Lipsmacks

This studio study continues the examination of how the presentation of popular music recordings is affected by the repurposing of audio after recording. Generally the practices described here represent a kind of “cleansing” of musical performances, as opposed to the “fixing” described in the previous application study. This study focuses on the nonverbal sounds that singers sometimes make as a part of the vocal production of singing. I have divided these vocal artifacts into three general categories: lipsmacks, mouth noises, and heavy breathing, though these are in no way intended as exhaustive.

I use the term lipsmacks to refer to a kind of nonverbal sound that sometimes occur between singer’s words. A lipsmack is most commonly caused when a singer opens her mouth in preparation to sing. The separation of the two lips can cause an audible “smacking” or clicking type sound. This may be an occasional occurrence that passes relatively unnoticed, or it may be a common occurrence that can become a major distraction in a vocal performance. One artist that I have worked with, the blues singer Freddie Hughes, is unable to stop himself from making frequent and loud lipsmacks between vocal lines. These sounds—distracting clicks that are not even identifiable as part of the vocal performance—have plagued his recordings. While they also occurred in live performance, extraneous sounds are much easier to ignore in a live setting where there are plenty of visual distractions. Prior to the advent of computer-based editing techniques, there were simply too many of these sounds, too closely placed to
Freddie’s sung vocal, to remove them from the final recording. In the computer it’s a relatively simple matter to remove the clicks and to create inaudible transitions around the places where the sound has been removed.

This simple action extends the sanction of the recording engineer beyond the kind of fixing discussed in chapter 1 and broadens my consideration of DAW-based alterations of musical performances. Given that nothing in a recording can be truly “original”—that is, identical to the source—what is essential to a singer’s performance? Are all alterations beyond the most transparent translation of source to recording to be considered anathema? Here at the beginning I have provided one example of performance alteration that might be seen as generally positive and benign in terms of its effect on the substance of the performance. The ability to remove lipsmacks has certainly been a great relief to Freddie Hughes, as it has allowed him to make a record that didn’t have a lot of distracting clicks and pops on it. The audio clips provide examples of Freddie’s vocal before and after the clicks have been removed (audio clips 3, 4, 5).

**Mouth Noises**

By mouth noises I am referring to sounds that the mouth makes while singing certain words—artifacts of word production that are not a normal or necessary part of the word itself. Two common such artifacts are popped “p’s” and excessive sibilance. Popped “p’s” are created when the explosive kind of exhalation that might accompany a particularly expressive hard consonant is sung directly into a microphone (“p’s” are the most common offenders, but any hard consonant might produce this effect). The effect of this exhalation is to vibrate the diaphragm of the microphone in a way that produces a relatively loud after-effect—a popped “p.” As with a lipsmack, this may be perceived as a distracting, nonverbal artifact of the recording process. A popped “p” is in some sense “natural”—a result of a natural vocal occurrence—but it is so amplified and altered by its effect on the microphone as to make it “unnatural” to the ear and unique to recordings.

Thanks to the digital processing power of DAWs it is usually possible to eliminate the effect of the popped “p.” This is done by isolating the popped “p” and filtering out most of the low frequencies. This generally removes the “pop”—the explosive sound that follows the initial articulation of the “p” sound—while retaining the high-frequency transients that make up the majority of the actual “p” sound. Bob Dylan’s original recorded performance of “Sad-Eyed Lady of the Lowlands” contains many popped “p’s”
on lines such as “Where the sad-eyed prophets say that no man comes.” The audio clips allow the listener to hear the original popped “p” and then the “cleaned-up” version (audio clips 6 and 7).

Is the application of this kind of editing clearly to be preferred? Is Dylan’s performance more or less appealing with the popped “p’s”? Does the vocal performance have more or less impact with the popped “p’s”? If the popped “p’s” had been removed before the record was released, would they have been missed by the listener? I will return to these questions, but first I will consider another of these types of recording artifacts.

Excessive sibilance is a part of vocal performance that may also be an artifact of recording. The enormous amount of high-frequency information in the sound of sung “esses” might be amplified by a variety of recording techniques. Compression, short delay, and high-frequency EQ may all generate an unnatural level of the “s” sound in the recording of a vocal performance and may further slur and distort that “s” sound in the process. As with popped “p’s” there are techniques for taming this excessive sibilance—primarily with a processing device called, appropriately enough, a “deesser.” De-essers automate a volume reduction in the sibilance, making the “esses” quieter and thus less intrusive and more in line with what their level might be in a live, acoustic performance. However, when the sibilance has undergone so much processing as to become smeared and distorted, it is not possible to “clean” it with a de-esser. The volume of the “s” may be reduced, but in these cases the recording process has changed the nature of the sound in a way that can’t be undone without undoing the process, which may mean rerecording the performance. Only if the heightened “s” sound is created by processing done after the recording has been made, such as during the final mixing, would it be possible to undo this effect by a remix of the same performance in which the processing responsible for the excessive sibilance was removed.

Many modern recordings have instances of very exaggerated sibilance that could be avoided in this later mixing process. Why isn’t it? Clearly some producers and artists enjoy the rather artificial effect of very pronounced sibilance. A good example is on the line “sound of hysteria” from a song by the band Green Day called “American Idiot.” The strong sibilance on this line can be heard on the audio clip (audio clip 8). Perhaps the added and unnatural sibilance serves to underscore singer Billie Joe Armstrong’s vitriolic sentiment and delivery. Incidentally, this degree of exaggerated sibilance would not have been possible in the age of the vinyl disc; high-frequency transients require very jagged grooves in vinyl in order to be repro-
duced, and if they are too prominent the needle is unable to track the grooves and the record will skip. In the age of vinyl recordings engineers had to monitor and reduce transient levels, especially sibilance, to avoid making LPs that skipped. The use of heightened sibilance as an effect is another example of how DAW production allows new forms of presentation.

In many instances the performer’s relationship to the microphone—to the technology—has supplanted the primacy of his relationship to the audience, so that “while pop performers sing to an audience, real or imagined, they always sing first and foremost to the microphone. In return, the microphone reveals, in intimate detail, every nuance of the performer’s vocal style.” As noted here, the microphone’s interaction with the voice and the recording process may cause “unnatural” elements such as popped “p’s” and excessively sibilant “esses.” These sounds extend beyond the performer’s vocal style in the traditional sense and may even be created without the performer being aware of them.

In regards to these kind of unnatural vocal sounds, how many people actually notice anything odd or distracting in Dylan’s “p’s” or Billie Joe’s “esses?” Do Billie Joe’s exaggerated “esses” and Dylan’s explosive “p’s” produce a positive or negative experience for the listener? Unlike with Freddie Hughes, where I think it’s clearly an advantage to have rid ourselves of those distracting sounds, perhaps the sense of passion created by unnatural “p’s” and “esses” is preferred by most listeners. There might be a heightened experience from the feeling that the listener is almost inside Dylan’s mouth—and indeed it is the microphone being exceedingly close to the mouth when the recording is made that causes this effect. Are these unnaturally heightened vocal artifacts a kind of manufactured passion? Perhaps, but in all likelihood the listener simply accepts the sound of the recording without questioning whether or not it is “natural” or preferable. Whether to “clean up” the recording of these sounds or to leave them for their effect comes down to an aesthetic decision. For the practitioner it is an advantage to now have the tools available that allow this choice to be made based on aesthetics, rather than having been saddled with these artifacts of the recording process whether we like them or not.

**Heavy Breathing**

The most obvious and prevalent nonverbal sound produced by singers is the breaths between phrases. It is also here that recordists implement the most obvious and prevalent kind of manipulations of these sounds. The
sense of exaggerated breathing in vocal performance may simply be part of an exaggerated performance (perhaps even a relative of the kind of heavy breathing that the phrase might first bring to mind). Vocalists may wish to emphasize the physical production of singing by dramatic emphasis of the breaths between phrases. However, there are technical aspects of recording that may also significantly affect the volume level of the singer’s breaths. The most common of these is compression.

Compression reduces the overall dynamic range of a vocal performance and thereby increases the volume of quiet sounds relative to loud sounds. This means that the singer’s breaths, normally quieter in volume than the actual singing, will be raised in volume and sound relatively louder. Compression is used in part to even out a vocalist’s performance, to make it more consistent in level and thereby easier to follow both lyrics and vocal nuance. It is also used to add dramatic presence, to put the vocal more “in your face”—that is, by reducing the dynamic range the vocal presence is more constant and therefore feels closer. This is partially a product of close miking techniques, but it is magnified by the use of compression. The increase in volume of the breaths is part of the added presence and perhaps of the sense of drama. Is this affectation? The relative volume of the breaths is created artificially but the breaths themselves are critically real.

This phenomenon is not new to the digital age of recording and reproduction. The spectacular vocal performance on the 1970 recording of Stevie Wonder singing “Signed, Sealed, Delivered, I’m Yours” features very prominent between-line breathing that is clearly pumped up through compression (audio clip 9). However, contemporary recordings digitally created and reproduced may show even more obvious effects of compression. This is partially due to new tools of compression in the digital domain (especially a very powerful kind of compression called “brick-wall limiting”) and aided by the ability of CDs to reproduce denser audio than was possible on vinyl records. There are many examples of this kind of deep compression in contemporary popular music. While the effect may be heard on the entire recording it is most noticeable on the vocal—especially in the volume of the breaths. A clear example of this can be heard on the audio clip from the track “Cruel” by Tori Amos (audio clip 10).

The superior editing capabilities of computer-based audio recording have affected the status of singers’ breaths. The removal or replacement of singers’ breaths has become commonplace. This is a result of the kind of close scrutiny done to vocal performance—listened to and analyzed in isolation (a cappella)—combined with the simplicity of excising any element
cleanly and completely. Under these circumstances I often get singers asking me to remove certain breaths, or sometimes almost all of the breaths, because they become self-conscious about how the breaths sound. If a particular passage sounds too empty without the offending breath I might grab a shorter or quieter breath from a different part of the performance and insert it. Frequently choices between breaths must be made as one edits together different takes of a vocal performance. In editing together two lines that were not actually sung in sequence one has the choice of the breath at the end of the first line or the breath at the beginning of the second. Manipulation or elimination of breaths has become very common in contemporary production. A new piece of software called DeBreath has been released that is directed solely at this process.\(^5\) One might consider some or all of this activity to be harmless, perhaps meaningless, or one might consider the widespread removal of breaths to be a kind of sterilization of vocal performance. As such this activity would be in line with the contemporary obsession with depilation, deodorants, and the like—even something approaching a realization of the nineteenth-century fantasies of autonomous, disembodied music. In any event, this is another example of the more dramatic kinds of manipulation of performance made possible by computer-based recording platforms.

What would seem clearly an advantage in removing distracting sounds on the Freddie Hughes vocal, and a judgment call on Dylan and Green Day, where artifacts of the recording process may or may not be perceived as desirable, runs the gamut when it comes to breaths. From the removal of breaths, to the replacement of an awkwardly sounding breath, to the alteration of the sound or level of breaths, one’s judgment on the appropriateness of such behavior may range from emphatically positive to extremely negative. Contemporary production opens many such complex questions in regards to the relationship of original performance to final recorded presentation. Some of the production techniques discussed here may cause us to reevaluate these relationships, or at least to confront them under new circumstances.

A Theoretical Basis

Roland Barthes is a valuable point of reference in attempting to draw meaning from the manipulations of recorded audio that I’ve just described. In his seminal essay “The Grain of the Voice” (1977) Barthes attempts to describe what he considers the most important qualities of the
“sound” of the voice. He chooses to call these qualities “grain.” Although Barthes speculates about whether he is the only one hearing this grain—he wonders whether he is hallucinating this quality that he is attempting to describe—what he does know is that if it exists, it is at the margin of our ability to describe it, and as such it is “able to bear traces of significance, to escape the tyranny of meaning.”

This phrase “tyranny of meaning” evokes the ineffable element of the music experience. Barthes is acknowledging that while he may be attempting to describe the indescribable, he remains committed to the venture.

Barthes appeals to concerns beyond the phenomenon of vocal production (what he calls the “pheno-voice”) to that place where the sound of the voice encounters language (“geno-voice”). If genotype is the genetic makeup of an organism, as opposed to its physical characteristics, then geno-voice is the underlying coding or “DNA” of vocal production. It is the essence of the voice that Barthes seeks to describe and comment on. Barthes identifies this critical subtext as residing in the kind of nonverbal aspects of vocal production that we’ve just been describing. What Barthes wants to hear from his singers is “the tongue, the glottis, the teeth, the mucous membranes, the nose.” For Barthes these are the elemental qualities of vocal production. Certainly the physicality of Dylan’s “p’s” and Billie Jo’s “esses” participates in this subtext of vocal production. The fact that the recording process substantially alters these sounds (consciously or not) opens questions of intentionality but doesn’t alter their participation in Barthes’s grain. And the issues that might surround their intentionality are as present for the choices made by the live performance in the studio as they are for the manipulations of the recording engineer either during or after the recorded performance. Popped “p’s” and sibilant “esses” are not just artifacts of recording; they would not be created by the process if there weren’t a certain kind of emphasis (or intentionality) on the part of the singer.

When it comes to breaths, however, Barthes isolates them as separate from the genotypical sounds made in the throat, the mouth, and the nose. He identifies breaths as part of the pheno-voice, part of vocal production, and makes it clear that the grain that he craves from singers is not just a bodily function but also physicality and sexuality. This distinction comes in the context of Barthes’s discussion of two well-known concert music singers, Fischer-Dieskau and Panzera. In dismissing Fischer-Dieskau as without grain he notes that his singing is “beyond reproach . . . yet nothing seduces . . . (the diction is dramatic, the pauses, the checkings and re-
leasings of breath, occur like shudders of passion) and hence never exceeds culture: here it is the soul which accompanies the song, not the body.”

Whereas with Panzera, with whom Barthes perceives this elusive grain, “you never heard him breathe but only divide up the phrase.” Clearly it is not transcendence that Barthes seeks from his singers, but what is it about Panzera’s unheard breath that yields this grain? I don’t believe it is the actual breath that he’s referring to but rather the reliance on breath for expression. He refers to the pedagogy that elevates singer’s breaths to a “myth of respiration,” and it is this interpretation of breaths that he wishes to deflate. He is reclaiming (at least) his own pleasure in listening to vocalists by privileging the more mundane realm of the physical and the visceral over the vaunted classical “breath.”

It is difficult for me to relegate Stevie Wonder’s breathing in “Signed, Sealed, Delivered, I’m Yours” to some glorified realm of spirituality or transcendence, despite Barthes’s desire to dismiss singers’ breaths as pheno-voice. If any nonverbal sounds call up Barthes’s geno-voice, this part of Wonder’s performance would certainly qualify for me. Perhaps this difference from Barthes can be understood by accepting pop vocal performance as essentially different from classical. The distinction Barthes is making is in the perception of the listener, and breathing has not achieved this mythic status of supposed transcendence in pop vocal pedagogy (to the extent that such a thing exists). This might leave Wonder’s breaths more free to participate in Barthes’s grain, as they certainly do for me. A similar divergence from classical vocal performance is heard in the propensity toward vibratoless singing in pop music. This not only dramatically differentiates it from classical vocal style but also speaks to an aversion to the kind of pretension of interpretation that I think Barthes is hearing in classical vocalists’ breathing.

Unfortunately, some of the computer-based work on nonverbal sounds tends to eliminate anything below the level of surface meaning—bits of both Barthes’s pheno-voice and geno-voice are excised. The ease of computer-based editing, combined with vocalist’s tendency toward self-consciousness and the recording culture’s sometime obsession with a kind of perfection that promotes sterility, means that a significant number of popular recordings are “cleansed” of breaths and other artifacts either partially or completely. Yet there are also recording techniques that heighten many nonverbal sounds—the effects of compression, equalization, delay, and the like produce some of the artifacts such as the “p’s” and “esses” described
above. What is never really possible is a simple reproduction of what actually happened inside the singer’s mouth. However, this conscious manipulation of vocal sounds that are produced at the visceral level takes us well beyond core issues of original versus copy or basic questions regarding the status of reproduction. The results are specific and unique to vocal recording and are also part of the evolution of the culture’s acceptance of sounds that didn’t previously exist. That is to say, the culture has come to accept the effects of heavy compression or equalization as “normal” within the context of recorded music. Our culture has absorbed these anomalies into its aural vocabulary.

Paradigms Unraveled

Barthes wants to “disengage this ‘grain’ from the acknowledged values of vocal music” and uses a “twofold opposition . . . theoretical, between the pheno-text and the geno-text, [and] paradigmatic, between two singers.” Further analysis may be balanced in this same twofold way, focusing on the practical examples above that I have observed in popular music recordings.

In his essay on musical collage, Nicholas Cook draws a link between Barthes’s two essays “The Grain of the Voice” and “The Third Meaning” (which is about film analysis). Barthes calls his “third meaning” “the one ‘too many,’ the supplement that my intellection cannot succeed in absorbing, at once persistent and fleeting, smooth and elusive,” and he labels this “the obtuse meaning.” For Barthes this is “outside (articulated) language while nevertheless within interlocution”—that is, part of the discussion but nonverbal—“a signifier without a signified, hence the difficulty in naming it.” Although he identifies this “obtuse meaning” within the context of his film analysis, he is clearly suggesting the same thing when describing the grain of the voice. Cook identifies this obtuse meaning as being repressed by the overt meaning—which is to say, in the context of the current discussion, that the message of the nonverbal sounds is missed as our attention is focused on the delivery of the verbal portion. What is missed in this process, according to Cook, is the “defamiliarization of the everyday; and its ineffability.” For Cook this resonates with other modern forms of creative expression such as surrealism. For me this calls to mind the “p’s,” “esses,” and breaths under discussion. And though these sounds may be simply lost to listeners if they focus only on the words being sung, they may also be actually removed by the power of digital audio editing. Such removal may mean that even the possibility
of tuning in to the ineffable is lost to the hegemony of the everyday (the lyrics alone).

Ultimately Barthes acknowledges that his idea of grain is a part of the ineffable quality of music, and, as much as he attempts to identify its qualities, music in general and grain in particular continue to resist any such identifications. Barthes opens his essay with a rhetorical question: “How, then, does language manage when it has to interpret music?” to which he answers: “Alas, it seems, very badly.” And to the extent that Barthes can locate grain, it is not connected primarily to linguistic expressiveness but in the close association of voice to physicality and secondarily to the intersection of the physical and the verbal. And again, the “p’s,” “esses,” and breaths under discussion are certainly a step away from musical expression (Fisher-Dieskau) and toward the corporeal (Panzera).

In Performing Rites Simon Frith balances the destructive and constructive forces of the recording process in a manner that also privileges something that may be akin to Barthes’s grain. In analyzing the interaction between the listener and the performer he notes, “The presence of even a recorded sound is the presence of the implied performer.” The use of “even” in this context belies the possibility of an enhanced presence of the recorded voice, and the subtext is the diminished capacity of recordings to call forth true “presence.” Here again I think that we have to consider the ways that recordings might augment a singer’s presence, without discounting the effect of the disembodied voice in recordings. Frith adds that the recorded performer’s voice does have a strong physical presence for the listener: “the performer [is] called forth by the listener—and this is clearly a sensual/sexual presence, not just a meeting of minds.” And, although it is in the context of differing qualities of popular music genres (in this case the reference is to “pop”), Frith acknowledges that recordings may surprise us with an added layer of intimacy: “It is as if the recording of music—its closeup effect—allows us to recreate, with even greater vividness, the ‘art’ and ‘folk’ experiences which the recording process itself destroys.” I especially like the use of “close-up effect” and “vividness” here to describe qualities of recordings (separate from the qualities of live performance) that might be closely linked to Barthes’s vocal grain. They also describe qualities that are easily associated with the prominent “p’s,” “esses,” and breaths that I have noted. Here Frith balances the unique powers of recording against the unique powers of live performance in a way that contradicts his previous use of “even” as a qualifier of recorded performances’ ability to create presence.
Personal Hallucinations

Early in Barthes’s “The Grain of the Voice” essay he asks, “Are we condemned to the adjective? Are we reduced to the dilemma of either the predicable or ineffable?” His answer is no—instead it is “better to change the musical object itself, as it presents itself to discourse, better to alter its level of perception or intellection, to displace the fringe of contact between music and language.” And Barthes proceeds to do just that, not only to challenge the language that is used to describe the experience of a singer’s voice but to challenge what is valued in “the musical object itself” to suit his own experience (and love) of music. His analysis, however, fails to recognize (or perhaps accept) the various ways the application of recording techniques may be participating in his relationship to this experience.

Here I have argued that the conscious manipulation of vocal sounds by the recording process participates in Barthes’s notion of grain. I have shown that Barthes’s idea of grain may be heard to reside both in the singer and in the recording of the singer. As the experience of grain is an individual response, this is a personal vision—but it is reinforced by the culture’s acceptance of sounds that have become so altered or exaggerated as to be considered new phenomena—a result of the digital age of recording and reproduction. Slowly but inexorably they have become a part of our aural vocabulary. As these sounds, especially nonverbal vocal sounds, take on new forms, they also have the potential to take on new meanings.

How did Barthes feel about the effects of recording? On the surface not very positively: “today, under the pressure of the mass long-playing record, there seems to be a flattening out of technique; which is paradoxical in that the various manners of playing are all flattened out into perfection: nothing is left but pheno-text.” But Barthes is only looking at the surface effects of the recording process on some performers. Were he to have allowed his own radicalizing viewpoint that created the idea of the “grain of the voice” to be focused on some of these artifacts of vocal recordings that I have described, he might have come to a broader appreciation for the recording process. Perhaps this would have happened had he turned his attention to certain popular singing traditions, as opposed to Fischer-Dieskau and Panzera. This would have forced him to encounter the effects of technology: “A whole tradition of popular singing, from crooning to bossa nova, is unimaginable without the microphone.”

If at the outset of the essay Barthes boldly launches the discourse out past the clearly predicable, in the midst of it he feels compelled to question
his endeavor. Is he reading qualities into voices? Is he the only one perceiving these sounds? “Am I hearing voices within the voice? but isn’t it the truth of the voice to be hallucinated?” He certainly doesn’t claim to have exhausted the significance of the phonetics that he is analyzing, and thus is perhaps opening the door for the arguments I am making. For Barthes, his work is most valuable in holding “in check the attempts at expressive reduction operated by a whole culture against the poem and its melody.”

He is trying to inject the creative element into his own analysis of creative expression. One may be accused of hallucinating if one strays beyond that which is easily predictable, but Barthes insists on the value, indeed the necessity of this wandering, in the study of musical meaning. This is not simply a value judgment. To look at this physicality of the voice, this grain, yields individual evaluation, but it is not “subjective”: it isn’t the subject that is reinforced, rather the intention is to lose the subject. It’s the value that is outside of culture and “hidden behind ‘I like’ or ‘I don’t like.’”

From the very beginning of sound recording the technology has interacted with the vocal source to alter the nature of the voice itself. As detailed above, contemporary recording techniques may further alter certain aspects of vocal production—especially those sounds in the mouth and throat that are of particular interest to Barthes and his notion of grain. I am arguing that these nonverbal effects, these hyperreal sounds from the mouth, also have the potential to pull us back to Barthes’s poem and melody. In the end is Dylan better understood with his popped “p’s”—or is Green Day’s message made stronger by Billie Joe’s amplified “esses”? Do these obscure or enhance our experience of these singers? In the same nonsubjective way Barthes describes, these artifacts may heighten rather than flatten, favor geno-text over pheno-text, perhaps adding grain as a part of the processing of recorded audio—though of course I may be hallucinating this. Is this “true grain” or is this artificial “grain”? Without judgment I contend that at least I, for one, do hear grain as a part of the ways recording has altered these vocal performances.

So the technologies of recording may feed Barthes’s abstract notion of grain. The technologies that today allow active participation in the manipulation of audio provide opportunities to eliminate or enhance elements within the ineffable grain. On the one hand the ideal of “perfection” as generated through musical “cleansing” and “fixing” is set against the notion of Barthes’s elemental, geno-typical grain of musical expression. This control of performance in the face of the unknown is characteristic of a culture that has been steeped in positivism and now rides the dual crest of
technology and information. On the other hand the ineffable mysteries of musical expression may be expanded in the interaction with technology. Some of the “unnatural” sounds of recording may become contributors to the underlying geno-type of Barthes’s grain. Ultimately, the cultural imperatives that are exemplified by the “fixing” and “cleansing” of musical performances weigh in on both sides of the scale that balances the art of Barthes’s geno-type and the artifice of his pheno-type.