

Singing, Speaking, and the Difference

Jeanette Bicknell, Independent Scholar

The differences between singing and speaking seem obvious at first sight. When people are asked to specify those differences, they typically point to two kinds of things. The first are the physiological differences in what a vocalist does when she speaks or sings. The second are differences in the resulting sounds.

I have several aims in this paper. One is to show that the differences between singing and speaking are not, in fact, obvious, and to examine some seemingly ambiguous cases. These *borderline* cases merit philosophical reflection on the categories being used. Another aim is to argue that the distinction between singing and speech is not best made along physiological or auditory lines, or even solely along musical lines. Instead, the distinction between singing and speech is better made on cultural lines and on pragmatic grounds. What a singer or speaker is considered to be doing when he or she communicates, and how that communication is received, will depend heavily on social and cultural factors, and on the shared expectations of singers, speakers, and listeners. Different cultures understand the difference between speech and song in various ways. Finally, I aim to explore the ways in which composers, songwriters, and performers exploit the differences and the boundaries between singing and speaking.

With the distinction between song and speech made along cultural and pragmatic lines, we can better understand some of the borderline cases, both in artistic contexts and in social life more generally. This in turn will allow us to consider some of the wider implications: why does singing continue to have a place in nearly every musical genre, and why sing when speech usually is a more efficient means of communication than song?

Quora (www.quora.com) is an internet forum where anyone can ask and answer questions; the answers are edited, rated, and organized by the site's users. It is a great way to get a variety of answers on topics from current science and politics to popular culture, the arts, and mundane practicalities. Participants range from curious and naïve to people who are experts in their field. When someone asked, "What is the difference between speaking and singing?" the question did not arouse much controversy. The relatively narrow range of the answers was instructive, although not particularly surprising.²⁰ According

²⁰ "What is the difference between speaking and singing?," Quora, accessed on September 15, 2017, <https://www.quora.com/What-is-the-difference-between-speaking-and-singing>.

to Quora's users, the primary differences between singing and speaking are physiological. Both singing and speaking are physical processes in which breath is moved through the larynx and vocal folds, and the resulting sound is shaped into consonants and vowels with the tongue, teeth and lips. Singing demands a larger range of pitch and requires greater effort, together with proper breath support, and greater control of the vocal apparatus. The answers also mention that singing tends to be more "emotional" than speech, and that singers usually sing existing compositions, whereas everyday speech is "composed" moment-to-moment.

It was presumably such commonsense insights that prompted philosopher of art Frances Sparshott, long before Quora, in an article about the difference between singing and speaking, to claim that, "it is usually easy to tell when someone starts singing."²¹ There are several problems with this seemingly unproblematic claim. First, not all cultures make a distinction between speech or talking on the one hand, and song or singing on the other.²² So while it may seem straightforward enough for members of our own Western culture to make the distinction, it does not follow that it is easy for everyone.

Second, even if it is indeed "easy" to tell when someone starts singing (which I am not quite prepared to allow), it is so because we typically already have a good understanding of the context in which an individual vocalizes, and we judge on that basis. If we were to attend purely to the auditory features of an utterance, purely to the quality of sound that he or she makes, it would not always be easy. There are two kinds of complications. The first are cases where the vocal production falls between singing and speaking. The second are cases in which we need to understand the social and cultural context in order to determine what the individual in question understands herself to be doing, and how she expects that utterance to be received.

Let's consider the borderline cases first: many vocal productions fall between singing and speech, yet arguably belong clearly to neither. In a 1963 article, George List proposed a classification of vocal production based on modifications of speech intonation. (He limited his analysis to non-tonal languages, since including tonal languages such as Mandarin Chinese in his classification would present "innumerable" difficulties.) "Intonation" here refers to the melodic pattern of an utterance, which can be exaggerated or minimized by speakers. Speech intonation may level out and approach a monotone or be heightened and exaggerated. List defined song as a communication form exhibiting relatively stable pitches, possessing at least a minimally elaborate scalar (melodic) structure, and showing little influence of speech intonation.²³

²¹ Francis Sparshott, "Singing and Speaking," *British Journal of Aesthetics* 37:3 (July 1997): 199.

²² George List, "The Boundaries of Speech and Song," *Ethnomusicology* 7 (1963): 3.

²³ List, "The Boundaries of Speech and Song," 3.

Some of List's search far afield to find and speech. Non-articulate street sellers' calls and calls to prayer. In the "heightened" "talking" blues.

The second kind of qualities. I mentioned recognize what we mean function of these forms cultural activity, and so tations and takes place its role in society. What doing when he or she ve of everyone involved.

Consider the role of In cultures where music reluctant to describe their of a larger spiritual practice considered by adherents of passages from the Torah cantor or chazzan in Judaism.

Sprechstimme (from tion halfway between speech terms of two musical genres contrast these with spoken and artistic tradition, is not

²⁴ List, "The Boundaries of Speech and Song," 3.

Some of List's examples of borderline cases drew on ethnographic research, but we do not need to search far afield to find vocal practices, both in art and in everyday life, that fall sonically between singing and speech. Non-artistic examples include children's skipping and clapping rhymes, auctioneers' chants, street sellers' calls and cries, field and street hollers, the chants used in meditation and religious practices, and calls to prayer. In the arts, we find vocal practices that fall somewhere between speech and singing in the "heightened" speech of musical theatre, spoken word poetry, recitative, *Sprechstimme*, rap, and "talking" blues.

The second kind of complication revolves around cultural and social practices, rather than auditory qualities. I mentioned earlier that not all cultures distinguish between speech and song. Other cultures recognize what we may consider intermediate forms between the two (such as chanting), and the social function of these forms may vary widely.²⁴ While singing is a physical activity, it is also a social and cultural activity, and some singing is artistic. Singers' activity is subject to a variety of collective expectations and takes place within specific shared contexts and shared cultural understandings of "music" and its role in society. When we look at musical practice globally, what an individual is considered to be doing when he or she vocalizes will depend heavily on social and cultural factors, and on the expectations of everyone involved.

Consider the role that vocalizations of different kinds play in religious practices around the world. In cultures where music is understood as a secular pursuit or has little prestige, participants may be reluctant to describe their vocal activity as singing or as musical. Instead, they may see it as just one part of a larger spiritual practice. Islamic calls to prayer may sometimes sound like songs but are not usually considered by adherents to be singing performances. Similarly, in Jewish religious practice the chanting of passages from the Torah is referred to as "reading" or "recitation," rather than singing. The role of the cantor or chazzan in Judaism is seen to be primarily moral or spiritual rather than musical.

Sprechstimme (from the German verb *sprechen*, to speak) is the musical name for the vocal production halfway between speech and song. I'll first say something general about it, and then discuss it in terms of two musical genres (hip-hop and "talking" blues). I see all of these as forms of song. Then I'll contrast these with spoken word poetry, which I see as a related artistic practice that, for reasons of culture and artistic tradition, is not properly a musical genre and not a form of singing, but of speaking.

²⁴ List, "The Boundaries of Speech and Song," 3.

Peter Kivy described, in a paper delivered to the American Philosophical Association in 2013, how early opera composers were faced with the problem of how to set dramatic dialogue to music, given that the pace of music is slower than that of conversational speech. One response to this problem was the development of the genre melodrama, in which words are spoken, accompanied by music. While the genre did not survive much past the 18th century, the technique of having a character speak accompanied by music, instead of sing, is still used occasionally in opera. A well-known example is found in Beethoven's *Fidelio* (Leonore in Act II). Arnold Schönberg most fully developed *Sprechstimme* in his opera *Moses and Aron* and in his song cycle *Pierrot Lunaire*.

Kivy argues that the vocal practice in melodrama and *Sprechstimme* are best understood as a type of song or "singing in speech." He bases his argument on the sound quality of the vocal production: "We are drawn to the *sounds* of the words, *as if* they were a kind of musical sound, as we are not to the merely spoken dialogue in the opera..."²⁵ I agree with him that *Sprechstimme* is a type of singing, although I would emphasize historical and cultural reasons in support of this claim. *Sprechstimme* developed in a musical context (as opposed to a strictly theatrical context), was developed mainly by composers (not playwrights or theatre directors), and is a technique expected in the repertoire of singers (rather than actors or poets). To put it a little crudely, *Sprechstimme* is a form of singing because we treat it like one.

Rap can also be seen as a form of *Sprechstimme*. Like *Sprechstimme* in art music, it strikes the ear as something between singing and speaking and is usually accompanied by music. Interestingly (and despite the obvious differences between them), rap also arose as a response to challenges of artistic media. The origin of hip-hop dates to the mid-1970s with huge block parties in the Bronx and other areas of New York City. DJs broadcast popular dance music, manipulating turntables to isolate percussive passages and keep the crowds dancing. Another person, the MC, rapped over the music to keep the party going and the crowd entertained. MCs soon sought to out-do one another, and their rapping reflected their ambitions. While initially the focus in hip-hop was on dancers, not vocalists, rap is now a medium for personal expression and political comment.²⁶

Like *Sprechstimme* in art music, the vocal delivery of rap is stylized and does not sound either like regular speech or standard singing. Also, like *Sprechstimme*, rap can be appreciated for its auditory and musical qualities—its "flow" (rhythm and rhyme), and for the performer's skill, as well as for its lyrical content. Listeners are drawn both to the meaning of the words and to the musical qualities of their arrangement and delivery. And like *Sprechstimme*, there are historical and cultural reasons why rap is best understood as a form of singing, rather than speaking. For one, it is now firmly part of the music industry.

²⁵ Peter Kivy, "Speaking in Song and Singing in Speech: Another Paradox for the Paradoxical Art." Unpublished manuscript.

²⁶ I have oversimplified the musical and cultural development of rap. See James McBride, "Hip Hop Planet," *National Geographic*, April 2007.

Rap artists are consi
formed a freestyle (i
tainment Television
style." They didn't c
without instrumental

In the performing
of expectations, carr
different expectation
the other. Both comb
to or emphasize spee
or both. Yet how they
how they sound. The
milieu in which they

Talking blues arc
adhere to some form
instrumentalists or si
poetry recitation. Wo
formers are usually po
are low, including pop
cultural associations a

Compare two ex:
country music star Jo
describes the trip he a
plea for peace and fo
pressing political pos
suited to convey a lot
they can be readily ur
At a time when view
cated sentiments of st
of-fact storyteller, rec

²⁷ Jon Blistein, "W.
<http://www.rollingstone.com/music/album-reviews/jon-blistein-w508187>.

²⁸ PeterRabbit59. "2
Posted [May 2010]. <https://www.rollingstone.com/music/album-reviews/peter-rabbit-59>.

Rap artists are considered to be musicians rather than poets. For example, in 2017 when Eminem performed a freestyle (improvised) rap highly critical of President Donald Trump during the Black Entertainment Television awards, *Rolling Stone* magazine called his performance, “a ferocious a cappella freestyle.” They didn’t call it a poem or a speech, and “a cappella” is, of course, a term that refers to song without instrumental accompaniment.²⁷

In the performing arts, vocal practices that sound superficially similar may respond to different kinds of expectations, carry different associations, and arouse different meanings for listeners. Consider the different expectations and meanings surrounding talking blues on one hand and spoken word poetry on the other. Both combine rhythmic speech and instrumental music. Both practices have names that refer to or emphasize speech. Both can be used didactically, to state a political position, or for entertainment, or both. Yet how they are classified and how listeners respond to them involves many more factors than how they sound. These include most importantly the traditions from which they arose and the cultural milieu in which they are situated.

Talking blues arose from the blues musical tradition. Structurally, works tend to be formulaic and to adhere to some form of song structure, yet with a much simpler chordal structure. Performers are usually instrumentalists or singers. Spoken word poetry comes from a very different tradition, that of modern poetry recitation. Works are freer, more individualistic, and tend not to adhere to established forms. Performers are usually poets rather than instrumentalists or singers. The cultural associations of talking blues are low, including popular art, folk, and country music. Spoken word poetry, on the other hand, has high cultural associations and is often seen as a more elevated, esoteric pursuit.

Compare two examples, both with political themes, that date from around the same time. In 1971 country music star Johnny Cash had a hit with “Singin’ in Vietnam Talkin’ Blues.”²⁸ In the song, Cash describes the trip he and his wife June Cash took to Vietnam to sing for American troops. It ends with a plea for peace and for the war to end. While Cash wrote and sang other, more conventional songs expressing political positions, presumably he chose the talking blues form in this case because it is well-suited to convey a lot of words (compared with a conventional singing style), and to do so in a way that they can be readily understood by listeners (again, when compared to a more conventional singing style). At a time when views about the Vietnam war were highly polarized, Cash wanted to convey the complicated sentiments of supporting American troops while being anti-war. His vocal style is that of a matter-of-fact storyteller, recounting events as they happened. There are few vocal or musical flourishes drawing

²⁷ Jon Blistein, “Watch Eminem Demolish Donald Trump in BET Awards Freestyle,” accessed October 11, 2017, <http://www.rollingstone.com/music/news/watch-eminem-demolish-donald-trump-in-bet-awards-freestyle-w508187>.

²⁸ PeterRabbit59. “Johnny Cash–Singing in Vietnam Talking Blues.” Filmed [March 1971]. YouTube video. 03:12. Posted [May 2010]. <https://www.youtube.com/watch?v=EQAXTvcMoMA>.

attention away from the narrative. Again, all of this draws listeners' attention to the words and the message they contain.

My second example, from 1970, is Gil Scott-Heron's "The Revolution will not be Televised."²⁹ It first appeared on the album "Small Talk at 125th and Lennox," the cover of which announces Scott-Heron as "A New Black Poet." Like Cash's "Singin' in Vietnam Talkin' Blues," it is an overtly political work. Scott-Heron's vocal style is different from Cash's. He does not adopt a storyteller style but is more declamatory. The musical background of the two works is also different; while Cash accompanies himself on a guitar, Scott-Heron performs with flute, congas and bongo drums. I assume that Scott-Heron chose to perform his work and release it on an LP (rather than simply publish a written text) partly in an attempt to reach a wider audience.

While talking blues and spoken word poetry share many similarities, there is one fundamental difference between them. Talking blues is best understood and usually treated by listeners as a form of song. Spoken word poetry is a form of speech. This difference stems not only from vocal style or anything that could be measured with a spectrograph, but from audience expectations having to do with their respective traditions and cultural milieu. We distinguish one from another by these means, rather than on the basis of auditory cues.

Although I've argued that talking blues is a form of singing, there are other types of "speaking in song" that are unambiguously speech, and these also deserve our attention. I have in mind popular songs with the addition of ad lib (or "ad lib") speech (accompanied by music), and popular songs with spoken lyrics.

The most obvious examples of ad lib speech within a song are the times when a singer breaks off in the middle of the song to speak directly to the audience. Bruce Springsteen is well-known for this practice. He breaks off singing and speaks directly to the audience *in propria persona*, sometimes telling a story about his past that inspired the song he is singing. Usually the band plays on, accompanying him, and he resumes singing when he reaches the end of the story.

There are many reasons why singers might do this. First, they may be responding to practical considerations, especially in live performance. Springsteen is known for marathon performances, so the speaking interludes may give his voice a rest (singing being more physically demanding than talking).

²⁹ Acc Records Ltd. "Gil Scott-Heron—Revolution Will Not Be Televised (Official Version)." YouTube video, 03:06. Posted [October 2013]. <https://www.youtube.com/watch?v=vwSRqaZGsPw>.

As I have already n
may speak rather th
While the phen
a song seems to be f
great example can b
comes near the end

Well I guess by
You know I...
playing. Do you
is Louis's recor
four more bars a
else, I don't kno
ing this song ...
trying to wait for
"Tired."]

In this example,
her to comment on th
tired!"). It also conve
more bars and "sneak
interest, musically or
feather duster in the f
way of gently sending
in on the joke.

In these examples
argue that we understa
by music, the words ar
speech within the song

My last category c
the song, with or with
mind, the words are sp
Leiber and Stoller hit,
Peggy Lee in 1969. Sh

The verses describ
love for the first time.

³⁰ peggy4AL. "Pegg
2009]. <http://www.youtube>

As I have already mentioned, song is not a particularly efficient means of communicating words. Singers may speak rather than sing when there is something they want to be sure that the audience understands.

While the phenomenon of singers speaking directly to the audience within the context of performing a song seems to be found most often in live performance, it has been captured on recordings as well. One great example can be found on Pearl Bailey's 1956 recording of "Tired" with Louie Bellson's band. It comes near the end of recording, before the final chorus:

Well I guess by now y'all have the general idea: I am tired! I... I am in the wrong place, too. You know I... what I am I doing in this studio? Well I just passed by and heard "Tired" playing. Do you know I'm in the wrong place? I thought this was *my* recording session. This is Louis's recording session. Sounds so good, though, I think I'll stay here and at least sing four more bars and enjoy myself while I'm here. Come to think of it, you know something else, I don't know nobody here. They certainly look like nice fellas, though. They keep playing this song ... maybe they expecting me to come in later? Yeah, it's a long wait but I'm trying to wait for a few more bars and then I'll sneak in on 'em [she suddenly resumes singing "Tired."]

In this example, speaking allows Bailey briefly to step outside of the song's persona. This enables her to comment on the performance in progress ("I guess by now y'all have the general idea that I am tired!"). It also conveys something of her personality (as when she tells us that she's going to wait four more bars and "sneak in" on the musicians. It is also worth noting that "Tired" likely offered little of interest, musically or lyrically, to Bailey. She had already sung it wearing a frilly apron and wielding a feather duster in the film *Variety Girl* in 1947. Her spoken comments in the 1956 recording may be a way of gently sending up the song and amusing herself and her musicians, while also including listeners in on the joke.

In these examples (the Springsteen and the Bailey), and unlike in the talking blues example, I would argue that we understand the speech as speech, rather than as song. Although the speech is accompanied by music, the words are not generally pronounced in a stylized manner. Rather, the aesthetic effect of the speech within the song performance depends on listeners understanding what is being said.

My last category of speech in song occurs when singers do not ad lib but, instead, speak the lyrics of the song, with or without musical accompaniment. Unlike rap or melodrama, in the examples I have in mind, the words are spoken more or less "straight." My final example of speaking within a song is the Leiber and Stoller hit, "Is that All There Is?" It has been recorded by many performers, most notably by Peggy Lee in 1969. She recites the verses and sings the chorus.³⁰

The verses describe the narrator's experience of a house fire, a circus performance, and falling in love for the first time. Each of these leave her underwhelmed. In the chorus she asks, "Is that all there

³⁰ peggy4AL. "Peggy Lee—Is That All There Is? 1969." Filmed [1969]. YouTube video, 04:37. Published [Oct 2009]. <http://www.youtube.com/watch?v=LCRZZC-DH7M>

is?” and counsels us to “break out the booze and have a ball” if life really is as disappointing as it seems. It would be tempting, and yet unfortunate, to sneer at the song’s “dime-store existentialism.” Let me share some of its origin. The lyrics were inspired by Thomas Mann’s short story “Disillusionment” written in 1896. This places it in a similar cultural milieu to Albert Giraud’s cycle of poems, written in 1884, which inspired Schönberg’s *Pierrot Lunaire*. I would argue that in both works—the popular song and the atonal melodrama—using speech or *Sprechstimme* rather than straight singing—serves a similar purpose. These vocal performances are contrary to listeners’ expectations. Listeners expect singing, not talking, and not something in between singing and talking. The effect is uncanny, and I would argue, extremely apt to the lyrical content.

I’ve said something about the differences between singing and speech and how these differences are not necessarily obvious. I’ve argued that we do not (and should not) draw the line between singing and speech primarily on the basis of auditory cues. I’ve also said something about the artistic difference it makes when a performer speaks or uses *Sprechstimme* when the audience expects unambiguous singing. What of the larger differences between singing and speaking? Why does the practice of singing persist, and indeed thrive, in every musical genre in which it has a place?

Let me return to the article by Sparshott that I referenced earlier. He frames the problem in this way:

Singing is a fundamental use of a part of our built-in psychophysical apparatus, the voice mechanism. If the basic use of that mechanism is speech, why should it be susceptible of modification in a different way, proving to have possibilities that speech does not exploit?³¹

What I find interesting here is the assumption that speaking is primary (the “basic use” of the built-in mechanism) and singing is derivative (a “modification”). The fact is, we do not know what form the first human vocal communications took. Were they threats howled in imitation of an animal’s cadence, or perhaps in a mother’s reassuring syllables, crooned and repeated until a lilting melody took shape? No matter what the origins of human vocal communication, it is not outrageous to conjecture that singing and speaking evolved together.

Yet at the same time, something about Sparshott’s assumption seems fitting. Whether singing or speaking represents an earlier stage of human evolution, it is clear that today speech is the default mode of vocal communication. Singing, regardless of the social context in which we find it (and along with *Sprechstimme*, chant, and all of the in-between forms), is socially demarcated as special. To sing when speech is expected is to draw attention to oneself. Singing is a form of heightened expression, with words

³¹ Sparshott, “Singing and Speaking.” 201.

enhanced by affect,
pression, singing in

In 2015, Barack
was called upon to
Pinckney was murder
ton, South Carolina.
That silent pause was
of communication to
ever so gradually. Th
they recover, stand u

Obama is well-k
were no match for th
could not have done.

³² C-SPAN. “President
lished [Jun 2015]. <https://yo>

³³ For comments and di

enhanced by affect, emphasis, repetition, rhythm, and the power of music itself. And as heightened expression, singing invites intensified or special kinds of attention.

In 2015, Barack Obama faced surely one of the most difficult moments of his presidency when he was called upon to deliver the funeral address of South Carolina State Senator Clementa Pinckney.³² Pinckney was murdered, along with eight others, in a racially-motivated shooting at a church in Charleston, South Carolina. Near the end of his address, Obama paused, then started to sing "Amazing Grace." That silent pause was nearly thirteen seconds long. It was a silence in which the transition from one mode of communication to another, from the declaration of speech to the emotive language of song, was made ever so gradually. Those standing with Obama on the dais laugh at first, clearly not expecting this, yet they recover, stand up, and start to sing along. Soon everyone in the church is singing.

Obama is well-known and highly praised as an orator. Yet the power of his words, in this instance, were no match for the power of song. His gesture brought everyone together in a way that words alone could not have done. The best speakers understand when words alone are not enough.³³

³² C-SPAN. "President Obama sings Amazing Grace (C-SPAN)." Filmed [Jun 2015]. YouTube video, 02:30. Published [Jun 2015]. <https://youtu.be/IN05jVNBs64>.

³³ For comments and discussion, I am grateful to Joseph Agassi, Ian Jarvie, and especially to Jennifer Judkins.