pretend to keep *your* calm.

Attentive readers will already have noted that PUBLIC SPEAKING voice, PRIMARY SCHOOL TEACHER voice, a lilting PARENT voice (motherese),²⁹ the PSYCHOLOGIST voice ('vocal valium') and the EARNEST INTERVIEWEE voice are all aesthesic vocal descriptors, more precisely persona descriptors designating professions, roles or archetypes.³⁰ Those labels act as shorthand not just for a type of person (teacher, trader, psychologist, parent, etc.) but also for the type of voice associated with that type of person in particular circumstances. One final example of spoken vocal costume should clarify the issue once and for all.

Before I first went searching for vocal persona-related concepts in 2005, I'd never heard of the GIRLFRIEND VOICE. The online *Urban Dictionary* de-

28. Thanks to Alessandra Gallone (Milan) for answering questions about prerecorded phone voices in Italy [080225]. 'Stiamo cercando il vostro taxi. Restate in linea per non perdere la priorità acquisita' is what the flirting secretary voice says.

29. More about sing-song motherese on page 368.

30. See Table 10-1, p.357, subcategory 3d.

finest it as '[t]he change in pitch or tone of a man's voice when talking to their significant other'.³¹ The dictionary continues:

'The girlfriend voice is characterised by a higher pitch and a more effeminate tone with speech patterns scattered with pet names and childish words. This type of speech is usually frowned upon when used in the presence of other men'.... 'When he answers his phone and it's a guy, he uses his normal voice, but when he sees that it's his girlfriend calling, his voice instantly climbs several octaves and acquires a whiny, please-don't-be-mad-at-me tone. He's also the kind of guy who, when he gets on the phone with his girl, immediately walks away from the group, leaves the room, or tells everybody to shut up so he can talk.'

Even if 'several octaves' is a gross exaggeration, this explanation of the girlfriend voice provides a clear example of all three functions of vocal costume. It involves traits of phonation that firstly enable the man adopting it to *more easily carry out a particular activity*, in this case that of talking to his 'significant other' in the way he imagines will please her. Secondly, the same man *vocally assumes the role and acts the part* of boyfriend rather than that of 'one of the guys'. Thirdly, he signals that he belongs to the social sphere of the couple by *vocally conforming to the cultural norms* of conversation considered appropriate for that sphere of interaction, even to the extent of walking away from his male peers and telling them to shut up.

Sung costumes

Although pitch, loudness, timbre and tempo are parameters of expression common to both speech and music, and although prosody is a key element in music's cross-domain mode of representation (p. 62 ff.),³² there is apparently no language unable to distinguish in some clear way between what we call speaking and singing.³³ If that is so, what's the actual difference between the two?

31. urbandictionary.com/define.php?term=girlfriend+voice [080224].

32. For neurological basis of these similarities, see Özdemir *et al.* (2006).

33. '[I]t is usually easy to tell when someone starts singing. Anthropologists say this is true in all cultures.' (Sparshott, 1997: 199).

Singing as costume

Differences between speaking and singing can be understood in two general ways: [1] in terms of use, function, context and connotation; [2] in sonic terms. We'll start with the first of those.

If someone changes vocal mode from talking to singing you can say they 'burst into song' but no-one ever says that they 'burst into speech' from song because speech is in most situations the default vocal mode. The idea of song as an exceptional, special or heightened form of vocal expression can be understood in four ways.

1. *Being airborne*. This is the popular notion of song as vocal expression at literally a higher level, either *as air* (AIR is a synonym and *aria* (= air) the Italian for a tune), or as something carrying us up *into* the air, so that we are borne on the 'wings of song', 'flying' (*volare*), singing (*cantare*), 'in the blue' (*nel blu*), 'happy to be up there' (*felice di stare lassù*), etc.³⁴
2. *Special occasions*. People in the urban West tend to sing more on special occasions than in their day-to-day lives. We don't usually burst into song while filling out tax returns or having lunch with work-mates; but we might well sing at birthdays, weddings, funerals, the New Year, or on a night out in a karaoke club. We are also more likely to sing in patriotic or religious contexts where some aspect of ritualised transcendence is the order of the day.³⁵
3. *Heightened emotion*. Circumstances of heightened emotion such as lulling your little child to sleep, falling in or out of love, righteous indignation, erotic arousal, deep sympathy or sorrow, painful separation, great elation, bitter resentment, angry alienation, wondrous amazement, blissful contentment, etc. are more liable to bring on a

34. 'High, high, like a bird in the sky' (Abba, 1977b). For 'Wings of Song' (*Auf Flügeln des Gesanges*): see Mendelssohn (1833). 'Flying, singing', etc. is a literal translation of the hook lines in *Volare / Nel blu dipinto di blu* (Modugno, 1958). *Volare* is one of the most frequently covered postwar songs. The 1958 Dean Martin version includes other vocally airborne lines: 'Let's fly way up to the clouds, Away from the maddening crowds', ... 'No wonder my happy heart sings, Your love has given me wings', etc. See also the extremely popular Gipsy Kings version on the CD *Mosaïque* (1989).

35. Such vocal transcendence occurs in religious and nationalistic ritual (e.g. Christmas, international sports events), as well as in group-tribalist situations like football (soccer) matches, e.g. *You'll Never Walk Alone* (Liverpool FC, 1972).

song than what you feel when reading an instruction manual or attending a committee meeting. Put tersely, it can be 'worth making a song and dance' about some experiences but not about others.³⁶

4. *Religious chanting*. Before the advent of PA systems, speaking was for centuries replaced by chanting in reverb-rich venues like cathedrals and large mosques. The Word of God merely spoken by an officiant under such acoustic conditions could easily end up as an incomprehensible sonic blur in the ears of the congregation.³⁷ The fixed pitches and measured delivery of chanting helped overcome this prosaic problem. This historical observation reinforces the notion of song as 'transcendent', more 'otherworldly' than speech.

Although those four observations clearly suggest that song is a special or heightened mode of vocalisation, it could also be argued that singing is more down-to-earth, more somatic, or at least more directly emotional, than talking, the dominant or default mode of vocal interaction among grown-ups. However, just as falling in love can be regarded as regression to emotions of infancy and at the same time an important step forwards in the personal development of adults,³⁸ singing provides an instantaneous direct connection between, on the one hand, preverbal and/or nonverbal (infant and/or animal) vocalisation and, on the other, verbal vocalisation, all in the socially constructed cultural environment of a musical genre.³⁹

36. For more on basic differences between singing and talking, see Sparshott (1997),

37. This phenomenon can be still be observed today, even with PA systems in place. For example, announcements in large Victorian railway stations can be very difficult to understand. They are more decipherable in less reverb-rich places like airports and supermarkets, even in cathedrals if decent speakers are placed, as in York Minster, at head height on every pillar in the nave.

38. *Falling in love* (Wikipedia) referring to Gordon (2008: xiv-xv) and Salonia (1991: 58).

39. What passes as 'music' or 'singing' in one culture need not correspond with phenomena labelled similarly in another. For example, if I visit a Muslim-owned corner shop at the time of the evening call to prayer, I hear the recorded muezzin 'singing', but since, according to clerics like Imam Abu Hanifah, singing is *haram* (forbidden, an abomination) [Islamnewsroom.com/news-we-need/493](https://www.islamnewsroom.com/news-we-need/493) [120124], I cannot be hearing 'song'. Muslim definitions of song may, however, be changing, at least if the BBC report 'Istanbul's tuneless muezzins get voice training' (2010-05-11) is anything to go by [news.bbc.co.uk/1/hi/world/europe/8665977.stm](https://www.bbc.co.uk/1/hi/world/europe/8665977.stm) [120124]. In either case, the observation about sociocultural specificity of what passes for song is still valid.

Turning to sonic differences between speech and song, it's possible to make the following five *general* observations about *typical* traits.

1. Singing is more tonal than talking: sung pitches are longer and, if free from wide vibrato, more stable than spoken pitches.
2. When words are sung, vowels (and, sometimes, voiced continuants) tend to become longer while durations of non-continuant, unvoiced consonants remain much closer to those of speech.⁴⁰
3. Sung statements (phrases) tend to be longer and more fluid than those of speech.
4. Disjointed, staccato delivery containing short breaks is less common in song than in speech, while breaks between phrases or periods are generally longer in song than in speech.
5. Singing uses more regular and recurrent patterns of accentuation, metre and periodicity than does speech.

There are of course *hybrid vocal modes* mixing traits from both speech and song. I'm thinking here of four such modes: metric chanting, recitative, intoned chanting and *Sprechgesang*.

1. In METRIC CHANTING speech replaces the tonal traits of song while rhythmic and metric traits of song remain in tact, as in rap, in the scanned slogans of street demonstrations, and in some types of poetry reading.
2. In RECITATIVE (*recitativo* \approx sung solo dialogue in opera or oratorio) the tonal traits of song (fixed pitches) are retained and a full melodic tonal range is in operation but speech rhythm replaces that of song and there is no clear musical metre (*parlando; senza misura*).
3. In INTONED CHANTING, where, as in recitative, speech rhythms dominate and the tonal traits of song are in clear evidence, melodic range is either very restricted (sometimes to just one note) and/or highly formulaic (e.g. consisting of a start motif, a recitation tone and a final motif). Non-metric psalm and canticle singing, synagogue cantillation, as well as Qur'anic recitation and calls to prayer are all examples of intoned chanting. Incantation usually takes the form of intoned chanting.⁴¹

40. Voiced continuants: /m/, /n/ etc.; unvoiced continuants: /f/, /s/ etc; non-continuant voiced consonants: /b/, /d/, /g/ etc. Unvoiced consonants (/p/, /t/, /k/ etc.) consonants are proportionally even shorter because they cannot be sung.

4. In *SPRECHGESANG*, a technique used only by individual voices, pitch range can be extensive, the overall pitch profile of a phrase well defined and the rhythmic patterning more similar to that of song than speech, but the individual pitches of each syllable are unfixed and much closer to those typical of speech.⁴²

To end this section it's worth considering the use of sung tones on certain words in everyday speech. One of the most common examples in standard UK English must surely be the sudden application of sing-song motherese intonation, featuring a descending third⁴³ delivered in a highish register, on to a particular disyllabic in utterances like: 'Baby go *bye-byes!*', 'Oh-oh!', 'That's *naugh-ty!*', 'You'll be *sor-ry!*', '[I] *love you!*', 'Bo-*ring!*' (sing-song disyllabics in italics). This use of over-intoned 'kiddie-speak' can have effects ranging from humorous and childish to rude and patronising. How such effects are created and why they are used would be the subject of another entire book. The point here is that there is a momentary but marked change from normal speech into song, into a demonstrably different vocalisation mode to create a particular effect.

Talking is definitely more common than singing. That's why, when we burst into song, we're adopting a special human mode of vocalisation in a way that to some extent resembles changing clothes for a special occasion. It's in that sense possible to think of singing itself as a vocal costume. Now, there's more to it than that because there's a clear difference between the general 'singing costume' that we've all worn at some time and that of a singer performing for an audience. However, since music semiotics rather than psycho-social role analysis is at the core of

41. Enchantment and incantation both derive from Latin *incantare*, meaning literally to sing (*cantare*) someone into (*in*) another state of mind.
42. *Sprechgesang* (literally = 'speech song') originally meant what I just called *recitative*. *Sprechstimme* was the term Schönberg and Berg would have used to denote what most people today call *Sprechgesang*. Act 3 scene 4 of Berg's *Wozzeck* (1925) contains dramatic examples of *Sprechgesang* (e.g. 'Blut! Das Wasser ist blut!'). *Sprechgesang* is more widely known as the measured talking voice often used in asides by German cabaret artists like Lotte Lenya in *Dreigroschen Opfer* (Weill, 1928). The technique was parodied by Madeleine Kahn in 'Tired' from *Blazing Saddles* (1974).
43. Minor or major third. A descending minor third is the same interval as the ding-dong Friedland door chimes used in US sitcoms c. 1960: 'Honey, I'm ho-ome!'

this book, I'll leave issues of vocal stardom to colleagues in media studies and focus here on vocal costume and persona in terms of links between music as sound and its perceived meanings.

Suiting up for opera

Many vocal costumes used in singing relate to the first definition of COSTUME (p. 360) in the sense of what you wear to carry out a particular task (the 'swimming costume' function). CLASSICAL OPERA SINGING, for example, demands techniques of breathing, diction and phonation allowing the unmiked voice to be projected across the orchestra pit and stalls to reach listeners high up and far away in the opera house balcony. It can take years of training to master these somatic amplification and projection techniques. Inside that tradition there are costume variants like the dramatic soprano, the heroic tenor; and inside, or across, those categories there are idiosyncratic differences of vocal timbre and style letting you distinguish between, say, dramatic tenors like Pavarotti, Domingo and Carreras. If you enjoy and listen to a lot of opera you'll hear those differences instantaneously; if not, you may well hear no more than generic 'male opera singers'.⁴⁴

Although I ought to know better, I've always had a problem with classical opera's dislocation of vocal sound from narrative reality and psychological verisimilitude. I'm thinking here of the following two types of intimate scene. [1] On-stage lovers embrace and perform a duet declaring their undying devotion to each other. This patently private declaration is even more patently public because the soloists belt out the duet for the benefit of listeners fifty metres away in the balcony, not for the narratively realistic 'nearest and dearest' partner who, if the role and situation were real, would surely take offense if his/her beloved were to bellow in his/her ear. [2] A heroine in a small room breathes her last few faint breaths but nevertheless manages to muster maximum lung power to perform a final aria for a large crowd in a large audito-

44. A visual analogy: try distinguishing at a distance between uniformed soldiers in a group. If you don't know the niceties of rank indicated by minor differences on their uniforms, or if you can't see the relevant insignia, and if you don't know the soldiers as individuals, you'd be hard pushed to tell a lance corporal from a second lieutenant, let alone pick out Steve, Dave, Kieran or even Amy from the rest of the company.

rium.⁴⁵ Such operatic anomalies, however silly they may seem, are simply dramatic conventions that cause opera lovers no problem. After all, it could be argued, the sheer power and drama of operatic vocal costume can be heard as congruent with the power of emotions felt in such dramatic circumstances as falling in love or dying: both are in that sense 'worth shouting about'. The anomalies are in fact no more absurd than those of hearing extreme vocal close-ups carrying intimate lyrics that are sung, recorded and broadcast or sold to millions of people all over the world.

So why do I, and many others besides me, accept, without batting an eyelid, Peter Gabriel's dubbing of a whisper on to a full-throated vocal line — the voice simultaneously inside the head and out loud (p.311)— but not opera's way of dealing vocally with the dynamic between internal-private-subjective and external-public-objective aspects of expression?⁴⁶ I think my problem with opera treatment of that duality stems from being born a generation after the invention of coil microphones and the amplification techniques that brought singing voices up close to the ears of individual listeners. Having reached adulthood in the era of multitrack recording, I'm simply used to hearing a vocalist breathe, whisper, croon and so on, not just declaim, exclaim or proclaim. I expect intimacy to *sound* intimate.⁴⁷

The wealth of vocal detail audible, and manipulable, through multitrack recording is a prerequisite for the infinite variety of vocal personas which have become key elements in the aesthetics of popular music. This is a topic to which we'll shortly return (p. 376). Here, though, it serves as an example of how differences in the perception of vocal costume, and, by extension, in the functions and meaning of that costume, can arise. Put simply, lovers of classical opera hear operatic voices as standard VOCAL CLOTHING SUITED TO A PARTICULAR ACTIVITY

45. Examples: the love duet at the end of Act 1 in Verdi's *Otello* (1887); the heroine's death aria 'Con onor muore' in Act 3 of *Madame Butterfly* (Puccini, 1904). For illustration of the problem I have with attitudes to opera singing try the Florence Foster Jenkins collected recordings *Murder on the High Cs* (n.d./2003).

46. See also 'formal' and 'informal' zones of communication in Leeuwen (1999: 27).

47. See under 'Aural staging', p. 299 ff. My problem with opera singing isn't so much a symptom of codal incompetence as of codal interference (pp. 174, 179-185).

(singing opera) and differentiate easily between individuals, both performers and the roles they perform while wearing that vocal clothing, in the same sort of way that a laboratory assistant recognises the different roles and identities of other white-coated individuals working in the same lab. I mean: most of us will just see 'white coats in a lab' and think of, say, microbiology or genetics, unaware that one co-worker, Emily, gram-stains bacteria and likes hill walking, while another, Ryan, model-builds phenotypes and plays cricket. The semiotics of vocal costume are in other words dependent on degrees of familiarity with the real or potential variations of function and meaning inside the sphere of activity linked with the costume in question. Less familiarity and greater distance tend to shift the type of perceived vocal costume from SUITED TO A PARTICULAR ACTIVITY (more familiar) towards SIGNALLING GROUP IDENTITY (less familiar).

Group and genre identity costumes

The GROUP IDENTITY FUNCTION of vocal costume perception is perhaps clearest when vocal styles are heard by unfamiliar ears. In the urban West we often apply ethnic labels to singing styles — 'Arabic', 'Bulgarian', 'Indian', 'Mongolian', 'Native American', etc. as ETHNIC VOCAL COSTUMES, so to speak — because we seem to hear the unfamiliar singing voices primarily in terms of 'other people elsewhere'. That perception of otherness filtered through our own familiar frames of vocal reference tends to make us deaf to variants of style or genre that members of those foreign music cultures hear as distinctive and significant. Indeed, as we saw in the cross-cultural 'death music' experiment (pp. 49-50), we're liable to identify particular functions and meanings in a foreign music culture not with those functions and meanings — FUNERAL and DEATH in that case — but with the foreignness we perceive in the music — AFRICA, ARAB, CHINA, GREECE, INDIA, TURKEY, YEMEN, BAZAAR, DESERT, JUNGLE, etc.

We also tend to project the semiotic norms of familiar vocal styles on to unfamiliar ones. Hearing Bulgarian women singing traditional songs in semitone dyads as harsh and discordant rather than as standard procedure or good-natured fun (pp. 180-182) is one example.☛ Another is

when we talk about the Bollywood GIRLIE VOICE, even though Indian film's most famous female singers were in their seventies when they were still, quite recently, recording vocals for roles lip-synced by actresses in their twenties.⁴⁸ It's also worth noting that Lata Mangeshkar and Asha Bhosle, pre-eminent vocal doyennes of Bollywood, were trained in the Indian classical music tradition. In that tradition a strong, straight high-soprano voice is preferred because it traces a cleaner and clearer melodic profile against the overtone-rich instrumental drones than would a deeper, more mellow vocal tone and timbre subjected to Western-style vibrato.⁴⁹ If that is so, the GIRLIE VOICE notion makes little sense because we're not dealing with a particular female vocal persona (GIRLIE), but with a vocal *costume* SUITED TO A PARTICULAR ACTIVITY, that of presenting the female vocal line in tune and harmony with the drone-filled accompaniment so that the melody is clearly audible.

None of this means that we're 'wrong' to hear Bulgarian semitone diaphony as discordant or Bollywood female vocals as girlish any more than I am to hear operatic voices as tonally blurred, wobbly, loud and generally 'over the top'.⁵⁰ It's just that codal incompetence or interference is in action preventing us from hearing the unfamiliar sort of voice in an unfamiliar setting as we would if it were a familiar sort of voice in a familiar setting.⁵¹ Now, if you find such cultural relativity (or respect) uncomfortable, you might like to consider the work of Alan Lomax and his Cantometrics collaborators who, in *Folk Song Style and Culture* (1968),⁵² documented correlations between vocal style preferences and

48. A Google search for |+Bollywood +voice +("girly voice" OR "girlie voice") produced over 14,000 hits [120119]. As for the GIRLIE = OLD WOMAN issue, see, for example, young actress Gracy Singh lip-syncing vocalist Lata Mangeshkar's rendering of A R Rahman's *O Paalanhaare* in *Lagaan* (2001). Lata Mangeshkar was 72 in 2001.

49. This explanation and another suggesting that clear differences of sound between male and female vocalists was necessary because of mediocre audio playback when films were shown by itinerant movie projectionists in Indian villages are at ask.metafilter.com/168017/Shiva-me-Timbres [120119].

50. To be quite frank, I know for a fact that I'm not the only one to change channels as soon as I hear an operatic soprano on radio or TV. I find the sound overbearing and hysterical, often out of tune (due to excessive vibrato!) and generally unpleasant.

51. Codal incompetence and interference: see pages 179-189.

52. See also thousands of Lomax's recordings, free at research.culturalequity.org [120909].

modes of food production in different types of pre-industrial society in different parts of the world. Their findings describe how, for example, the hunting communities studied in the project tended to show a general preference for a raspy solo male sound, while the horticultural societies seemed more likely to favour mellow mixed-voice chorality. To conclude, *unlike* Lomax and his collaborators, that these observations demonstrate the existence of a universally viable vocal persona for 'the hunter' and another for 'the gardener' would be out of order but some of the project's findings could provide some ideas about crossovers between vocal costume and vocal persona.

Genre-specific vocal costumes

Male singer-songwriters

Fabrizio de André, Wolf Biermann, Jacques Brel, Johnny Cash, Leonard Cohen, Bob Dylan, Serge Gainsbourg, Socrates Málamas, Caetano Veloso, Tom Waits and Atahualpa Yupanqui, to name but a few, are all male singer-songwriters, each with a very distinctive voice. So, what vocal costume, if any, do they all wear that could possibly identify each one as belonging to the same overall genre?

'In the *canzone d'autore* [≈ singer-songwriting], things that might be considered as mistakes of intonation, delivery and bad pronunciation in other genres are accepted as characteristics of individual personality, which is of primary importance in this genre.' (Fabbri, 1982: 67)

Difference and non-conformity can in other words be understood as the singer-songwriter's vocal costume.⁵³ It's a sort of 'anti-uniform uniform' at the opposite end of the spectrum from the relative uniformity of operatic vocal costumes, as well as from that of all those young hopefuls given the Melodyne auto-tuning treatment on TV talent shows like *The X-Factor* (2011).⁵⁴ Being occasionally 'out of tune, or too shy, or too

53. Examples by the artists just mentioned: *Khorakhanè* and *Dolcenera* (Andrè, 1996), *Ermütigung* (Biermann, 1968), *La valse à mille temps*; *Le moribond*; *Ne me quitte pas* (Brel, 1959, 1961, 1972), *I Walk The Line* and *Ring Of Fire* (Cash, 1964, 1963), *Suzanne* and *Hallelujah* (Cohen, 1967, 1984), *Blowin' In The Wind* and *Subterranean Homesick Blues* (Dylan, 1963, 1965), *Les amours perdues* and *Glass Securit* (Gainsbourg, 1961, 1987), *Πιγιγκηπέσα* (Málamas, 2000), *Você é linda* (Veloso, 1981), *Martha* and *Shore Leave* (Waits, 1973, 1983), *Camino del Indio* (Yupanqui, 1973).

“shouty”, writes Fabbri (2005: 145), are vocal traits contributing to the singer-songwriter’s credibility as a ‘real person’, an ‘authentic voice’, a ‘true character’, complete with all the imperfections that inevitably come with every one of us and with our voices. It doesn’t seem to matter if the male singer-songwriter’s voice covers only a limited bass range (e.g. Cash, Cohen, Gainsbourg, Waits), or if he stays in mid register (e.g. Biermann, Dylan, Yupanqui), or if he covers a much wider range (e.g. De André, Brel, Málamas, Veloso). Nor does it matter if he sounds like a ranting preacher (Dylan), or a rueful ruminator (Cohen), or a gruff drunkard on sixty cigarettes a day (Waits), or like a degenerate rogue with little more than a DIRTY OLD MAN GROWL left by way of a voice (late Gainsbourg), or like a wise and simple but enigmatic bard (Yupanqui), or like a full-blooded but vulnerable thinker with a mellow voice that can break out into passionate exclamation (De André, Brel, Málamas).⁵⁵ Almost any voice will work, just as long as the following stylistic conditions are met: [1] the voice is no-one else’s and does not *appear* to conform to norms established through formal training or audio technology; [2] the words are intelligent or enigmatic, thoughtful or provocative, poetic or witty and usually *audible*: the artist’s voice is up front and centre stage; [3] the song, recorded or performed live, should not bear obvious traces of intricate arrangement, orchestration or audio signal processing even if it may well have been subjected to such types of treatment. And the singer-songwriter’s ‘no-frills’ performance, live or recorded, will be even more effective if reinforced by sartorial, behavioural, linguistic and other rules of the genre, especially if the lines between performing and non-performing persona are blurred. With all these attributes the singer-songwriter is easy to identify, not just as an ‘honest artist’ but also as the song lyric’s authoritative and authorial first person (Fabbri, 2005: 145).

54. For demonstration of auto-tuning see *X Factor: How Auto-Tune works* (V1 2010).

55. The thumbnail characterisations come partly from an informal phone conversation I had with Franco Fabbri in January 2012.

Other genre-specific vocal costumes

It goes without saying that other vocal genre costumes exhibit different traits to those of the singer-songwriter. Nevertheless, whether it be a *cantautore*, a *chansonnier*, a *fadista*, a *payador*,⁵⁶ or an opera diva, or female Bollywood singing star; or, in the anglophone world of popular song, a singer-songwriter, a death metal growler, a female gospel artist, a dramatic ballad star, a blues shouter, a crooner, a rapper, a mainstream jazz vocalist, a riot grrl or a folk revival songster, one thing is certain: every one of those different types of vocalist will be wearing some sort of vocal costume identifying him/her with the style and genre in question. As explained earlier, some vocal costumes may exist, at least partly, out of acoustic necessity (operatic voices, the Bollywood 'girlie' voice, intoned chanting etc.),⁵⁷ but every one of the vocal costumes just mentioned will be signalling some kind of genre group identity. 'But', as the advertisers say, 'that's not all'.

If you're familiar with the musical genre and style in question you'll not only recognise the vocal style as a genre costume: you'll also be able to distinguish the voices of individual singers and to recognise differences of vocal persona performed by those singers in those genres. Vocal genre costumes tend to be better suited than others to the presentation of certain types of vocal persona. For example, a death metal growler (e.g. Carcass, 1990) is incompatible with the smooth Mr Nice-Guy sort of persona a convincing crooner can create (e.g. Bowly, 1933); and a crooner, in his turn, would be not be much use as a hoodie gangsta-rapping about 'slappin' up de hoes 'n' bitches' (e.g. Eazy E, 1987), who in his turn would be useless as a 'sincere' lovestruck torch ballad persona (e.g. Houston, 1992), who would make a lousy riot grrl (e.g. Bikini Kill, 1996), and so on.

56. *Cantautore* (Italy), *chansonnier* (francophone world), *fadista* (Portugal), *payador* (Argentina, Uruguay, southern Brazil), *trubadur* (Sweden) are all types of singer sharing many traits in common with singer-songwriters of the anglophone world.

57. See the 'swimming costume' function 'suited to a particular activity', pp. 369-372. Intoned chanting is explained on pp. 367-367.

Grasping vocal persona

Despite the vernacular terms students use, often with considerable insight of lateral (transmodal) thinking, to describe the character of vocal sounds, I've also often registered blank faces in response to questions like 'What does the voice actually express here?' or 'What sort of person is singing to us?' I never interpret those blank faces as a sign of incompetence because I've learnt that all hearing individuals intuitively know, within the same broad music culture, what a voice is communicating and what sort of person is behind it. The blank faces seem rather to express a reticence that probably stems from the discomfort of being asked to verbalise personal impressions of emotions in front of a cohort of fellow students: no-one wants to risk making a fool of themselves by revealing too much of their emotional sensitivity in the company of peers. That peer pressure problem is compounded by the fact that talking about voice in terms like NERVOUS TEENAGER, BARBIE DOLL or SUICIDAL STUDENT isn't regarded as commensurate with the serious or grown-up sort of impression imagined appropriate in the supposedly serious grown-up context of a university analysis seminar. The reticence is in other words a symptom of the dual consciousness in which 'our sense of identity and agency in private is dissociated from whatever sense we may have of ourselves in the public sphere' (p. 2). For while we seem to accept that a successful artist can use voice to express all sorts of intimate, emotional and personal things (private) to millions of listeners all over the world (public), some individuals still find the *verbal description* of feelings and impressions evoked in them by the same artist's voice too personal, too private to talk about 'live', even in front of just a small group of people, and even though those subjective impressions are almost certainly shared by thousands of other human subjects. This contradictory vicious circle of dual consciousness has to be broken in semiotic music analysis. Discussion of vocal meaning tackles the problem head on, as we'll soon see, in a clear and tangible way. So, how can talking about vocal persona help break that vicious circle of dual consciousness? There are, I think, two main ways of approaching the problem, one theoretical, the other practical.

From the theoretical angle it's firstly reasonable to assume that familiarity with issues of dual consciousness (see Preface) and intersubjectivity (Chapter 6) will make the discussion of voice less embarrassing. That's because understanding the socially objective character of subjectivity (intersubjectivity) gives greater confidence in considering personal emotions and impressions in relation to those of others. Secondly, knowledge about psycho-somatic links between voice, mind and body can help liberate notions of subjectivity from their conceptual isolation and bring them out into contact with the external, objective, material world. Here are five broad categories of such links: [1] the vocal behaviour of trauma sufferers;⁵⁸ [2] the vocal characteristics of depression and of Parkinson's disease;⁵⁹ [3] connections between voice disorders and other physical or psycho-somatic conditions;⁶⁰ [4] gender variation and attractiveness in voice quality;⁶¹ [5] personality inference from voice quality.⁶²

Those are all areas in which it's absurd to act as if personal, subjective experiences had no empirically demonstrable connection with external, objective, physical realities.

58. Parson (1999) discusses 'the subjective elements of voice in trauma... for victims of extreme, catastrophic events.' These types of voice capture 'dissociated representational experience... replete with "trauma messages" from the depth of somatopsychic processes, expressed in the patient's "nonverbal talking" in gestures, tone of voice, posture, silences, facial expressions...'
59. See 'Vocal Indicators of Mood Change in Depression' (Hellgring & Scherer (1996); Breslow (2007) on Parkinson's disease; the article *Do you get depressed?* (n.d.), etc.
60. For a selection of books on this broad topic, visit books.google.co.uk/books/about/Understanding_and_treating_psychogenic_v.html?id=ShMBq6LwHp0C [120116]. See also 'Gastroenterological Conditions that can affect the Voice' (Bowen, 2012b), as well as studies of voice disorders in children (Bowen, 2012a; Hooper, 2004).
61. See 'Gender variation in voice quality' (Biemans, 2000) and articles discussing the traits of vocal attractiveness, e.g. DeBruine *et al.* (2005, 2006); see also Hughes *et al.* (2004) and Riding *et al.* (2006).
62. See 'Personality and Voice Inference' itself (Hellgring & Scherer, 1996); Scherer (1987) on the extravert voice, and 'Audiovisual Personality Cues for Embodied Agents' (Krahmer *et al.* 2003), which includes discussion of extravert and introvert vocal types; see also Hughes *et al.* (2004). There are numerous other types of connection between voice, mind, body personality that there is no room for here. More details can be gleaned from perusing the raw text file 'Unsystematic notes from vocal persona sources' at tagg.org/articles/VocPersUnsysNotes.txt [120116].

Turning to the practical side of analysing vocal persona, I've found the following ten simple steps useful in teaching situations.

1. Isolate a short passage in the AO where the vocal characteristics to be studied are particularly clear.
2. Play back that passage as a loop.
3. Listening eyes closed to the repeated loop, use your own voice to impersonate (i.e. to imitate and to appropriate) the vocal sound[s] whose meaning you want to focus on. You don't need to actually sing, just to make the general sound of the voice whose meaning you want to describe. Do NOT sing the lyrics at this stage! The object of this exercise is to understand the *connotative meaning of a vocal sound*, not the lexical meaning of words carried by that sound.
4. When you're reasonably satisfied that the sounds you're making sufficiently resemble the vocal sound in the loop, stop playback but carry on doing your vocal impersonation with your hands cupped round your ears as you continue to growl, moan, chirp, bellow, warble or vocalise in any other appropriate and convincing manner.
5. Still impersonating the appropriate vocal sound, run a quick poietic check. Are you using falsetto, head register or chest register? Is the sound you're producing at all nasal or guttural? Is your voice pitched high, low or in between? Are you using a narrow or wide pitch range? Does the pitch of your impersonation change often, suddenly, gradually, or not at all? Does your vocal impersonation sound loud or soft? Is your breathing short and fast or deep and slow, or in between? If you add words, how is your diction? Muffled and mumbling or crisp and clear? How much of your impersonation is like song and how much like speech?
6. Freeze face and body at some point while impersonating the recorded voice. Is your head held high, hung down, tossed back, leaning to one side? Are your eyes wide open, shut or squinting? Are they cast down, rolled upwards, looking straight in front or to one side? Is your mouth open or shut? Are your lips pursed? What shape is your mouth? Are your teeth clenched? Are your teeth visible? Are your face muscles taut and wrinkled or relaxed? Are you frowning? Is your chin pointing forwards or has your jaw dropped? Is there tension in your shoulders or are they relaxed, or drooping?

Are your arms outstretched, folded, by your side, or held in front of you? Are your fists clenched? Are your hands cupped? Are your fingers stretched and splayed or are they relaxed and together? Are the palms of your hands open and visible or closed and hidden? Do your posture and facial expression fit better with standing, sitting, kneeling, lying, walking, running, etc.? In short does anything in your facial and bodily expression correspond to any particular emotion, state of mind or attitude?

7. What words best fit the vocal sound you're imitating? Is it any of these? I LOVE YOU. I HATE YOU. LIFE IS POINTLESS. THIS IS FUN. I'M BORED. DON'T MESS WITH ME! DON'T YOU THINK I'M SEXY? I'M A CREEP. I'M COMING TO GET YOU. COME CLOSER! GO AWAY! YOU'RE GORGEOUS. YOU'RE STUPID. THIS MAKES ME LAUGH. I DESPISE YOU. I'M SICK OF IT. I'M WORRIED. I'M TERRIFIED. I WON'T GIVE IN. I DON'T CARE. THIS IS FANTASTIC. What words sound ridiculous or are impossible to say with the facial expression and body posture you've adopted to produce your impersonation? If there are lyrics, how does their meaning fit with the words you think best correspond to the vocal sound?
8. What sort of person (age, gender, nationality, occupation, etc.) might typically be *talking* in that way? Is it a lover, sister, brother, teacher, preacher, best friend, enemy, trickster, philosopher, or any of those listed in section 3 of the table on page 356? Or is it someone or something completely different? Perhaps it's an animal or a machine? Who might the vocal persona you're imitating be addressing? Him/her/itself or someone else? Just one other person, or several, or many? What sort of relationship could there be between the vocal persona and whoever they're addressing?
9. Where is the voice you're impersonating most likely to be heard? Indoors, outdoors or inside your head, or all three? In a bedroom or a church? In a bar, car or club, or at school? In the street or countryside? At the far end of a long corridor or breathing in your ear?⁶³
10. What words best describe the vocal sound you're impersonating? Is it any of the concepts shown in Table 10-1 (pp. 356-357)?

63. See pp. 298-303 for discussion of spatial parameters in music.

The main value of this ten-step exercise is that it tangibly relates non-verbal vocal sound with other types of expression inside the listening subject. Vocal impersonation concretises the attitude and emotional state of the voice under analysis. The exercise provides direct access to the identification and meanings of vocal persona and makes it easier to overcome the negative effects of dual consciousness.

And finally: parody

If, despite the tips just presented, the task of denoting vocal persona still seems difficult or embarrassing, why not try some humour? Just look on line for parodies of the sort of voice you're struggling to describe. Parody involves the humorous exaggeration of stylistic traits which, like caricatures, become larger than life and which make salient features of the style and genre extremely clear. Here are a few examples of vocal persona parody that I found useful while putting this chapter together: [1] Reggie Watts's rap spoof *Fuck Shit Stack*, his 'Irish folk ballad' *Fields Of Donegal* (both 2010a), and, sharpest of all, *Big-Ass Purse* (2010b); [2] vocal-instrumental gags by Bill Bailey, for example his Bryan Adams lampoon *Hats Off To Zebras*, or his Billy Bragg parody *Chip Shop*, or *Dr Qui*, the 'Jacques Brel/Belgian jazz' version of the *Dr Who* theme (all 2000); [3] Jon Lajoie's boy band parody *Pop Song* (2009), complete with obligatory rapper for 'a slice of the urban market' and a verse for the 'gay voice to let you know I'm sensitive'.

Then there are the acrobatic, ecstatic, post-gospel princess caricatures in Nile Rodgers' 'Soul Glo' spoof ad in *Coming to America* (1988) and in Stevie Van Lange's orgasmic 'Whoaa!' for the 1993 Bodyform TV ad (Tagg, 2008c). Add to that the looped coloratura phrase from The Queen of the Night's aria in *The Magic Flute* (Mozart, 1791) set to visuals of 'perfectly groomed young women in the back-arching, pupils-dilating throes of carnal abandon' (Service, 2008) for the Durex Play-O TV advert (2008), and you have a fascinating but gender-politically disturbing can of semiotic worms that should be, if it isn't already, the subject of a complete book discussing 'auditeurism', the audio equivalent of voyeurism (see Corbett and Kapsalis, 1996).⁶⁴ I can't deal with any of that here but the issue certainly suggests that the power of vocal persona should never be underestimated.

Of course, musical parody isn't just limited to the humorous exaggeration of vocal traits. Vocal, instrumental and compositional style are parodied in different ways by different entertainers who draw larger-than-life musical cartoons of sounds you may need to describe in your analysis. Therefore, to end this chapter on a usefully frivolous note, I take the liberty of listing a few artists from the anglophone world whose musical parodies might be useful if you need to pinpoint style-specific musical traits. In addition to the instrumental as well as vocal mannerisms parodied by Reggie Watts, Bill Bailey and Jon Lajoie (p. 380) those few examples would be Dudley Moore (1961), Peter Schickele (1967, 1971), Stan Freberg (1957) and Frank Zappa (1965, 1967, 1981). I would also recommend mockumentaries like *The Rutles* (1978) and *Spinal Tap* (1984), as well as sketches from *Mad TV*, not to mention style-specific novelty songs such as *Disco Duck* (Dees, 1976).⁶⁵ Finally, different stylistic versions of the same tune automatically draw attention to parameters of musical expression like instrumentation, vocal persona and aural staging that can be missed when the melodic line and its lyrics are the main focus of interest. One striking set of multiple examples of the same tune was broadcast in the Australian TV series *The Money or the Gun* (1989-1990). It featured a radically different version of *Stairway To Heaven* (Led Zeppelin, 1971) every week over a six-month period (see *Stairways to Heaven* 1992).⁶⁵ Who said music analysis was a drag?

64. How does this issue relate to: [1] songs like *Black Snake Moan* (Spivey, 1926) and *Moanin' Low* (Holman, 1929); [2] religious ecstasy and black female gospel singers from Bessie Johnson (e.g. 1927) to Mahalia Jackson (e.g. 1947); [3] the secularisation of gospel female vocal style into soul (e.g. Aretha Franklin, 1967) and its overt sexualisation in disco (e.g. Donna Summer's 1975 *Love To Love You Baby*)? And what about Gloria Gaynor (e.g. 1978) returning to gospel from disco? Then, what about opera vocal acrobatics, *les danseuses du Corps de ballet* and the simultaneous privatisation and prostitutionalisation of Covent Garden and *L'Opéra de Paris* in the nineteenth century? Another question: what link is there between the diva as vocal persona and the diva (male or female) as typical of the 'narcissistic-aggressive type' (Benis, 2005) and, if there is any, what could that tell us, if anything, about gender relations? Finally, why are the male's orgasmic grunts and yells seemingly so much less interesting than the sounds of a sexually aroused woman when it comes to selling a recording, a performance or any other product?

-
65. • On *Beyond The Fringe* (1961) Dudley Moore parodied: [1] Benjamin Britten's song-writing and Peter Pears' tenor voice in *Little Miss Muffet*; [2] an inter-war German cabaret song complete with *Sprechgesang* passages (*The Weill Song*).
- Peter Schickele, alias P D Q Bach (1967, 1971), recorded the *Schleppet in Eb Major* (classical chamber music gone mad), *New Horizons in Music Appreciation* (Beethoven's Fifth Symphony with sports commentary), *What's my Melodic Line?* (the same one-chord extract from fictitious *Concerti grossi*), cigarette commercials as Purcellian ground-bass arias (*Do You Suffer?*, *If You Have Never*), the *Toot Fugue* (Bach fugue version of the *Volga Boatmen* song on calliope), the cantata *Iphigenia in Brooklyn*, fake madrigals, *The Stoned Guest* (half-act opera) and *The Seasonings* (oratorio).
 - Stan FREBERG (1957) lampooned 1950s teenage pop in *Heartbreak Hotel*, *The Great Pretender* and *Rock Around Stephen Foster*.
 - Large parts of ZAPPA's early recordings were devoted to parody, for example 'Wowie Zowie' and 'You Didn't Try To Call Me' on *Freak Out!* (1965) and virtually all of *We're Only In It For The Money* (1967); see also *You Are What You Is* (Zappa, 1981).
 - *Spinal Tap* (1984) is a rock mockumentary which spoofs every conceivable aspect of heavy metal while Neil Innes' Beatles pastiches for *The Rutles* (1978) are so convincing that they have overwritten several of the Beatles originals in my head.
 - *Mad TV* have parodied Britney Spears (*Lick My Baby Back Behind* [¶¶JkXuckuJ0w](#)) and Shakira (*Whatever Don't Matter* [¶¶w8QH93jWZbk](#)).
 - French and Saunders (2002) spoofed Alanis Morissette in *Aimless Morris Minor*.
 - 'Frank Satsuma, the Japanese crooner' (n.d.) parodied Frank Sinatra (and Japanese pronunciation of English) in *I Want You To Get Under My Skin* ([¶¶j4dHPS8gvLA](#)).
 - Peter Sellers (1958) sent up British 'folk' mannerisms in *Suddenly It's Folk Song*.
 - The funniest or most convincing 'Stairways to Heaven' in the Australian TV show were, I think, by Rolf Harris, The Australian Doors, The Beatnix (Beatles tribute band), The Whipper Snappers (à la Bangles/Pretenders), The Fargone Beauties (bluegrass) and Vegimite Reggae.
 - S T Sanders' 'band shreds' are hilarious remusicalisations of rock videos (e.g. Rolling Stones, Eagles, Bruce Springsteen) [¶ stsanders.com/www/pages/videos.php](#) [121018].
 - *Flight of the Conchords* (2007-2009) was a TV series based largely on musical parody.
 - I'd also like to recommend a spoof ad for the fictitious 'Best of' album *Arnie Schöenberg and his Second Viennese School* (1977); but it's not really relevant here.