

Nigel Fabb

## When Words Stop: Omission in Songs

### The omitted voice in “The Royal Scam”

The Steely Dan song “The Royal Scam” (1976) begins with instruments playing, then the first verbal line is sung accompanied by the instruments. This verbal line lasts seven seconds. After this, the instruments play without any words being sung for ten seconds. This omission of the verbal text is the omission which is the focus of this article. The next verbal line lasts seven seconds, followed by an instrumental section of ten seconds, which is again longer than the verbal line. Most songs in which instruments accompany the voice have an alternation between verbal+instrumental sections and just instrumental sections. But the pattern we see in “The Royal Scam” is unusual: here the instrumental sections are longer than the verbal+instrumental sections. In contrast, in most songs the instrumental sections are either shorter than the verbal+instrumental sections or, particularly in the style of blues songs, the instrumental sections and the verbal+instrumental sections are about the same length. The instrumental sections in a song are a type of verbal omission, because there are no words sung during this part of the song. In this essay I ask how the instrumental section can nevertheless contribute to the meaning of the song, even though it has no words. In particular I explore the rather specific effect, at least on me, that is produced by this pattern in “The Royal Scam,” in which the instrumental section is consistently longer than the verbal+instrumental section.

In this first part of the paper, I discuss the structure of “The Royal Scam.” The song was written by Walter Becker and Donald Fagen, and is the final song on the Steely Dan album of the same name, *The Royal Scam* (1976). While this song has a rare type of alternation, the other songs on the same album have three other types of alternation which can easily be found elsewhere. There is one song which is mostly instrumental with occasional and unsystematically placed verbal+instrumental sections (“The Fez”) and one song which has the blues pattern of equal length verbal+instrumental and instrumental sections (“Sign in Stranger”). In the other six songs on the album the verbal+instrumental sections are longer than the intervening instrumental sections: relatively short instrumental sections come between each of the lines of verse.

“The Royal Scam” tells a story of migrants from Puerto Rico to New York, and appropriates the voices of its characters, and Others them; Clements describes the narrative as “ironic.”<sup>1</sup> It is 6'30" (six minutes thirty seconds) long, with the large-scale structure shown in Table 1.

---

<sup>1</sup> Paul Clements, “Cultural Legitimacy or ‘Outsider Hip’?: Representational Ambiguity and the Significance of Steely Dan,” *Leisure Studies* 28.2 (2009): 189–206, 200.

Table 1. General structure of “The Royal Scam”

start	end	length (secs)	section
0'00"	0'33"	33	instrumental (introduction)
0'33	1'51"	68	stanza 1
1'41	1'51	10	instrumental (between stanzas)
1'51"	3'00	69	stanza 2
3'00"	3'52"	52	instrumental long section
3'52"	5'01"	69	stanza 3
5'01"	6'30"	89	chorus repeated, with breaks, and outro

At this level of structure, the pattern is common; it is only once we get inside the stanzas that we can see that something odd is happening. Table 2 shows this for the first stanza; the other two stanzas have an identical pattern.

Table 2. The structure of stanza 1 (including chorus) of ‘The Royal Scam’

start	end	length (secs)	section	length (bars)
0'33"	0'40"	7	line = verbal+instrumental	2.5
0'40"	0'50"	10	instrumental	3.5
0'50"	0'57"	7	line = verbal+instrumental	2.5
0'57"	1'07"	10	instrumental	3.5
1'07"	1'14"	7	line = verbal+instrumental	2.5
1'14"	1'24"	10	instrumental	3.5
1'24"	1'31"	7	line = verbal+instrumental	2.5
1'31"	1'36"	5	instrumental	1.5
1'36"	1'43"	7	chorus line = verbal+instrumental	2.5

In the first four lines, the alternation is clear: there is a shorter verbal+instrumental section of about seven seconds or two and a half bars, and a longer instrumental section of about ten seconds or three and a half bars. After the fourth line of the stanza there is a five second instrumental section before the second chorus; this five-second break is the shortest instrumental break in the song’s overall structure, and it gives the five-line group of stanza plus chorus a balance which I think is odd, because for a song with long pauses between lines, it is unexpected when the chorus rushes in after the last line of the stanza. This change at the chorus might function to emphasize even further the length of breaks between the lines of the stanza.

Each stanza is divided into four quite long lines, each of which is a clause (or subclause). In some transcriptions of the lyrics, the lines are broken up and presented as shorter lines (e.g., [genius.com] presents the stanza as eight lines whereas [sdarchive.com] presents the stanza as twelve lines). These transcriptions seem to be assuming that the four

lines are unusually long and so must be broken up. So we can add the unusual length of the lines to the unusual balance of lines to instrumental as a second oddity in the song. Each line is performed with five stressed syllables, matching the first and third beats of each of the two and a half bars. Overall the lines vary between sixteen and nineteen syllables, at least in the first stanza, so the metrical structure of the text is loose (i.e., based on counting stressed syllables not overall number<sup>2</sup>). The third and fourth lines of each stanza rhyme with each other, and this is the clearest evidence that the stanza consists of four long lines, rather than a larger number of short lines as in some of the transcriptions.

In terms of the musical form of the song it is worth noting that the melody of the solo verbal+instrumental part has a very narrow range, using only four notes B-flat, C, E-flat and F, all within the same octave. The choral verbal+instrumental line is slightly wider in range. The song is in the key of c minor, and each vocal line has its final syllable on the tonic note C, a note which is insistently present throughout the instrumental music. The instrumental section in contrast uses all seven notes, and over a much wider range of four octaves. This creates what will be called a “registral silence” in the vocal part, where most of the registers are silent or omitted; I return to this notion below.

I chose to discuss this song because I experience it as odd, and have done so for the nearly half a century since I first heard the song. There is something about the balance of verbal+instrumental to instrumental sections which feels significant. My response is what I call a type of “epistemic feeling,”<sup>3</sup> a feeling of knowing something which feels significant, without being able to specify what is known. Here it feels as if the music of the instrumental parts adds meaning to the words of the verbal+instrumental parts, perhaps commenting on them; and yet the musical parts are not expressive in the way of word-painting, or otherwise clearly commenting on the text. The effect seems to arise not just from having long instrumental sections but also from the fact that in these instrumental sections the words are explicitly omitted. Omission itself thus produces this epistemic feeling.

## Types of omission in songs

Songs exist in time, and so omissions of parts of the music have durations over which those omissions extend, and so the song can exploit those durations. Silence is the most complete kind of omission, and the use of silence has been widely explored in music, relative to the surrounding materials.<sup>4</sup> Consider for example the joke in the final movement of Haydn's String Quartet in E Flat, Op. 33, No 2, where there are harmonically complete apparent endings followed by silences, each interrupted by the music restarting, and the whole piece concluding in an actual ending which is harmonically incomplete as though there should be more to come.

When we consider the verbal and instrumental aspects of a song and how they might change in the course of the song, we can distinguish between three main types of section.

- (i) The verbal and instrumental components exist at the same time, as when someone sings while an instrument plays. This is a verbal+instrumental section.

---

<sup>2</sup> Nigel Fabb and Morris Halle, *Meter in Poetry: A New Theory* (Cambridge: Cambridge UP, 2008).

<sup>3</sup> Nigel Fabb, *Thrills, Sublime, Epiphany: How Literature Surprises Us* (London: Anthem Press, 2022).

<sup>4</sup> Thomas Clifton, “The Poetics of Musical Silence,” *The Musical Quarterly* LXII.2 (1976): 163–181; Andrew Edgar, “Music and Silence,” in *Silence: Interdisciplinary Perspectives*, ed. Adam Jaworski (Berlin; New York: Mouton de Gruyter, 1997): 311–328.

(ii) The instruments play but there is no verbal component, as might arise between lines or stanzas of the song, or before or after the singing. This instrumental section involves the omission of the verbal part.

(iii) The voice sings words but there are no instruments playing. For example, in James Brown & The Famous Flames' song "Please Please Please" (1956) there are moments where all the instruments fall silent, and we hear only James Brown's voice. This is a verbal section which involves the omission of the instruments (not discussed in this article).

There are other possibilities. For example where there are several instruments, we might find that a particular instrument only plays when there is no singing, as for example when a performer either sings or blows harmonica but cannot do both at the same time, so one or the other is omitted at any time. There can also be an alternation between two different types of verbal section, between the main voice and the chorus's voice if they are distributed in sequential sections. For example in James Brown & The Famous Flames' "Please Please Please," there is at all times a verbal component which is either solo or chorus, in alternation.

It is also possible for the voice to produce nonverbal sound, so that during the nonverbal instrumental sections, the singer might still produce sounds, including whoops, cries, vocables (meaningless word-like items), and so on. The boundaries between verbal and nonverbal vocal behavior are not always completely clear, and so the notion that the verbal is "omitted" can sometimes be a matter of interpretation. Consider for example The Clash song "Complete Control" (1977); in one of the longer instrumental sections during a guitar solo, the singer calls out "you're my guitar hero" (at 1'08"). Strictly speaking, this is a verbal+instrumental moment because these are words, and furthermore, this interjection relates to the overall meaning of the song, which is about songs and the music industry. However, these words are performed as if improvised, and are not part of the stanzas of the song, and so have an uncertain status as verbal components of the song. The lyrics as represented on [azlyrics.com] do not include this text, but they are included in the lyrics as represented on [songmeanings.com], and this difference illustrates the problem of deciding whether this verbal interjection is part of the song or not. The complexities here do not undermine the three-part typology suggested above; in fact, they depend on this typology to produce the ambiguity of this moment in the song. As another example of the complex role of nonverbal vocalizations, consider another song by The Clash, "White Man in Hammersmith Palais" (1977). At the very end of the song, after the verbal text has ended, there is a moment in which the instruments fall silent and in that silence the singer just breathes out loudly and emphatically. This is not a word, so not an example of a purely verbal section, but it feels quite expressive, even if its exact meaning is unclear.

Certain genres of song exploit the balance between verbal+instrumental and instrumental sections in a systematic way. This is most obviously true of some blues songs, which may alternate equal-length (or almost equal-length) verbal+instrumental and instrumental sections. The singer sings a line, and then the instrument or instruments play for the same length of time.<sup>5</sup> For example in Bessie Smith's performance of "St Louis Blues" (1925) the verbal+instrumental and instrumental sections are similar in length, though not exactly the same: the first line and the first instrumental are each about eight seconds, the second line and second instrumental each about six seconds, and so on, sometimes varying so that the two parts are not exactly the same length. More regular instances of this alternation

---

<sup>5</sup> Samuel A. Floyd Jr., *The Power of Black Music: Interpreting Its History from Africa to the United States* (Oxford: Oxford UP, 1997): 96.

can be found in rock 'n' roll songs, such as Big Joe Turner's "Shake, Rattle and Roll" (1954) in which lines of approximately three seconds alternate with instrumental sections of approximately three seconds, i.e., the verbal+instrumental and instrumental sections are of the same length; this changes in the chorus, where the instrumental sections are greatly shortened. Another regular example is Ray Charles' "I Got a Woman" (1954), where there is an alternation of equal length verbal+instrumental and instrumental sections, approximately two seconds long. It is easy to find songs with a similar structure: The Stooges' "Down on the Street" (1970) alternates equal-length verbal+instrumental and instrumental sections, as do the David Bowie songs "Station to Station" (1976) and "Breaking Glass" (1977).

Songs can have other types of long instrumental sections outside the alternation of text and music. Popular songs may have a single long instrumental section, sometimes an instrument solo, in the second half of the song. Songs also begin with long instrumental sections; Bowie's "Sound and Vision" (1977) has only occasional vocalizations until 1'28", after which the verbal part of the song dominates. Bowie's "Station to Station" (1976) has an instrumental section until 3'18" when the voice starts unexpectedly, and the song then follows a blues pattern of equal length verbal and instrumental sections. In Pink Floyd's "Wish You Were Here" (1975) the instrumental section extends until 8'42" after which the voice starts, and thereafter we hear the normal pattern in which verbal sections tend to be longer than intervening instrumental sections; this appearance of the norm after a very long introduction may make the introduction, retrospectively, even more marked. Songs can also have long final instrumental sections, as for example in Talking Heads' "The Big Country" (1978) whose final line ends almost a minute before the end of the song at 4'28" after which the singer vocalizes without words for almost thirty seconds until 4'55" followed by a fully instrumental section for the final thirty seconds to the end at 5'25". These long instrumental sections can appear to communicate meaning in the context of the song. In this Talking Heads song, language is explicitly disavowed by the final line "Its not even worth talking about those people down there" before the instrumental section, and in this context the omission first of words and then of the voice can both be taken as communicating a meaning related to the final line. In "Wish You Were Here" the long instrumental section includes a particularly prominent and expressive guitar, as though the guitar is 'speaking,' an interpretation I return to shortly.

Another type of omission in songs is complete silence, where neither voice nor instruments can be heard. In principle there may be silence before and after a song, and there can be brief silent sections in songs. It is worth noting however that even some silences still have musical content, so are not entirely empty. Consider the beat structure of a song. A beat structure is a psychological construct, not an acoustic one, but nevertheless part of the musical form of the song even though it cannot strictly be heard. The beat structure can continue during a silent section, and this can be shown on a musical score by using timed symbols for rests instead of symbols for notes. There are also symbolic ways of thinking of music as existing unheard in silence as though silence can still have musical content, as in Keats' "Heard melodies are sweet, but those unheard are sweeter" ("Ode on a Grecian Urn," 1819), or Paul Simon's "The Sound of Silence" (1965). And of course silence has an important role in contemporary classical music, as most significantly theorized by John Cage.<sup>6</sup> So the radical type of omission in music which is silence, even though it has no verbal or instrumental content, can sometimes also appear to produce meaning, though this meaning always arises in a context.

---

<sup>6</sup> Edgar, "Music and Silence."

Clifton and Edgar discuss another kind of omission which is found in all kinds of music, which is what they call a registral silence.<sup>7</sup> For example, the high notes, these being the notes in an upper register, may not appear at all in a piece of music, or may appear only very rarely; this constitutes a ‘silence’ in the upper register, a type of omission, even though there is sound in the lower registers. Clifton says that “certain registers are not occupied all the time – this enables us to identify structurally important tones,”<sup>8</sup> when for example a register which is not generally used is occasionally used, producing a markedness in the music. Edgar expands this notion to suggest that music can be thought of as the removal of aspects of sound, leaving just the sounds which compose the piece, like a sculpture which is produced by removing the surrounding stone: “music can be seen as the result of the paring away (or silencing) of all but a few possible sounds, so that the sounds that remain have their meaning due to the silences that have been created around them.”<sup>9</sup> In “The Royal Scam” the vocal melody manifests this type of ‘silence,’ compared with the instrumental melody which involves more different notes over a wider range. So, the words are omitted in the instrumental sections, and the vocal lines omit much of the musical range: these are two types of omission in alternation.

Clifton’s focus is on registers which are mostly silent but occasionally occupied, and he emphasizes that the effect on the listener, and the possibility of producing meaning, comes not from the omission in itself, but from the contrast between omission and presence of notes in the particular register. In this way, registral omission relates to the broader issue of change as a characteristic of a linear artwork such as a verbal text or a piece of music. Changes can have various effects on an audience, including emotional effect and implied meaning. As Huron shows, a change can be a surprise, and surprise involves an emotional response.<sup>10</sup> I have found, when exploring thrill responses to songs with my students, that textural thinning and thickening in a song are among the triggers of thrill responses. This is most obvious when there is a crescendo where the omission of loud sound is replaced by its presence, and this is extensively documented in experimental studies of the triggering of chill responses.<sup>11</sup> I find that thrills are triggered also by a sudden thinning out as when most of the instruments are omitted; my go-to example for this is the ending of Shostakovich’s fifth symphony. The key is D major and most of the instruments sustain the dominant note A during the final part, while the timpani drum plays an alternating pattern of the tonic D and the dominant A, then all the instruments cut out except for the timpani which keeps playing this pattern, and is joined at that moment by the bass drum. The effect is that for three beats all you hear is the thin texture but loud noise of the three beats of the drum. Then the rest of the orchestra comes back, for the last two bars to conclude on the tonic. I almost invariably get a chill response from this, and when measured experimentally, discovered that at this moment my breathing is shallower and longer (i.e., a freeze response); I think it is the sudden thinning of the texture which produces this, though the fact that this is a key structural moment in the piece, and there are other meanings swirling around, undoubtedly contributes to the experience. This is an example where a change involving an omission, in context, can produce a powerful response.

This is a change, and any change can also demand increased processing effort, because something new has to be processed, and increased processing effort itself can

---

<sup>7</sup> Clifton, “The Poetics of Musical Silence,” 171, and Edgar, “Music and Silence,” 312.

<sup>8</sup> Clifton, “The Poetics of Musical Silence,” 173.

<sup>9</sup> Edgar, “Music and Silence,” 314.

<sup>10</sup> David Huron, *Sweet Anticipation: Music and the Psychology of Expectation* (Cambridge, MA: MIT Press, 2006).

<sup>11</sup> Jaak Panksepp, “The Emotional Sources of ‘Chills’ Induced by Music,” *Music Perception: An Interdisciplinary Journal* 13.2 (1995): 171–207.

produce an emotional response.<sup>12</sup> In relevance theory, increased processing effort, when demanded by a communicative act, licenses the derivation of increased cognitive effects.<sup>13</sup> This means that any change in a text, so long as it is ostensive, can give the reader or hearer permission to derive additional meanings beyond the literal. All the changes discussed here, including registral and textural changes and the changes from verbal+instrumental to instrumental parts of a song, might in principle demand increased processing effort, with its emotional and interpretative consequence. This is the general relevance-theoretic principle of how style can produce meaning without coding it: by demanding increased processing effort, stylistic devices including stylistic changes, can thereby legitimate further interpretation of the text. But it cannot be that all changes produce these effects, because a song or other text tends to change all the time, in various ways; so the change must itself have something noticeable about it, which draws our attention. Since increased attention is equivalent to increased processing effort, it is the noticeable changes, the oddities, which in principle can produce the greatest effects, including the epistemic feelings which I suggest are produced by the odd alternations of verbal+instrumental and instrumental sections in “The Royal Scam.”

### **How do the non-verbal parts of a song produce meaning?**

I take a relevance theory approach to meaning.<sup>14</sup> I begin by assuming that songs are (at least in part) communications, and so fall under relevance theory: the hearer is induced to engage in interpretative effort with the guarantee that the effort will be rewarded by cognitive effects. This means that the meaning of the song is the set of thoughts which the hearer attributes, with strong or weak confidence, to the singer or composer as what they intended to communicate. The text of the song, and other aspects of the song, all contribute evidence which the hearer uses to derive the intended meanings of the song. The evidence is interpreted in context, including real-world knowledge, knowledge of the band, the other songs on the album given that the whole album has the same name as the song, and the album’s cover picture which specifically relates to this song. For our purposes, the two most important parts of the song are its verbal text and its musical form.

The words in their syntactic context provide the most specific evidence for meaning, because words have coded meanings and the syntactic structure also contributes to the coded meaning of the text. These coded or literal meanings may not be the intended meanings: for example, the literal meaning is not the intended meaning in a metaphor. Furthermore, a general characteristic of songs is that they tend to have a small vocabulary with much repetition and an attenuated syntax, with juxtaposition and listing replacing complex sentence structure.<sup>15</sup> (However, neither are true of this song which has a rich vocabulary and complex sentences, with only the chorus line repeated.) Nevertheless, language gives the listener the richest and most detailed of the clues needed to make sense of the song. The clues encoded by the verbal text then form part of the basis for inferring what meanings the singer or composer intend to communicate. The particular focus of this paper is on the places where words are omitted and whether the instrumental sections between lines also contribute to the

---

<sup>12</sup> Nigel Fabb, “Processing Effort and Poetic Closure,” *International Journal of Literary Linguistics* 5.4 (2016): 1–22; Nigel Fabb, *Thrills*.

<sup>13</sup> Dan Sperber and Deirdre Wilson, *Relevance: Communication and Cognition*, second ed. (Oxford: Blackwell, 1995).

<sup>14</sup> Sperber and Wilson, *Relevance*.

<sup>15</sup> Nigel Fabb, “Why is Verse Poetry?” *PN Review* 189 36 (2009): 52–57; Elizabeth Hellmuth Margulis, “Repetition and Emotive Communication in Music Versus Speech,” *Frontiers in Psychology* 4.167 (2013): 1–5.

implied meaning of the song, and particularly the extra-long instrumental sections in “The Royal Scam.”

This raises the familiar question of how music in general is able to communicate meaning. Following Schlenker,<sup>16</sup> we can say that there are two ways in which music can provide evidence for the communicated meaning of the song. They both depend on inference, because in music there is no equivalent to the lexical and syntactic codes which translate verbal sound into verbal meaning.

The first way by which music can communicate meaning is by what Schlenker calls a source-based semantics. When we hear music, we draw on our general acoustic knowledge of how things sound in the world, by which we can judge that a certain kind of sound might imply a possible source. A loud sound can come from a potentially large source; a crescendo can come from a source approaching the listener, and so on. We do not necessarily turn off these inferences when we are listening to music and so the acoustic properties of music can be attributed to imagined sources of various kinds, even if we know that the actual sources of the music are quite different, that the music is produced by musical instruments. In this theoretical approach, listeners to a song do not switch away from ordinary auditory cognition into a specialized musical cognition; instead the ordinary auditory cognition is always present and impinges on our aesthetic experience. This fits with some aspects of contemporary cognitive anthropology<sup>17</sup> and Fabb<sup>18</sup> argues for example that doubles have an uncanny effect on us because they contradict deeply embedded knowledge that the world is characterized by constant variation and not exact repetition. It is in contrast to approaches to aesthetic experience which see them as separated off from ordinary experience, and where ordinary experience is not prior or always underlyingly present. This holds for example of the frame-based approach to aesthetic experience, according to which, when we listen to a song we listen to it relative to specific frames, as an aesthetic object, as a performance, and so on.<sup>19</sup>

The source-based semantics helps us understand how the musical aspect of a song can seem to involve word-painting, where some aspect of the music appears to parallel the meaning of nearby words, perhaps by a type of source-based semantics in which we recognize the parallel between an acoustic phenomenon and some other real-world phenomenon. In the Talking Heads song noted earlier, “The Big Country,” when the word “good” is sung, the melody goes up in pitch (at 3'07"). While there is no coded relation between a rise in pitch and the notion of good, it is possible to think of a rise in pitch as somehow related to goodness, perhaps via a source-based semantics in which happy people produce sounds which are higher in pitch. In Thomas Weelkes' song “Thule, the Period of Cosmography” (1600), the words “doth vaunt” are sung to a rising melody; here the source-based semantics attributes something which rises (vaunts) as analogous to a rising melody. In Michael Tippett's setting of Hopkins's “The Windhover” (1942), the word “morning” is accompanied by a rise in pitch; again, it is possible to find a relation between the two, perhaps via a more indirect source-based semantics which links the rising of the music with

---

<sup>16</sup> Philippe Schlenker, “Prolegomena to Music Semantics,” *Review of Philosophy and Psychology*, 10.1 (2019), Vol.10(1), 35-111.

<sup>17</sup> Dan Sperber, “Why Are Perfect Animals, Hybrids, and Monsters Food for Symbolic Thought?” *Method & Theory in the Study of Religion* 8.2 (1996): 143–169; Pascal Boyer, “What Makes Anthropomorphism Natural: Intuitive Ontology and Cultural Representations,” *The Journal of the Royal Anthropological Institute* 2.1 (1996): 83–97.

<sup>18</sup> Fabb, *Thrills*.

<sup>19</sup> Richard Bauman, “Verbal Art as Performance,” *American Anthropologist* 77.2 (1975): 290–311; On silence and frames: Adam Jaworski, “Aesthetic, Communicative and Political Silences in Laurie Anderson's Performance Art,” in *Silence: Interdisciplinary Perspectives*, ed. Adam Jaworski (Berlin; New York: Mouton de Gruyter, 1997): 15–35, 17.



the morning rising of the sun. Schlenker discusses how the beginning of Richard Strauss' *Also Sprach Zarathustra* (1896) can be interpreted as its rising music resembling a sunrise. He suggests that these resemblances depend on "preservation principles"<sup>20</sup> whereby certain kinds of musical change (such as a rising melody) involve relationships which are analogous to certain kinds of nonmusical change (something gets better, the sun rises, someone leaps, etc.). It is worth noting that while all these examples happen to have a rising melody, nevertheless they can have different meanings. This is a reminder, if needed, that musical meaning is always inferential, and that if the instrumental sections in "The Royal Scam" have meaning, it must be inferential meaning not coded meaning.

The second way in which music can provide evidence for the communicated meaning of the song is by the form of the song, and in particular how the song progresses, either according to expectations, or violating expectations. Edgar notes that there is a long tradition of interpreting the meaning of a piece of music in terms of the musical structure of the work (and in fact refers to relevance theory).<sup>21</sup> Meyer and Huron emphasize how the meaning of the musical piece depends on manipulating the specifically musical expectations of the listener.<sup>22</sup> This is relevant to "The Royal Scam" in that the song has a structure of a type which is very familiar, in which verbal lines alternate with instrumental lines, but it consistently violates this structure by having the instrumental lines consistently longer than the verbal lines; in this way, the song may provide evidence for meaning by manipulating the expectations which we have developed on the basis of our knowledge of music.

One of the specific questions to be asked about this song is what generic expectations we bring to it; these will influence how we interpret both the words and the musical form. We might bring to the song our expectations of songs in which the intervening instrumental parts are significantly shorter than the verbal lines. Or we might bring to the song our expectations of songs (in the blues tradition) in which the instrumental parts are about the same length as the verbal lines. Both types of song have already been heard on the album by the time we get to this final song, and it makes sense that both types of expectation may play a role in producing meaning from this song. If we start by interpreting it as a variation on the blues structure, then have available to us the notion that made of the blues structure involves a verbal call and an instrumental response, as though the instrument is speaking back to the words.

Floyd draws on Gates's notion of Signifyin(g) (1988) to describe this: "A twelve-bar blues in which a two-measure instrumental 'response' answers a two-measure vocal 'call' is a classic example of Signifyin(g). Here the instrument performs a kind of sonic mimicry that creates the illusion of speech or narrative conversation."<sup>23</sup> Similarly, Evans says of the alternation of verbal+instrumental and instrumental sections in blues, "Blues also elevated the role of the musical instrument within popular song making it a second voice, integral to the song itself, punctuating, commenting upon, and answering the vocal line."<sup>24</sup> So if we interpret "The Royal Scam" as a variation on the blues pattern, then a meaning we can infer from the song is that the instrumental lines respond to the verbal lines but in this particular song, because they are longer, the instrumental parts have more to say in response than the verbal lines. The structure of the stanzas thus emphasizes to the listener that the words are only a part of the meaning of the song, and that there is much more that is not explicitly said but can only be inferred. This fits well with a broad interpretation of the song as involving

---

<sup>20</sup> Schlenker, "Prolegomena to Music Semantics," 74

<sup>21</sup> Edgar, "Music and Silence," 321; reference to relevance theory in a footnote on page 324.

<sup>22</sup> Leonard B. Meyer, *Emotion and Meaning in Music* (Chicago: Chicago UP, 1956); Huron, *Sweet Anticipation*.

<sup>23</sup> Floyd, *Power of Black Music*, 96.

<sup>24</sup> David Evans, "The development of the blues," in *The Cambridge Companion to Blues and Gospel Music*, ed. Allan Moore (Cambridge: Cambridge University Press 2002): 20-43, 21.

meanings which are concealed or not fully expressed in the lyrics, and in particular that the song is communicating that the lure to move from Puerto Rico to New York is a “(royal) scam.” This fits also with the narrative in the final stanza where an old man back home reads a letter which misdescribes the Puerto Rican protagonist’s experience, presenting it as positive when in fact it was negative, and fails to say what needs to be said. But we can also read the song as a variation on a different genre, the type of song in which lines are long and instrumental sections short. In this reading, the song violates a structure where the text is long and the instrumental short. This reading produces the interpretation that the instrumental sections are particularly long, and this, added to the interpretation of the instruments as responding further emphasizes that the parts in which the words are omitted have the most to say. If we draw on a source-based semantics we might further associate the greater size of the instrumental sections with a meaning which is more important by the analogy of size with importance. There are two alternative genres, or song structures, relative to which this song can be interpreted; in principle the listener could interpret the song relative to both of them, particularly as the readings in the two cases reinforce each other.

Floyd’s and Evans’s characterization of blues songs suggests that the instrumental part is heard as if it is in language. Songs in other genres can be interpreted in a similar way, as though the music is language. For example Bullock discusses a song by Tchaikovsky in which a “non-existent dialogue is expressed by the mirroring of the descending vocal part by an ascending line in the piano accompaniment.”<sup>25</sup> Hearing music as speech, but in the absence of speech, seems to be common, and perhaps it draws on a source-based semantics, whereby characteristics of speech are preserved in the comparison with characteristics of music. Our knowledge of the characteristics of speech is presumably more basic than our knowledge of the characteristics of music. This suggests that we do not ‘turn off’ our tendency to hear sound as speech, even when we know it is coming from an instrument. Juslin makes the interesting claim that in listening to music we seem to be hearing a language whose syntax we understand, but not its semantics:

In listening to music, there is a strong sense that something ‘highly structured’ and ‘meaningful’ is being said, but our brain cannot make out what it is. The resulting feeling might be construed as ‘mild excitement’ mixed with ‘confusion’ as our brain is continuously ‘hooked’ by this subtle yet inexplicable language-like structure that we call ‘music.’<sup>26</sup>

Expressing the same basic idea, here is Robert Browning, a poet who had a great interest in music (from “A Serenade at the Villa,” 1855):

What they could my words expressed,  
O my love, my all, my one!  
Singing helped the verses best,  
And when singing's best was done,  
To my lute I left the rest.

A final general point to make about instrumental sections in a song relates to time, processing and memory. The verbal line is followed by a nonverbal part, and in principle the omission of words in this part could allow for the line to be remembered and processed more deeply, compared with if another line immediately followed. In this song, by having such

---

<sup>25</sup> Philip Ross Bullock, “Ambiguous Speech and Eloquent Silence: The Queerness of Tchaikovsky's Songs,” *19th Century Music* 32.1 (2008): 94–12, 114.

<sup>26</sup> Patrik N. Juslin, “From Everyday Emotions to Aesthetic Emotions: Towards a Unified Theory of Musical Emotions,” *Physics of Life Reviews* 10 (2013): 235–266, 261.

long instrumental sections, more time is given to the listener to interpret the lines; the predictability of the pattern enables processing effort to be managed, with the expectation that there will always be space to further understand each line. Deeper processing allows more meaning to be inferred from the lines, and also enables the line to be better remembered.<sup>27</sup> A complicating factor is that the line is not followed by silence but by a musical section which also demands processing effort. This will take away some processing from the preceding line, but on the one hand, by increasing processing complexity, the line will be more deeply processed and better remembered. The second factor specifically in this song is that the instrumental section does not vary much from section to section, and so in principle requires less effort each time we hear it.

### **How does meaning arise from the instrumental sections of “The Royal Scam”?**

In conclusion, it is possible to suggest various ways in which the instrumental sections can contribute to the meaning of the song, even though words are omitted. The first points to make are that all such meanings are derived by inference, that a range of possible meanings can be derived, and that it is difficult to generalize over the meanings. A related point is that it may be that the hearer feels that meanings have been implied, but that they are so indeterminate that it is impossible to fix on any of them; this might give the sense of an epistemic feeling of knowing something without being able to say what it is. The third thing to say is that there may be a tendency, theorizable under a source-based semantics, to hear instrumental sections in a song as though they are in language; this further adds to the feeling that meaning has been implied but without saying what that meaning is. As part of this way of hearing the instrumental sections as verbal, we might interpret them as responses to the verbal sections with which they alternate. This particular song has particularly long instrumental sections, which has two kinds of consequence. First, it makes the instrumental sections particularly marked, noticeable, and ostensive, and this adds to the sense that the singer or composer is trying to say something to us during these sections. Second, the length of the sections means that they can be interpreted as saying more than the shorter verbal lines can say. However, it is worth noting also that the instrumental sections tend to be quite similar throughout the song and so we might also attribute variations on the same response to all these instrumental sections, as though each line is commented on in the same way.

At the beginning of the essay I described my response to this song as a type of “epistemic feeling,” a feeling of knowing something which feels significant, without being able to specify what is known. It is precisely the fact of omission, in which the verbal component which supplies richest kinds of evidence for meaning is omitted, which produces this effect.

---

<sup>27</sup> David P. McCabe, “The Role of Covert Retrieval in Working Memory Span Tasks: Evidence from Delayed Recall Tests,” *Journal of Memory and Lang* 58 (2008): 480–494, 481.