

## Vocal pop pleasures. Theoretical, analytical and empirical approaches to voice and singing in popular music

**Martin Pfeleiderer**

[martin.pfeleiderer@hfm-weimar.de](mailto:martin.pfeleiderer@hfm-weimar.de)

**Institute for Musicology, Liszt School of Music Weimar / Friedrich Schiller University Jena, Germany**

### Abstract

Voices and vocal expression play a crucial role in popular music. Voices are linked immediately to the human body and various affective states. But vocal expression could be conventionalized and manipulated deliberately by learned cultural patterns of expression, too – in everyday speech as well as in music. Two approaches to investigate voice and singing in popular music are outlined: Firstly, an analytical approach is illustrated by a characterization of a vocal performance by Ray Charles using spectrograms and transcription. Secondly, an empirical approach to the effects of vocal performances is illustrated by listener surveys asking for ascriptions of emotions and personality traits to music excerpts.

**Key words:** voice, singing, methodology, emotions, personality, psychology.

### Introduction

Since songs stand right in the centre of most popular music genres, voices and singing play an outstanding role in popular music. Songs offer a direct access to the music through the topics of the lyrics on the one hand, and through singing and vocal expression on the other hand. While only few listeners are able to play an instrument, everybody has a voice and can sing or hum along with the music. Though voices and singing are an intriguing subject for popular music research, there are only few investigations published on the topic by now.<sup>1</sup> My intention is to present some perspectives for future research on voice and singing in popular music. The text falls into three parts. After theoretical considerations, I will illustrate an analytical approach to vocal expression with a music example, an early song of Ray Charles. Finally, I will outline the design and some preliminary results of an empirical investigation of the interrelationships between vocal expression and the impression or ascription of affects, emotions and personality traits.

## Voice, body, personality – some theoretical considerations

In his seminal study of love song lyrics in the 1950s Donald Horton draw the following conclusion:

They [love songs] offer the opportunity to experiment in imagination with the roles one will have to play in the future and the reciprocal roles that will, or should be, played by the as-yet-unknown others of the drama. [...] it may be function of the popular singer, in dramatizing these songs, to show the appropriate gestures, tone of voice, emotional expression [...] for transforming mere verse into personal expression. The singer is at the same time available as an object of vicarious identification or as a fancied partner with whom in imagination the relationships and emotions of the future may be anticipated (Horton, 1957, p. 183).

What Horton writes about love songs and how adolescents use them to find their roles in gender relations could be generalized to most genres of popular music: For many listeners not so much the contents of the lyrics but rather their expression with an appropriate tone, sound and gesture stands in the focus of experience – or as Simon Frith put it: “Song words are not about ideas (‘content’) but about their expression” (Frith, 1996, p.164). Vocal expression is comprehended and understood in vocal utterances of others independently from linguistic contents. Therefore voices stand right in the centre of popular music: “Voices, not songs, hold the key to our pop pleasures” (Frith, 1996, p. 201).

In order to find an explanation for this significance of voices, first one has to look at the voice as a part of the human body. Voices refer to the physical expression of sensitivities, affects, needs and desires. As bodily beings listeners are able to comprehend immediately the emotional expression of voices and even to imagine the personality of the singer. In many respects, human vocal expression has preserved the emotional signal character that it possesses among animals. Think for example of howling and bawling as an expression of aggressiveness, screaming as an expression of fear, shouting as an expression of anxiety and pain, friendly grunting and laughing as an expression of well-being, or crying and weeping as an expression of sadness. Besides, larynx and the lower jaw are a kind of an emotional valve that holds back incriminating and scary emotions – for example by pressing teeth and lips together for annoyance and sorrow, or by holding back breath for fear. Another foundation of vocal expression is the pre-linguistic communication between parents and little children. By imitating the vocal expression of their parents little children learn to express moods and emotions. Moreover, screaming and babbling are two poles of the infant’s relation to its own voice – screaming as an extrovert behaviour that sends vital, even aggressive impulses, and babbling, as a joyful introvert and even auto-erotic play with relaxed lips, often in sucking position.

Later in life, singing yields pleasure too: pleasure by playing with one’s own voice as well as enjoyment of listening to vocal utterances of others. Moreover, while singing along with a singer and a song – silently or loudly – we learn to identify ourselves with the vocalist, with her/his emotions and personality.

We make her/his song our own song and her/his voice our own voice. Therefore, one can comprehend listening to a singing voice as a kind of close interaction with a high degree of affective involvement.

In general, affects come along with involuntary body changes: for example, alteration of breath frequency and muscle tension, which conversely influence vocal utterances. According to the psychology of emotions these immediate effects are called “push“-effects (Juslin and Scherer, 2008). But vocal expression can be controlled and employed strategically, too, resulting in so called “pull“-effects. With conventionalized stereotypes of vocal expression utterances could be made more comprehensible. Stereotyped vocal patterns could serve as means of self-presentation in order to intentionally create a certain image or effect – and even in order to deceive others.

As in everyday speech, singers can deliberately choose vocal means in order to express certain emotions and characters. Thereby they could fall back upon conventionalized cultural patterns of vocal expression which are broadly understood. According to Clyde Kluckhohn (1962, p. 73) cultural patterns are “patterns, explicit or implicit, of and for behavior acquired and transmitted by symbols”, for example by certain vocal cues. Relying on such patterns and cues, singers are able to express not only a certain mood and emotion, but to create with their voices a certain personality or image, for example the vocal image of an angel or a child, a nymph or a virgin, a storyteller (griot, folk singer, blues singer, rap MC), a seducer, a preacher, the voice of religious or erotic ecstasy, a frenetic voice (singing in a buzz), a diabolic voice, a prosaic voice of the working man, a political voice (educational, accusing, or intellectual) and many more. As Philip Tagg observed several years ago, many if not all of these patterns are consistently inferred by listeners from vocal recordings – at least by listeners of the same age and same cultural background. Tagg names those cultural patterns “vocal persona”. He writes:

When exhorted to speak the words of a particular vocal line emulating its approximate pitches, dynamics, timbre, duration, accentuation and rhythm, students are not slow to come up with words which pinpoint what might be called a *vocal persona*. The kinds of vocal persona heard in analysis classes tend to receive such vernacular epithets as: ‘little girl’, ‘nice boy’, ‘complete bitch’, ‘utter bastard’, ‘desperate rebel’, ‘macho git’, ‘sexy whisper’, ‘confident woman’, ‘evil child’, ‘worried man’, ‘death voice’, ‘voice of Satan’, ‘hound of hell’, ‘nervous hiccuping teenager’, ‘angry scream’, ‘Harpie shriek’, ‘frustrated squawk’, ‘bedside discoman’, ‘football hooligan’, ‘one of the lads’, ‘nudge-nudge, wink-wink’, ‘friend and confidante’, ‘tired and abused’, ‘resigned’, ‘depressed’, ‘demoralised’, ‘cynical’, ‘angry’, ‘hysterical’, ‘out of breath’, ‘stressed out’, etc. The list seems virtually endless and intersubjective agreement in analysis classes about the connotative traits of the vocalist(s) and vocal line in question is usually very high (Tagg, 2009, p. 13).

However, in many popular music genres listeners do not appreciate stereotypes and conventions, but rather an “undiluted”, immediate personal expression, something like a “personal voice” of a singer. Therefore listeners

tend to distinguish between deliberately chosen vocal patterns and strategies of expression on the one hand, and the “personal voice” of a vocalist on the other hand. Then, the personal voice is conceived as something staying constant in different performances of a singer, depending on her or his specific voice qualities due to individual anatomical characteristics. In addition, there are persistent personal dispositions and habits of phonation and articulation as well as regionally and ethnically socialized voice-settings – something like “biographical traces” of a voice. Admittedly, often a “personal voice” couldn’t be separated from certain effects of sound manipulation typical for a singer, for example a certain reverb or delay (think of the slap-back echo connected with Elvis Presley’s voice in his early recordings), and from certain cultural patterns or images which a singer personifies. Therefore, one could comprehend a certain voice as a net of affects, emotions and personality traits as well as their stylized vocal expression, of physiological peculiarities and biographical traces as well as of conventionalized cultural patterns (Bielefeldt, 2008, p. 212).

In the following, I will outline two approaches to explore voices and vocal expression in popular music: an analytical approach in order to characterize vocal styles and an empirical approach in order to explore the affective effects of voices.

### **Analytical approach to vocal expression – Ray Charles “Come Back, Baby” (1954)**

The intention of an analytical approach is to describe typical features and peculiarities of a vocal style – be it a personal, a regional or a historical style. Roughly speaking, there are three basic dimensions for musical analysis of a vocal style: voice quality, melody, and rhythm. Voice quality, or the sound of a voice, could be divided in features due to phonation, which is the typical shape of the oscillations of the vocal folds, and peculiarities due to articulation according to specific resonances of the vocal tract (Laver, 1980; Sundberg, 1988). Besides neutral phonation there are several other characteristic kinds of phonation used in popular music: for example, a breathy phonation, a hoarse and rough phonation, a creaky phonation, and falsetto. There are several kinds of articulation, too: for example, a nasal articulation, a pressed or a relaxed articulation. Moreover, there are regional peculiarities of articulation along with regional accents, special pronunciation etc. Additionally, there are technological means to shape the sound of a voice, for example the use of microphone reverb and delay, doubling of voices, harmonizer etc. Often voice qualities are connected to certain singing techniques e.g. belting. Melody could be described according to the melodic contour of a phrase, specific choices of tones and intervals, and relations of melodic lines to the harmonies of the accompanying instruments. Additionally, several means of ornamentation play a crucial role for melody: for example, slurs and glides, melismas, glissandi and peculiarities of intonation etc. Rhythm encompasses durations and accentuations of notes, the use of certain rhythmic patterns and of syncopations or offbeat-beats, the

prevailing tempo, the relation of the sung rhythm to meter and rhythmic patterns of the accompanying instruments, and last but not least the peculiarities of micro-rhythmic play or the degree of metric exactness. Of course, additional aspects have a strong impact on vocal performances and therefore have to be considered, too: the lyrics, their topics and dramaturgy as well as the context of the song, the performance, the recording and the visualisation (video clip, cover art) and the biographical background of the singer or band. But for now, I will focus on voice quality, melody, and rhythm and on two possibilities to represent those features graphically.

Let me illustrate the analytical approach to singing with an example, a song by Ray Charles – one of the most influential singers in the history of popular music. Charles combines means of vocal expression from several genres and traditions: from blues and gospel as well as from jazz and country music. “Come Back, Baby”, recorded in November 1954, is an early recording of Ray Charles and was published as back side of “I’ve Got a Woman”.<sup>2</sup> It’s a slow twelve-quaver-ballad, the tempo is about 45 bpm. The built-up of the song is simple: there are four eight-bar verses each with a short refrain. Verse four is accompanied by a stop chorus of the whole band. While some piano figures could be associated with gospel music, the lyrics refer to a typical blues topic. The protagonist has been abandoned by his baby and he’s now furious about her. But at the same time the lyrics have another turn, since lines like “let’s talk it over one more time” and “well I admit, baby, that I was wrong” are rather unusual for blues lyrics and foreshadow some new attitudes in gender relations – though, in the last verse the singer threatens “it won’t be long, (...) and I’ll be gone”.

Ray Charles uses a variety of voice qualities and singing techniques. Often he slightly presses certain syllables. At times he employs a deep, almost creaky voice, and at times his voice snatches up to falsetto, particularly at the end of phrases (for example “love you”, “know-ow” in the first verse). He regularly screams and shouts, for example “yeah, yeah” before “let’s talk it over”. In a spectrogram of that line (fig. 1)<sup>3</sup> the screamed “yeah” could be distinguished very well from the sung “yeah”. In general, spectrograms offer helpful options to represent timbre and voice quality, and they are easy to read and understand: the relative greyness or blackness of frequency components (vertical axis) represents the relative loudness of those frequencies in the overall sound over time (horizontal axis). Horizontal lines represent the tones played or sung. Normally there are several parallel lines, which represent the harmonics of one tone. The more and the higher the harmonics are, the brighter is the timbre of a certain tone. Vertical lines and clouds of greyness represent noise that is characterized by a wide frequency spectrum. For example all the strokes of a drum set are represented as vertical lines. In the excerpt represented in figure 1 the first “yeah” has a rough and noisy quality, represented by clouds of greyness over a frequency band until 4 kHz, whereas the following sung “yeah” has a wider harmonic spectrum represented by parallel lines. Because of that sudden change of voice quality, it should be clear that Charles is not just overwhelmed by his feelings while singing, but that he controls his vocal expression and deliberately changes between noisy shouts and perfectly sung phrases.

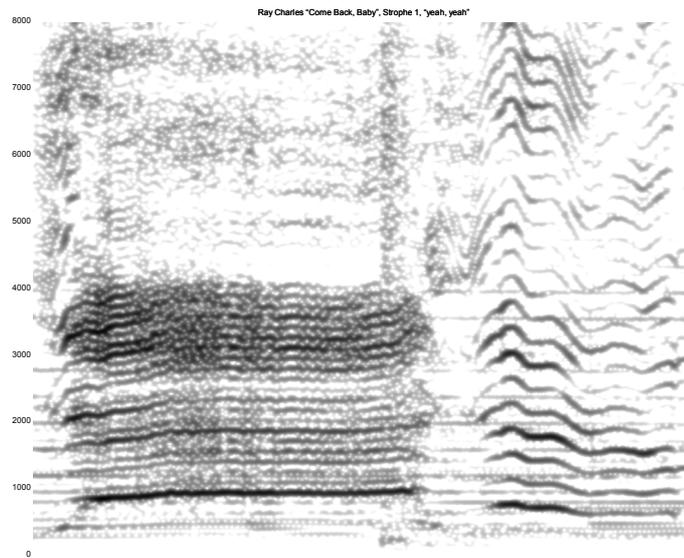


Fig. 1: Spectrogram of an excerpt from Ray Charles “Come Back, Baby” (1954): roughly (left) and neutrally (right) sung “yeah”.

Moreover, there are falsetto-shouts (“huh”), too. As it could be seen in the following spectrogram (fig. 2), falsetto notes (on the left) have less high harmonics than the following tones sung with a neutral phonation.

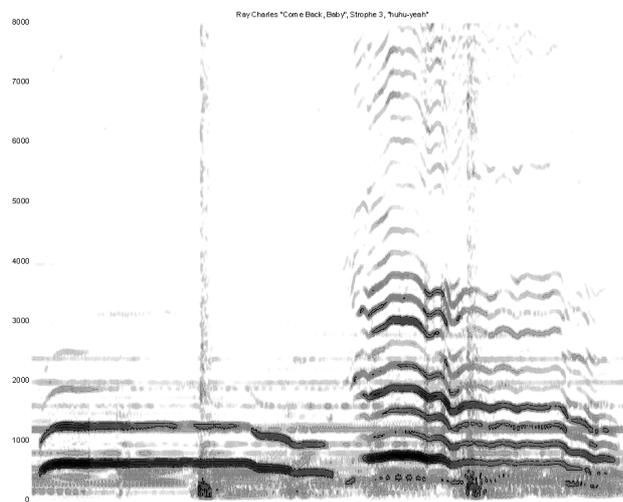


Figure 2: Spectrogram of an excerpt from Ray Charles “Come Back, Baby” (1954): falsetto “huhu” (left) and neutral phonation “yeah” (right).

In the last example (taken from the third verse), Charles uses falsetto phonation with a high degree of roughness or hoarseness (fig. 3). Again, he switches deliberately between two different techniques of phonation: first he

screams with a rough falsetto voice, then he sings with a slightly rough voice without falsetto phonation.

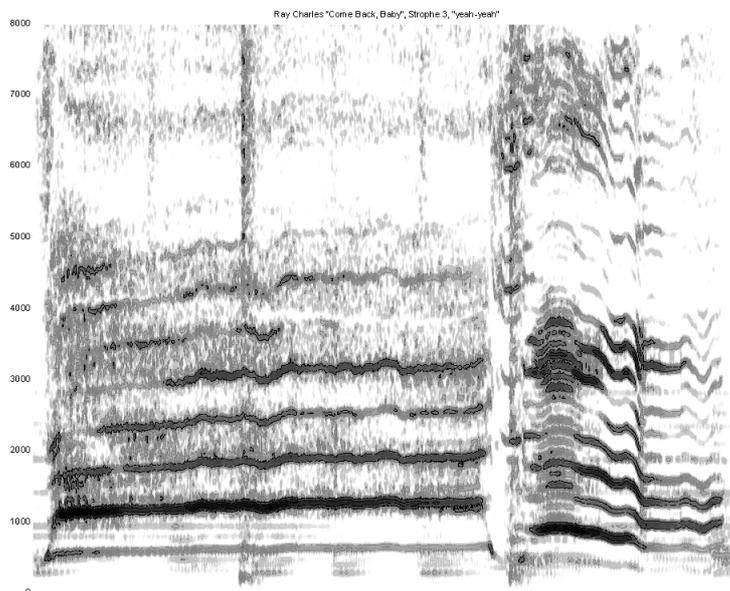


Figure 3: Spectrogram of an excerpt from Ray Charles “Come Back, Baby”: rough phonation of the syllable “yeah” with falsetto (left) and without falsetto (right).

In order to analyse and to visualise rhythmic and melodic features of Charles’ singing, I transcribed the first verse of “Come Back, Baby” (fig. 4). Although, transcriptions are always just approximate representations of the real sounds, transcribing is very helpful to delve into the details of a performance as well as to represent and to communicate the melodic and rhythmic structure (Winkler, 1999).

oh come back ba - by oh mama please don't go yeah for the way I love  
 you child you never know oh o-oh come back ba - by yeah yeah let's talk it  
 o-ver oh one mo-ti-me oh, no-no

Figure 4: Ray Charles “Come Back, Baby” (1954), transcription of the first verse.

Charles regularly accentuates and pushes up syllables just before the beat is laid down by the band. This striking offbeat-phrasing gives a certain drive and urgency to the whole vocal performance. Moreover, note the micro-rhythmic peculiarities of the vocal performance – in particular the ornamental lines freed from a metric reference, such as those, for example, in line two (“oh, mama please don’t go”). Here, the transcription could only give a vague impression the micro-rhythmic accelerations actually articulated by Charles. In general, the melodic lines are descending towards the end of a phrase. The lines are diatonic, even pentatonic most of the time, but with an unstable third – I often didn’t know whether to write a major or minor third (g or g-flat). In general the high degree of melismatic ornamentation at the end of each phrase is striking. Ornamentation could be visualized very well with a spectrogram, too. In figure 5 one can see very well that the singer is steplessly slurring and sliding between pitches.

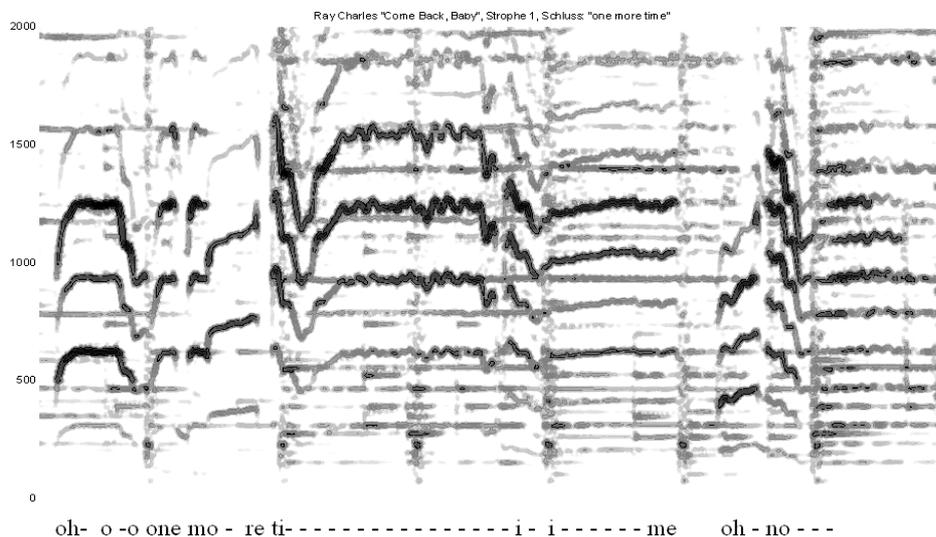


Fig. 5: Ray Charles “Come Back, Baby” (1954), spectrogram of the last line of the first verse.

It’s an intriguing challenge to analyse and describe singing and vocal expression of various vocalists in order to come to a history of vocal styles in popular music. Alongside with the establishment of appropriate methods of analysis and representation, there are several issues that still are to be investigated, for example: which vocal means represent a style of singing? Are certain singing styles congruent with certain styles and genres of popular music? And how could a personal vocal style or the personal voice of a singer be characterized?

## Vocal effects – ascriptions of emotions and personality traits to singers

Beyond that, there is another intriguing question concerning the reception of popular voices: which affects – emotions, moods, dispositions of personality – are actually communicated by a vocalist? Or to put it another way: which affective states and personality traits are ascribed to the vocalist by listeners? To approach this issue, I conducted a listener survey. I played the first verse and refrain of “Come back, Baby” sung by Ray Charles to fourteen participants of an introductory course in music psychology at university of Basel, Switzerland, and asked them: “Please listen attentively to the voice in the music example. Which effects has the singer in your impression? He is ....”, followed by forty-four adjectives. For each adjective one had to rate her/his degree of agreement between “not at all” (1) and “very much” (5). I selected the forty-four adjectives in two steps. First, I took all adjectives for basis emotions and many (not all) adjectives for personality traits from the personality circle of Hans Jürgen Eysenk as well as from the five-factor-model that is common in the psychology of personality (Zimbardo and Gerrig, 2004, pp. 606-608). Then, from a list of about one hundred adjectives all those items were excluded, that were hardly applicable for music, voice and singing in the opinion of the participants of a pre-test. Moreover, some adjectives were added, which participants of the pre-test missed, e.g. ironical, erotic, and seductive. Table 1 shows the results of the survey. The adjectives are listed according to the mean degree of agreement (from five to one).

Table 1 – Ascriptions of affective states to the first verse of “Come Back, Baby” by Ray Charles.

ascriptions of affects	N	mean	standard deviation
<i>aus sich heraus gehend</i> (going beyond one's constraints / letting one's hair down)	14	4,36	,842
<i>ehrlich</i> (honest)	13	4,08	1,115
<i>natürlich</i> (natural)	14	3,86	,663
<i>nachdenklich</i> (contemplative)	14	3,86	1,027
<i>energiegeladen</i> (energetic)	14	3,79	1,311
<i>niedergeschlagen</i> (depressed)	14	3,79	,699
<i>offen</i> (openminded)	13	3,77	1,013
<i>durchsetzungsfähig</i> (assertive)	14	3,71	,994
<i>mitfühlend</i> (compassionate)	14	3,71	,726
<i>verzweifelt</i> (despaired)	14	3,64	1,216
<i>freundlich</i> (friendly)	14	3,50	,760

<i>erregt</i> (aroused)	14	3,50	,941
<i>lässig</i> (easy-going/cool)	14	3,50	1,160
<i>impulsiv</i> (impulsive)	13	3,46	1,050
<i>aktiv</i> (active)	14	3,43	1,158
<i>friedlich</i> (peaceful)	13	3,38	1,121
<i>traurig</i> (sad)	14	3,29	1,139
<i>verführerisch</i> (seductive)	13	3,15	1,068
<i>ruhig</i> (calm)	13	3,15	1,214
<i>beherrscht</i> (restrained)	14	3,14	,770
<i>kontrolliert</i> (controlled)	14	3,00	,877
<i>intellektuell</i> (intellectual)	10	3,00	,816
<i>emotional instabil</i> (emotionally instabile)	12	2,92	,793
<i>launisch</i> (capricious)	14	2,71	1,139
<i>ausgeglichen</i> (balanced)	14	2,71	,914
<i>euphorisch</i> (euphoric)	13	2,69	,947
<i>erotisch</i> (erotic)	13	2,69	1,251
<i>optimistisch</i> (optimistic)	13	2,62	1,193
<i>freudig</i> (glad)	13	2,46	,877
<i>wütend</i> (angry)	14	2,36	1,082
<i>ironisch</i> (ironical)	12	2,33	1,371
<i>glücklich</i> (happy)	13	2,31	,751
<i>ängstlich</i> (anxious)	14	2,29	1,139
<i>oberflächlich</i> (superficial)	14	2,14	,770
<i>unintelligent</i> (not intelligent)	10	2,00	,816
<i>reserviert</i> (reserved)	14	1,93	,616
<i>aggressiv</i> (aggressive)	14	1,86	,663
<i>schüchtern</i> (shy)	14	1,86	,663
<i>gleichgültig</i> (indifferent)	14	1,79	,975
<i>gelangweilt</i> (bored)	14	1,79	,802
<i>sorglos</i> (carefree)	14	1,64	,497
<i>verachtend</i> (contemptuous)	13	1,54	,519
<i>kalt</i> (cold)	14	1,50	,650

The singer of “Come Back, Baby” was appraised as someone who is to a high extent going out of one’s constraints or letting one’s hair down (*aus sich herausgehend*), who is honest (*ehrlich*), natural (*natürlich*), energetic (*energiegeladen*), and who is contemplative (*nachdenklich*) and depressed (*niedergeschlagen*), too. (Ray Charles was not mentioned during the survey and actually only two participants stated that they have recognized him.) Therefore one could conclude, that probably different emotions and affects could be expressed in one and the same performance of a song – affects that even could contradict each other, such as, for example, energetic and depressed. At the same time the singer was not or only weakly rated to be cold (*kalt*), contemptuous (*verachtend*), carefree (*sorglos*), bored (*gelangweilt*), indifferent (*gleichgültig*), shy (*schüchtern*) or aggressive (*agressiv*). In light of differing musical preferences and listening biographies among the participants of the survey the accordance of their appraisals was astonishingly high – with mean deviations between 0,335 (cold) und 1,300 (despaired) on the five-point scale. Therefore, in case of the Ray Charles song, listeners do have a general tendency to infer from singing to emotions and personality traits they ascribe to the singer.

This raises another question: which features of singing exactly yields certain ascriptions and ratings? According to which vocal cues listeners conceive and ascribe a certain emotion or personality trait to a vocalist? In order to get some hints for an answer to this exciting issue I designed another listeners survey that now falls into two parts. First, an analytical checklist for the analysis of voice and singing in popular music was designed. Almost half a century ago, Alan Lomax designed something similar with his cantometric scales (cf. Lomax, 1962, 1968). Lomax’ cantometric scales were a starting point, but in long discussions with my students the items proposed by Lomax were modified and complemented. The resulting checklist included thirty-five vocal features with items for several voice qualities, for rhythmic and melodic features, for dynamics and ornamentation as well as for recording facilities like reverb and delay. Then, excerpts from ten songs, each lasting about thirty seconds, were selected. Intentionally, a variety of styles were chosen in order to get a broad overview of different means of vocal expressions in various popular genres.<sup>4</sup> Six experts for popular music singing – actually some of my students and me – rated the excerpts according to the items of the checklist. In a second survey the ten music excerpts were rated by thirty-six students in a course on popular music and media at university of Paderborn, Germany, according to the list of adjectives from the questionnaire mentioned above.

The overall intention of the pilot study was to test empirical methods in voice research and at the same time to formulate some hypothesis concerning the relationship between vocal expression and its effects on the listeners articulated in ascriptions of affective states and personality traits. Therefore, I computed multiple linear regression models for all the forty-four adjectives in relation to the thirty-five vocal features using the means of all ratings of the two questionnaires. The aim of regression models is to predict the best combination of weighted independent variables or predictors – here: features of vocal expression – which yields an effect on a dependent variable, in our case an affect

ascription or the mean of ascriptions by several listeners.<sup>5</sup> In other words: we are looking for the best combination of vocal features to predict a certain ascription. In table 2 some of the regression models are shown. There are four columns: the ascription on the left, then the voice cues or vocal features, their standardised weightings and, last but not least, a measure for the overall quality of the regression model. The quality of each prediction or regression model could be characterised by the percentage of the overall variance the model explains in comparison with the remaining variance, that has to be ascribed to other, yet unknown determinants. However, that quality measure is not meaningful for models with more than one predictor, since it is distorted by correlations or multicollinearity between the predictor variables.

Table 2: Regression models for several ascriptions.

ascription	vocal feature	weighting	explained variance
aggressive	shouting/screaming	b = 0,85	r <sup>2</sup> = 0,68
seductive	breathy phonation	b = 0,91	r <sup>2</sup> = 0,81
erotic	breathy phonation	b = 0,85	r <sup>2</sup> = 0,68
assertive	vocal effort	b = 0,88	r <sup>2</sup> = 0,75
bored	vocal effort	b = -0,91	r <sup>2</sup> = 0,80
shy	vocal effort	b = -0,89	r <sup>2</sup> = 0,76
restrained	shouting/screaming rubato rhythm	b = -0,84 b = -0,31	[r <sup>2</sup> = 0,92]
aroused	phrase ending high fast tempo neutral phonation pressed	b = 1,31 b = 1,27 b = 0,85 b = 0,48	[r <sup>2</sup> = 0,998]
easy-going / cool	short note values phrase ending high pressed	b = 0,76 b = -0,57 b = -0,58	[r <sup>2</sup> = 0,98]

Some of the relations between affect and vocal expression presented in table 2 are not very surprising, whereas other models contain new insight. For example, the ascription “aggressive” could be traced back to shouting or screaming – or to be more precise: from the data used in the survey one could infer that with a comparable music excerpt the impression of aggressiveness, rated on a five-point scale, could be deduced from the degree of screaming,

rated on a 7-point scale and multiplied with the factor 0,85; one does not need any other vocal features to explain two thirds of the variance (68%) in the ratings concerning the degree of aggressiveness. This relationship between screaming/shouting and the degree of ascribed aggressiveness is not astonishing at all, neither are the relations in the next examples: there's a strong relation between "seductive" and a breathy phonation ( $b = 0,91$ ,  $r^2 = 0,815$ ) as well as between "erotic" and a breathy phonation ( $b = 0,85$ ,  $r^2 = 0,68$ ). By the way, the correlation between the ratings for "seductive" and "erotic" is very high (Pearson's  $r = 0,96$ ) that means that ascriptions of the two adjectives are almost perfectly linked. There are some evident relations between the degree of vocal effort in the music excerpts and ascriptions of certain personality traits, for example "assertive" ( $b = 0,88$ ,  $r^2 = 0,75$ ), "bored" ( $b = -0,91$ ,  $r^2 = 0,80$ ) and "shy" ( $b = -0,89$ ,  $r^2 = 0,76$ ). Note that there's a negative prefix for two of the weightings, which means that the ascription to the singer of being bored or shy is higher if the vocal effort is weaker.

Some other affect ascriptions have to be modelled with a combination of several vocal features. For example, "restrained" could be inferred from low degrees of shouting/screaming ( $b = -0,84$ ) and rubato singing, that is: singing not according to meter but with a lot of micro-rhythmic play ( $b = -0,312$ ). Therefore, a singer seems to be a restrained person, if he does not scream and sings very precisely on the beat, whereas not screaming is much more important than rhythmically exactness of singing. For being aroused there are several determinants: phrase ending on a high note ( $b = 1,31$ ), fast tempo ( $b = 1,27$ ), neutral phonation ( $b = 0,89$ ) and pressed phonation ( $b = 0,484$ ). And a last example: the ascription to a singer to be easy-going or cool could be predicted from the predominance of short note values ( $b = 0,76$ ), a negative weighted factor on "phrase ending on a high note" ( $b = -0,57$ ) that actually means, that phrase ending have to be on a low note, and no pressed articulation ( $b = -0,58$ ). In other words: the impression of easy going and being cool goes along with short notes, a relaxed articulation and low notes at phrase endings.

Of course, the preliminary results of this pilot study have to be broadened and enhanced by taking more music excerpts from a variety of genres and more listeners with different cultural backgrounds. Additionally, the regression models have to be cross-validated by comparing model predictions for new musical excerpts with actual ratings of listeners.

## Conclusions

To explore voice and singing in popular music is a huge challenge for everybody intrigued by the manifold voices in various popular music genres. Several approaches and methods could complement one another and contribute to a profound understanding of vocal expression and its reception in popular music. Of course, the approaches proposed in this paper should be completed by, if possible, interrogations of the singers in order to explore their strategies and techniques of vocal performance. Together with analysis of recordings this

could be the basis for a history of several vocal styles that link vocal expression to certain cultural patterns as well as to the personality of a singer and to his biographical background. In order to escape from a subjective interpretation of the emotional and affective contents of a vocal performance I proposed an empirical approach to explore how listeners ascribe certain affects and personality traits to certain features of vocal expression. Of course, this approach could be enhanced in manifold ways, and there are many further empirical designs that could be employed, too: for example, multi-dimensional scaling of different music excerpts or qualitative enquiries of listeners experience. Moreover, since many instruments imitate or copy some features of vocal expression, probably research on popular voices could be transferred to expression by music instruments, e.g. electric guitars or horns (Pfeleiderer, 2010), too.

## Notes

1. See Frith (1996, pp. 183-202), Middleton (2000), Moore (2001, pp. 44ff.) Potter (1998, pp. 87-112 and pp. 133-189), Banfield (2000) and Bielefeldt (2006, 2008). David Brackett (1995) explores several aspects of vocal performances in his analysis of singers Bing Crosby, Billie Holiday, Hank Williams, James Brown and Elvis Costello.
2. Peter Shapiro (2000, p. 24) characterizes the recording as “one of the two or three most significant singles ever released”, since the single “marked the true beginnings of soul music”. Beside singer and piano player Ray Charles the band consists of Wesley Jackson, guitar, Jimmy Bell, bass, Glenn Brooks, drums, Joe Bridgewater and Charles Whitley, trumpets, and Don Wilkerson and David Newman, saxophones.
3. The spectrograms are computed with Praat 4.6.36, an audio freeware that has been developed by Paul Boersma and David Weenink at the Institute of Phonetic Sciences, University of Amsterdam (cf. www.praat.org). I used a FFT (Fast Fourier Transformation) with a Gaussian analysis window (0,03 s window length), a dynamic range of 40dB and a frequency representation from 0 to 8 kHz (fig. 1-3) and 0 to 2 kHz (fig. 5).
4. Excerpts are taken from the following songs: Muse “In your world”, TokTokTok “The Breakdown”, Amanda Woodward “Binaire Et Lisible“, Primus “Seas of Cheese“, Shivaree “Goodnight Moon“, Sizzla “Boom and Go through“, Christina Aguilera feat. Herbie Hancock “A Song for You“, Dead Kennedys “Bleed For Me“, DFA1979 “Sexy Results“, and Cypress Hill “Make a Move“.
5. Multiple linear regression is a standard procedure based on correlation statistics. Unfortunately, I can not delve into statistics here. For details see e.g. Aron/Aron/Coups, 2006, pp. 497-543, or any other statistics guide.

## Acknowledgements

Many thanks to all participants of the surveys, to Martha Uihôa for inviting me to publish this paper in IASPM@journal and to an unknown reviewer for helpful comments and suggestions.

## References

- Aron, A., Aron, E. N. & Corpus, E.J. 2006. *Statistics for Psychology*, Pearson Prentice Hall, Upper Saddle River, NJ. Fourth Edition.
- Banfield, Stephen. 2000. "Stage and Screen Entertainers in the Twentieth Century" in Potter. J. ed. *The Cambridge Companion to Singing*, Cambridge University Press, Cambridge, pp. 63-81.
- Bielefeldt, Christian. 2006. "Hiphop im Candyshop. Überlegungen zur populären Stimme" in Helms, D. & Phleps, T. eds. *Cut and paste. Schnittmuster populärer Musik der Gegenwart*, transcript, Bielefeld, pp. 135-152.
- Bielefeldt, Christian. 2008. "Voices of Prince. Zur Popstimme" in Bielefeldt, C., Dahmen, U. & Grossmann, R. eds. *Popmusicology. Perspektiven der Popmusikwissenschaft*, transcript, Bielefeld, pp. 201-219.
- Brackett, David. 1995. *Interpreting Popular Music*, Cambridge University Press, Cambridge.
- Frith, Simon. 1996. *Performing Rites. Evaluating Popular Music*, Oxford University Press, Oxford.
- Horton, Donald. 1957. "The Dialogue of Courtship in Popular Songs". *American Journal of Sociology*, Vol. 62, pp. 569-578 (reprint in Frith, Simon ed. *Popular Music Vol. 3, Popular Music Analysis*, Routledge, London 2004, pp. 173-185).
- Juslin, P. N. & Scherer, K. R. 2008. "Vocal Expression of Affect" in Harrigan, R. Rosenthal, J. A. & Scherer, K. R. eds. *The New Handbook of Methods in Nonverbal Behavior Research*, Oxford University Press, Oxford, pp. 65-135.
- Kluckhohn, Clyde. 1962. *Culture and Behavior. The Collected Essays of Clyde Kluckhohn*, ed. by Richard Kluckhohn, New York.
- Laver, John. 1980. *The Phonetic Description of Voice Quality*, Cambridge University Press, Cambridge.
- Lomax, Alan. 1962. "Song Structure and Social Structure". *Ethnology*, Vol. 1, pp. 425-251.
- Lomax, Alan. 1968. *Folk Song Style and Culture*, American Association for the Advancement of Science, Washington.
- Middleton, Richard. 2000. "Rock Singing" in J. Potter ed. *The Cambridge Companion to Singing*, Cambridge University Press, Cambridge, pp. 28-41.
- Moore, Allan F. 2001. *Rock. The Primary Text. Developing a Musicology of Rock*, Ashgate, Aldershot. Second edition.
- Pfleiderer, Martin. 2009. "Stimmen in populärer Musik. Vokale Gestaltungsmittel und Aspekte der Rezeption" in R. Bader ed. *Musical Acoustics, Neurocognition, and Psychology of Music / Musikalische Akustik, Neurokognition und*

*Musikpsychologie* (= Hamburger Jahrbuch für Musikwissenschaft, Vol. 25), Peter Lang, Frankfurt, pp. 233-274.

Pfeleiderer, Martin. 2010. "Singin' the Blues. Vokale Expressivität im instrumentalen Jazz" in W. Knauer ed. *Tension. Darmstädter Beiträge zur Jazzforschung*, Vol. 11, Wolke, Hofheim (in print).

Potter, John. 1998. *Vocal Authority. Singing Style and Ideology*, Cambridge University Press, New York.

Shapiro, Peter. 2000. *Soul. 100 Essential CDs. The Rough Guide*, Rough Guides, London.

Sundberg, Johan. 1988. *The Science of the Singing Voice*, Northern Illinois University Press, Dekalb.

Tagg, Philip. 2009. "Music analysis for 'non-musos'. Popular perception as a basis for understanding musical structure and signification", Paper for Conference on Popular Music Analysis, University of Cardiff, 17 November 2001, < <http://www.tagg.org/xpdfs/CardiffLBH2.pdf> > [Accessed 29 April 2010]

Winkler, Peter. 1999. "Writing Ghost Notes. The Poetics and Politics of Transcription" in Schwarz, D. ed. *Keeping Score. Music, Disciplinary, Culture*, University of Virginia Press, Charlottesville, pp. 169–203.

Zimbardo, P. G. & Gerrig, R. J. 2004. *Psychologie*, Pearsons, Munic. Sixteenth edition.

## **Discography**

Ray Charles. 1994. 'Come Back Baby', *The Best of Ray Charles. The Atlantic Years*, Rhino.