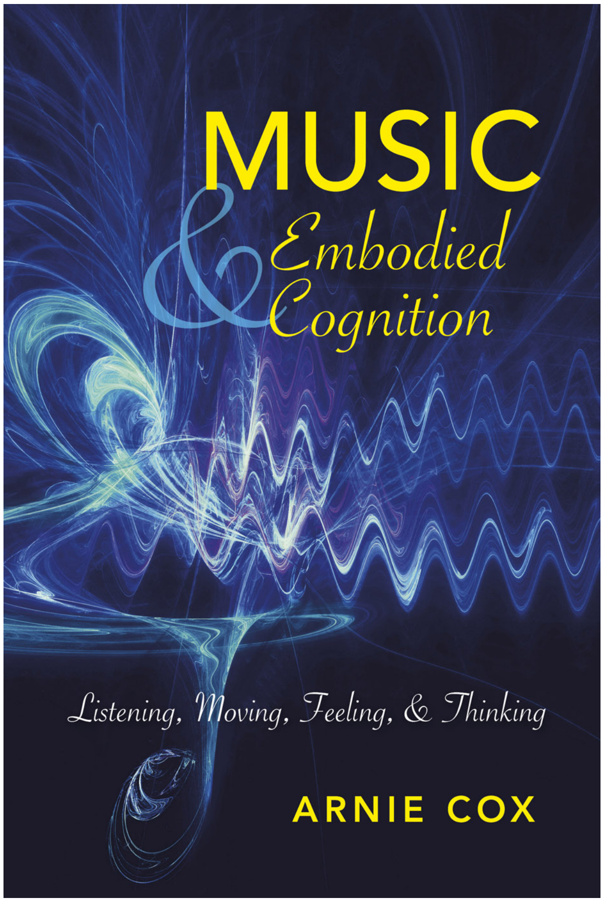

1 Mimetic Comprehension

Imitation is not just the sincerest form of flattery—
it's the sincerest form of learning.

George Bernard Shaw

When people are free to do as they please,
they usually imitate each other.

Eric Hoffer

If music cognition is embodied in a musically meaningful way, in the flesh of experience, then we ought to be able to specify just how this occurs. One way begins in imitation of musical sounds and of the physical exertions that produce them. This bodily comprehension of sounds and of sound-producing actions is one of the bases of embodied cognition of music, and it is the central basis that we will be exploring in the following chapters.

The issue of musical embodiment may be relatively straightforward in the case of performers, in the sense that performing, planning, and otherwise thinking about musical performance are tied to the bodily actions of performance. But the situation is less straightforward in the case of listeners: How and why would listening to or thinking about music, apart from planning or recalling one's own performance, have anything to do with embodiment beyond the operations of the auditory system? The answer offered here is that listening to, recalling, or otherwise thinking about music involves one or more kinds of vicarious performance, or imitation (or simulation), and that the role of this imitation in music is a special case of its general role in human perception. The gist of this idea is not new, but the details of how it actually plays out in music comprehension will take some time to describe.

By imitation I mean not only the overt behavior of "monkey see, monkey do" but also covert imitation that occurs only in imagination. These forms of imitation occur whenever we attend to the behavior of others, whether in the performing arts or athletics, or in learning a particular skill from someone else's demonstration, or in merely taking an interest in what others are doing. When we imitate overtly or covertly, in effect we are responding to two implicit questions:

12

What's it like to do that? and its twin question, *What's it like to be that?* We answer these questions in part by overtly and covertly imitating the behavior of others.

Overt imitation is plainly evident in children but it is also evident later in life. Music lessons and foreign language classes, for example, involve a measure of deliberate overt imitation. But imitation also regularly occurs covertly, involuntarily, and without our awareness, and I will try to clarify the importance of this. Because the term "imitation" bears unhelpful connotations, such as a lack of originality and/or lack of sophistication, I will in most cases describe imitative behavior as *mimetic*.[1] By *behavior* I mean not only overt actions, as in singing along with music, but also the behavior of muscle-related portions of the brain. Since overt mimetic behavior is plainly evident, it will not require much investigation in this context. The covert processes, however, some of which are conscious and some of which are not, will require more attention.

I will refer to overt mimetic behavior as *mimetic motor action* (MMA), and for the relevant muscle-related brain processes that do not manifest in overt actions I will use the term *mimetic motor imagery* (MMI): *mimetic* for imitative, *motor* for muscle related, and imagery for "thought," "imagination," or "mental representation." I intend *imagery* to include not only voluntary and conscious forms, but especially those forms that occur automatically and with or without our awareness. The involuntary and nonconscious forms of MMI are in some respects the most significant in the construction of musical meaning.[2]

It is important to distinguish *imagination*, as the term is commonly used, from *imagery*. When I imagine playing the cello, for example, this is normally a conscious and deliberate enactment of motor imagery, and when I imagine playing the cello like Jacqueline du Pré, this is conscious and deliberate MMI and is thus a special case of MMI generally. MMI is grounded in motor-related brain processes that occasionally become conscious and occasionally are initiated deliberately.

As a whole, the various forms of mimetic behavior (MMA and MMI) constitute the core of the *mimetic hypothesis* (Cox 2001, 2011), whose initial principles are the following:

- Part of how we comprehend the behavior of others is by imitating, covertly (MMI) or overtly (MMA), the observed actions of others.
- Part of how we comprehend music is by imitating, covertly or overtly, the observed sound-producing actions of performers.

Both of these propositions lead immediately to many questions. In the case of mu-sic these include the questions of how this might apply to, say, electronic music, in which the sounds are not produced directly by human exertions, or to ensemble music, where the various performers may be doing quite different things—for example, do listeners somehow imitate the percussion, the winds, and the strings of an orchestra all at once? They also raise the question of how differences in performing experience shape the different experiences of individual listeners. For example, listening to violin or fiddle music will offer a different

13

mimetic experience for string players than it will for other listeners, and the same applies to every performance medium. A theory of embodied music cognition must accommodate such variables, and I address these and other questions in the next chapter, where I describe twenty principles of the mimetic hypothesis. The value of the hypothesis and its implications, however, depends on the evidence for the hypothesis in the broader context of general mimetic behavior, and that is the purpose of the present chapter.

Because it will be a few pages before I get to the details of the hypothesis, I offer a preview of its principles here. The first nine apply to mimetic comprehension generally, while the others are more germane to music, and for the most part all are listed in order from more general to more specific. While each principle is integral to the hypothesis, principles 6–7, 9–13, and 16–20 are among the most significant for the approach to musical experience and embodied cognition described in subsequent chapters.

1. Sounds are produced by physical events; sounds indicate (signify) the physicality of their source
2. Many or most musical sounds are evidence of the human motor actions that produce them
3. Humans understand other entities (animate or not, human or not) and events in their environment in part via mimetic behavior (MMI and MMA)
4. MMA and MMI are bodily representations of observed actions
5. Mimetic comprehension is based on visual, auditory, and/or tactile information:
   - The observed behavior can be seen but not heard (the sight of action)
   - The observed behavior can be heard but not seen (the sound of action)
   - The observed behavior need not be seen or heard (the feel of action)
6. Musical imagery is partly motor imagery
7. Mimetic behavior (MMI and MMA) involves the variables of volition, consciousness, and overtness:
   - Mimetic behavior can be voluntary, but often it is involuntary
   - It can be conscious, but often it is nonconscious (beyond awareness)
   - It can be overt, but often it is covert (occurring only in imagery)
8. MMI and MMA are more strongly activated in observation of goal-directed actions
9. MMI and MMA occur in real time, recall, and planning
10. MMI and MMA take three forms:
    - Intramodal, or direct-matching (e.g., finger imitation of finger movements)
    - Intermodal, or cross-modal (e.g., subvocal imitation of musical sounds generally)
    - Amodal (abdominal exertions that underlie limb movements and vocalizations)

14

11. Any and all acoustic features can or will be mimetically represented: pitch, duration, timbre, strength (acoustic intensity, or "volume"), and location
12. Different kinds of music "invite" (motivate) different kinds of mimetic engagement, and this contributes to the different feel (quale) of different kinds of music
13. Music is sometimes found to "resist" mimetic participation
14. Ensemble music offers simultaneous multiple "invitations"
15. MMI and MMA can be stronger in live performance than in recorded performance
16. MMI and MMA vary in strength and accuracy among different people
17. Mimetic participation results in a sense of belonging and shared achievement
18. Mimetic participation is a central source of musical affect
19. MMI and MMA motivate and constrain conceptualization (metaphoric and otherwise)

20. Mimetic comprehension is part of human cognition generally

I am referring to this as a hypothesis because most of the principles are empirically testable but for the most part have yet to be tested directly in musical contexts. Nevertheless, the evidence presented below leaves little question that music is comprehended mimetically, and instead it leaves only the more specific questions of (1) the extent to which this is so, (2) the manner in which it plays out in different contexts, and (3) its implications for musical meaning.

Readers who happen to be familiar with theories of entrainment and/or simulation (e.g., Barsalou 1999 and 2009; Jeannerod 2001) will find overlap with the mimetic hypothesis. The much discussed "mirror neurons" (e.g., Iacoboni 2008) are also relevant, although we will consider some of the complexities that arise in trying to specify their likely role. Within music scholarship, the mimetic hypothesis is similar to ideas in Lidov (1987), Todd (1995), Cumming (1997 and 2000), Mead (1999), Leman (2008), and numerous others.[3] Of particular note are two ideas in Cusick (2006). The first is her description of a listener's desire to be the music, which is also one of the implications of the hypothesis and one that will force us eventually to define *the music* in light of mimetic engagement. The other is the notion of responding to an *invitation to participate*, which for all intents and purposes is identical to principle 12.[4]

What distinguishes the mimetic hypothesis from related writings is the more comprehensive view of imitation in music perception and of its role in our affective-cognitive responses to music. I begin by considering some of the evidence for the hypothesis.

Evidence for the Mimetic Hypothesis

The majority of the evidence comes from areas outside of music, in the form of psychological studies of overt mimetic behavior (MMA, mimetic motor action)

and neurological studies of covert mimetic behavior (MMI, mimetic motor imagery). In order to keep the focus on music, I have selected studies that are most closely related to music comprehension. I have grouped the evidence into the following overlapping categories:

1. Psychological Studies of Imitation
    1.1. Child-Caregiver Interactions
    1.2. Social Interactions in Adulthood
2. Neurological Studies of MMI and MMA in General
3. MMI and Auditory Perception: Neurological and Psychological Studies

3.1. Speech
3.2. Vocal and Instrumental Music

Discussion of the hypothesis requires a couple of novel terms. *Mimetic comprehension* refers to the portion of music comprehension that involves MMA and MMI. *Mimetic participation* emphasizes the joining-in and taking-part that result from MMI and MMA. *Mimetic engagement* refers to the more general aspect of merely being engaged with the music as a listener, and one of the claims to be explored is that whenever we are engaged in listening, normally we are *mimetically* engaged whether we are aware of it or not.

Although *mimetic perception* might be an apt term, I will speak most often of *mimetic comprehension* because the familiar use of "perception" is largely if not entirely nonmimetic (that is, not involving the mimetic processes that I am describing here). Once the arguments of the following chapters have been made, mimetic perception can then be understood as a form of perception that is complementary to our more traditional understanding of music perception. Along these lines, I am taking *cognition* to be the sum of the processes of coming-to-know and coming-to-understand and to thus subsume all forms of perception, comprehension, and conceptualization.

Finally, it will be helpful on occasion to use the term *mimetic representation*. This is defined in the discussion of principle 4 in the next chapter, but for now we can think of mimetic representations as activity in the muscles (MMA) and/or the motor-related portions of the brain (MMI) that involve imitation as a direct response to music—for example, singing along with a melody or dancing to a song are two kinds of mimetic representations of music. A mimetic representation is thus a kind of copy that we make, or that we embody, as part of how we perceive and comprehend something exterior to us.

1.1. Psychological Studies of Imitation: Children and Caregivers

Human development is saturated with overt imitation, where "monkey see mon-key do" describes a significant portion of our social lives as children. I want to suggest, in a manner similar to Walton's discussion of the practice of make-believe (Walton 1990), that the overt imitation we practice as children (MMA)

16

remains a part of how we participate with and understand others in the world, and that rather than outgrowing imitation as adults, a greater proportion of imitation gradually takes the form of MMI. The development of mimetic motor imagery occurs in tandem with the development of nonmimetic motor imagery (imagined actions that are not directly imitative of an observed action), and together the two constitute the more general category of *motor imagery*: imagination of one's own actions, including planned, recalled, and otherwise imagined singing and playing instruments. With this in mind, let us examine the nature and breadth of imitation in our early lives, first in general and then in musical contexts.

General Imitative Behavior

Infant studies confirm what might seem rather unremarkable: as infants, we imitate the vocalizations, facial expressions, and gestures of others around us.[5] These studies make plain the pervasive and fundamental role of imitation in how we learn to take part in and make sense of the world from the very start.[6] There is one significant feature of infant-caregiver interactions, however, that may not be obvious at first, and it is the mutuality of imitation in these situations: not only do infants imitate parents and other caregivers, but parents and other caregivers likewise imitate infants (Malloch 1999–2000). As Ulric Neisser puts it, "What is perceived is not merely the other's behavior, but its reciprocity with one's own. Both participants are engaged in a mutual enterprise, and they are aware of that mutuality" (Neisser 1976, 10). While it might not be surprising that we should imitate others as part of the process of learning to be fully human, we should ask why adult caregivers would imitate an infant. One answer is that, for infants, a caregiver's imitative behavior demonstrates at least two things: that as infants we are capable of generating a like response in others, and, since a like response demonstrates a basic level of understanding, in observing such a response we implicitly learn that we are capable of being understood. If it is a basic human desire to understand and to be understood, then mutual imitation helps satisfy this desire for both parties.

In imitating an infant, a caregiver feels something of what it is like to behave in the particular ways of the infant. In thus answering the implicit question of what it's like to do and be what this other little person-in-development is doing and being, adults enact mimetic exertions that produce mutual facial expressions, vocalizations, and other movements. These exertions have an affective dimension, in the feeling of what it is like to move in a certain way, and because such feelings normally correlate with emotional and mental states, it contributes to adult inferences as to the emotional and mental states of infants. One especially significant state is that of desire, as indicated by the actions of looking and reaching, and on the basis of which adults infer intentionality on the part of infants (Stern 1985). Mimetic comprehension of the actions of others contributes to such inferences in adult-adult interactions as well.[7] For example, I can believe that you understand me when you demonstrate that you share my state—when you show that you “feel” me by physically mimicking my facial expression, posture, gestures, and vocalizations (more on this below). In adult interactions with children and with other adults, mutual mimetic comprehension contributes to the affective reward that comes with mutual understanding.

Vocal-Musical Mimetic Behavior

The special case of imitating the music-making actions others begins in infant-caregiver vocalizations. Parents spontaneously model vocalizations that are especially for infants, with features that include the exaggerated melodic modulations of infant-directed speech (“motherese”), and they actively encourage and reward imitation (Papoušek and Papoušek 1982; H. Papoušek 1996; M. Papoušek 1996). As noted above, parents also imitate the infant’s vocalizations, which not only demonstrates understanding but also models mimetic participation: “Do what I’m doing, which is imitating my interlocutor.” In other words it is not simply that the impulse to imitate is innate, but that we learn that mimetic interactions with others is normal and emotionally rewarding behavior.

As infants we begin mimetic and nonmimetic cooing around eight weeks (Meltzoff and Moore 1994), on our way to the more complex and continuous vocalizations of speech and song.[8] “Infants usually stimulate an affectionate adult, male or female, to extended poetic or musical speech, which often moves into wordless song, or imitative, rhythmic and repetitive nonsense sounds. This distinctive style of adult speech is … attended to and responded to with much pleasure by infants. It varies with the age and stage, and motives and emotions of the infant partner” (Malloch 1999–2000, 30). Malloch goes on to describe how infants and parents attune their mutual vocalizations in timing, contour (between low and high pitch), and timbre, and that these combine with nonvocal bodily movements in creating a shared narrative. This behavior is both educational and aesthetic, and as much as one’s abilities may change from infancy to adulthood, it is difficult to miss the physical and emotional similarities with adult-adult mimetic participation in sing-alongs and other performance situations where we take part via some form of shared embodiment. Dissanayake (2000) makes similar observations in presenting a case for mother-infant interactions as the origin of the temporal arts.

Cross-Modal and Amodal Mimetic Behavior

One principle of the mimetic hypothesis is that imitation also occurs cross-modally, as demonstrated in numerous videos on YouTube that feature infants and toddlers spontaneously dancing to music: the singing is in one physical modality, the playing is in another, and the dancing is in yet another, and yet somehow they all fit together. One could say that all of them match “the beat,” but it will be helpful to be more specific. In such cross-modal imitation there is a pattern of exertions (rhythm), with a particular intensity (strength), that recurs at some rate (tempo). This composite can be manifest in the specific forms of singing, playing, and/or dancing, and at the core of such cross-modal mimetic participation

18

is a theoretical exertion schema: a shared pattern that can be manifest in various muscle groups.[9] Although culture shapes our responses to music, this is a shaping of what appears to be a spontaneous, innate response. The next example supports this conjecture more directly.

Many or most children who move to music have previously observed, or are concurrently observing, others moving to music, and so their dancing could be partly or largely intramodal imitation of dancing. But consider the example shared by Trevarthen (1999–2000) of a video of a mother cradling and singing to her blind infant. The infant is moving her right arm in time with the song in a manner that loosely resembles conducting or simply beating time. At one point the song is interrupted when someone enters the room to speak with the mother, and the infant’s arm movements, too, are likewise interrupted. The infant makes a couple of isolated conducting movements, as if wanting to resume the pattern, and when the mother eventually resumes her song, the infant resumes the matching pattern of arm movements.[10] In this example the mother and child are using different muscle groups—those of the voice and those of the arm—but they are sharing the same temporal pattern, rate of recurrence, and intensity: the gentle exertions of the singing and the gentle exertions of the arm movements.

In the cross-modal imitation just described, notice that there is an additional form of mimetic exertion shared by both participants: the abdominal exertions that anchor both vocalizations and arm movements, which I will be referring to as an amodal form of mimetic behavior. In one sense abdominal exertions are of course modally specific, in that they involve the specific muscle group of the core, but I am referring to them as amodal because they are activated in most if not all of the limb movements and vocalizations, including those involved in musical performance.[11] Abdominal exertions often are not as salient as the limb and vocal exertions that they anchor, and their relevance is thus easily overlooked, but trauma to the abdomen, via injury and/or surgery, will give salience to this crucial muscle group. One can also increase awareness through ad hoc experiments, such as lifting something moderately heavy while either seated or standing, or by trying to raise one’s arms while a friend offers resistance via pushing down on one’s hands. Any lifting of the arms in musical performance, and all singing, involves activation of the abdominal muscles.[12] Mimetic comprehension of musical performance thus always involves abdominal exertions to some extent or another, and this contributes to musical experience by way of the correlations between abdominal exertions (tightness and relaxation) and emotional states.

Imitation and Music Perception

Our ability as infants to detect (perceive) differences in timbre, pitch, contour, and rhythm likely has an innate foundation (Trehub 2003). Although perception of such features involves nonmimetic processes, infants also imitate timbre, contour, and rhythm as early as six weeks of age (Malloch 1999–2000), and

19

one of the principles of the mimetic hypothesis is that imitation enhances perception. This proposition is supported by the findings of Phillips-Silver and Trainor (2005, involving children) and Phillips-Silver and Trainor (2007, involving adults), in which comprehension of heard rhythms was enhanced through rhythmic movement training. In plain terms, participants were better at comprehending heard actions (the sound of musical performance) that shared an exertion schema with actions that they had performed previously. These findings are also consistent with the proposition that rhythm perception involves real-time MMI, or mimetic enactment of the rhythms heard, based on previous overt imitation (MMA).

There is one notable perceptual ability, however, that must be reconciled with the role of mimetic comprehension: absolute pitch (AP). Its most distinctive features are the abilities to recognize the specific pitch of a note by ear and/or to sing a specific pitch without the aid of an instrument. Traditionally this is implicitly understood to involve nonmimetic processes, but if AP truly is nonmimetic, then this has implications for the overall relevance of the mimetic hypothesis, particularly since many music academics possess this ability. I explore this matter in appendix I.

1.2. Psychological Studies of Imitation: Social Interactions in Adulthood

Social interaction among humans and among other animals involves individuals comprehending and responding to the gestures of others. These include limb movements, gaits, postures, facial expressions, and nonlinguistic vocalizations, all of which contribute to “body language” or nonverbal communication. Among humans and some other species, comprehension of these gestures involves both mimetic and nonmimetic processes. In this section I describe how mimetic comprehension plays out in both functional and aesthetic human experiences, including comprehension of musical gestures.[13]

Because gestures signify something of the state and/or intention of the gesturer, let us refer to them as gestural signs. While comprehension of a gestural sign is sometimes conscious, more often it is either nonconscious or only marginally conscious. For example, the subtle gestural signs, or microexpressions (Ekman 2001), that signify attempted deceit may register explicitly in consciousness and provoke a corresponding response (such as a belief that the speaker is lying), or they may register only implicitly (nonconsciously) and provoke a subtler response (such as a feeling of doubt), or they may not be recognized at all and thus provoke no response (the deceit goes undetected). The issue here concerns the role of MMI in comprehending these and other gestural signs.

We can begin by acknowledging that imitation is not directly relevant in comprehending all gestures. For example, imitation is not needed for the comprehension of, and appropriate response to, aggressive gestures that threaten immediate bodily harm, as in the cross-species case of walking past a parked car and being surprised by a dog suddenly barking at me: my startled response is, I want to say,

entirely nonmimetic. Despite the significance of such cases, however, they are exceptional. The majority of gestural cues that we read in other people and animals are more subtle and nonthreatening, and comprehension of these cues regularly involves imitation. Chartrand and Bargh (1999) review the history of studies of social imitation and remind us of how easily, for example, one picks up a regional accent upon moving to a new state or country. In unintentionally adopting an accent, one adopts a way of behaving, via speech gestures, that results in becoming like the other members of the community. This extends even to apparently superfluous gestures in one-on-one interactions, as in the following example.

Chartrand and Bargh conducted an experiment in which participants were paired first with one confederate and then another, in a task that involved taking turns describing photographs. (“Confederates” are part of the team of experimenters and pose as participants.) One of the confederates actively performed gestures that were superfluous to the task of describing the photos (including smiling, face rubbing, and foot-shaking), while the other was more neutral, and the experimenters found that participants tended to imitate the superfluous gestures. Although irrelevant to the task, such imitation is not irrelevant to the larger social value of enacting mutuality while participating with others.

“Air Guitar” and Other Air Instruments

Overt mimetic musical behavior, such as singing along and/or playing along with recordings, is plainly evident in various contexts and for the present purpose does not require the support of clinical evidence. Instead, let us simply consider how it relates to the mimetic hypothesis.

Intentional (deliberate), overt imitation is common among musicians, whether in copying a teacher's demonstration, learning a song by ear, or transcribing a solo. Among nonmusicians, *air guitar* is a common form of recreational mimetic behavior: pretending to hold and play a guitar in imitation of a performer, whether in real time or recall. As Godøy, Haga, and Jensenius (2006) note, air guitar requires no expertise in actual guitar playing: the imitation does not need to be exact in every detail in order for someone to enjoy the reward of mimetic participation. Air guitar has a counterpart in the video games *Rock Band* and *Guitar Hero*. The imitation in these games originally involved button pushing, which facilitated participation by nonguitarists, but *Rock Band* subsequently developed a real guitar as a controller. One notable feature is that air guitarists tend to vocally imitate the guitar sounds, thereby adding cross-modal imitation to flesh out the experience. In the case of the video games, however, the gamers are not required to make any mimetic sounds since the sounds are produced by the game, and yet commonly they do sing along. This mimetic singing is superfluous as far as the game is concerned, but we can understand it as satisfying the urge to sing along and to thereby more fully participate in the music-making experience.

Guitar Hero also has a vocal version in which the player must match the singing in the recording. The aesthetic-competitive rewards of successful imitation are similar but occur within the more straightforward context of intramodal

vocal imitation. The genre of *karaoke* is akin to video games but without the point-oriented scoring, and its wide appeal can be understood similarly in connection with the reward of successful mimetic participation in the performance of favorite songs. (Both activities involve a combination of intramodal imitation of recalled singing and real-time cross-modal mimetic participation with the accompaniment.) At another level, according to the hypothesis, audience members in turn mimetically comprehend the amateur performer's singing and gain a reward that varies according to the quality of the performance and the expectations—for example, an out-of-tune performance can be found to be dreadful or delightful.

A related form of mimetic participation is the practice among preteens and teens of mimicking popular singers, including not only the singing but also the gestures, postures, and poses. Soon after YouTube was launched, people took the opportunity to share audiovisual recordings of their lip-synching mimes, sometimes with wildly popular results. Two examples on YouTube are a 2005 video by the "Chinese Backstreet Boys," Wei Wei and Huang Yixin, based on the song "I Want It That Way," and Gary Brolsma's 2006 "Numa Numa" video.[14] In a broad sense of the term, these are "covers" of the original songs, but we should ask why anyone would want to watch a lip-synched version of the original recording. One answer is that some viewers might find them humorous (more on this below), but in fact some viewers find the mimes' version more compelling, in a nonironic way, than the originals. One plausible reason for this is that the mimes are enacting what some people either enjoy enacting themselves (mimicking professionals) or might like to enact but do not feel comfortable doing themselves. In this situation we have first-order overt imitation (MMA) enacted by the mimes, and second-order covert imitation (MMI) enacted by viewers of the mimes' videos. In some cases, the second-order imitation enacted by viewers transforms into overt and deliberate imitation in the form of tribute videos, created by fans of the mime's videos and uploaded to YouTube.

At the same time, some other viewers derive pleasure in laughing at the lip-synching performers. Although this might seem to be a nonmimetic response, more likely it involves a combination of mimetic participation and aesthetic distance. If I give my attention to a music video, my comprehension will involve MMI, as part of my implicitly asking what it would be like to do what the performers are doing. At another level, I evaluate the embodied experience of what it would be like, which then results in either derision or admiration or indifference. Accordingly, the evaluation is based not merely on what I have seen and heard but on what I have imagined doing. In the concluding chapter I suggest that aesthetic evaluation of music in general is based in part on mimetic parti-cipation.

Old Man River and the Mimetic Imperative

More private forms of mimetic singing, whether covert (subvocalizing) or overt (singing aloud), also offer rewards of self-expression and mimetic participation, and the urge to overtly join in with a live public performance can sometimes

overpower normal social constraints. I was once at a recital by the bass-baritone Thomas Quasthoff, who sang "Old Man River" as one of his encores. As he was singing I thought I heard a kind of echo, and then I realized that an elderly gentleman two rows in front of me was singing along. His wife shushed him, but after a few bars he joined in again. It may be that this was one of his favorite tunes, and/or he may have taken the popular nature of the song as granting some license to join in, but it seems clear to me now that he was forced to choose between two incompatible social behaviors: to sit quietly like everyone else, or to give in to the mimetic urge. In this case the mimetic urge became for him an imperative. In my own experience at concerts and recitals, classical and otherwise, I regularly find myself singing along in my head and wanting to move with the music in some way or another, as I imagine many other listeners do. In subsequent chapters I describe how such mimetic engagement contributes not only to the immediate experience but also to our conceptualizations of musical experience generally.

Sentics: The Work of Manfred Clynes

Manfred Clynes (1977) measured overt physiological features of emotional states in connection with his theory of sentics. In a number of experiments, participants were asked to express one of several emotions (anger, joy, and so forth) while the middle finger of their dominant hand was attached to a device that measures changes in finger pressure. Changes in breathing and heart rate were also measured, as was direction of the finger pressing either away from or toward the body. One of these experiments concerned responses to music, in which participants were asked to recall one of various classical musical works. Clynes compares the finger movements to conducting, and he describes the pattern of movement as an expression of the music's "inner pulse." From the perspective of the mimetic hypothesis these are overt, deliberate mimetic responses that are congruent not only with the rate of exertions in the imagined performance (the finger, arm, and torso movements that would produce the imagined sounds), but also with other features of movement that are commonly indicated in performance instructions for this music: strength of effort (*piano* versus *forte*), speed and strength of onset (*legato* versus *marcato*), and pattern of exertion (*crescendo* versus *diminuendo*, *sostenuto* versus *staccato*, etc.). Given the classical repertoire, for some participants these finger movements may well have been influenced by images of conductors. However, conductors' exertions are nevertheless congruent with the sound-producing exertions of the performers and vice versa, so that imitation of them can be understood as indirect imitation of the performers. In effect the conductor says not only "Play at this tempo," but also "Play in a way that matches the manner of my exertions." A conductor's movements are an invitation to the audience (welcome or not) to feel the music in the way that the conductor is demonstrating.[15] As with other cross-modal mimetic responses, most of the finger movements measured by Clynes are not direct imitations of the particular exertions of the performers, but instead are cross-modal imitations of the pattern, rate, and intensity of the exertions.[16]

2. Neurological Studies of MMI and MMA in General

Evidence in this section comes primarily from measurements of the motor-related brain activity that occurs when participants observe the actions of others. Motor-related portions of the brain are those portions related to muscle action and movement (the musculoskeletal system). Activation of these portions of the brain while observing the actions of others suggests that part of how we comprehend observed actions is by imagining performing those actions. This imagining need not be deliberate or conscious. It is *imagery*, more specifically *motor* imagery, and because it underlies potential actions that would mimic the observed actions, it is *mimetic* motor imagery (MMI). The techniques for measuring brain activity in the studies cited here include fMRI (functional magnetic resonance imaging) and PET (positron emission tomography).

Mirror Neurons

Since overt mimetic behavior is a normal part of human behavior, there must be corresponding processes in the brain. The subsystem in which MMI occurs is sometimes referred to as the mirror system, since it involves mirroring observed behavior. MMI is also closely related to various versions of simulation theory (for example, Gallese and Goldman 1998, Gallese 2005; see Decety and Grèzes 2006 for a review), the most relevant feature of which is the relation between perception and action, with the principle that perception of the behavior of others activates simulation of the observed behavior in one's own motor imagery.

Among the relevant findings are those involving mirror neurons. As is now widely understood, these are neurons that fire both when an action is observed and when the same or closely analogous action is executed. The behavior of this category of neurons is more complex than is often acknowledged outside of neuroscience, including the classification of different types of mirror neurons according to function.[17] When it comes to understanding their role in specific musical contexts the complexities only multiply, and here I would like to identify what I believe are some of the challenges.

Consider the case of listening to, for example, a flute performance. We should ask whether the mirroring neurons are devoted to the fingers, the arms, the embouchure, the tongue, and/or the abdomen; or whether there are generic mirror neurons that are activated by any exertion regardless of the specific physical modality; or whether there might be groups, such as those involving the fingers and arms, or the lips and the tongue. We would also want to account for the significant variation in both the physical modality and the strength of mimetic responses among different listeners—for example, among flute players, among all other musicians, and among music lovers who do not perform on a regular basis. We should also investigate possible cross-modal mirroring, such as mimetic representations of instrumental music in voice-related areas of the brain. And we should distinguish between representations of pitch and representations of rhythm, timbre, and attack (tone onset) and intensity (strength). I believe that such specifications may be possible, but as yet these details remain unexplored.

MMI and Visual Perception

The majority of studies on the relation between perception and action focus on visual perception and not auditory perception. However, because much or most music is performed by the visible or visualizable exertions of human performers, these sources are relevant to the mimetic hypothesis.

In a paper focusing on involuntary simulation, Decety and Grèzes (2006) cite one study (Calvo-Merino et al. 2005) in which fMRI scans of expert dancers showed stronger activation of premotor cortex (a part of the brain involved in action planning) when watching other dancers, than in the brains of novices watching the same dancers.[18] One way to understand this is that this activation was stronger in expert dancers because of the greater congruency between the observed actions and their own experience. This process of understanding via simulation should apply beyond dance to expertise in other artistic, athletic, and everyday actions. For example, imagery activated in a two-year-old in watching someone tie a shoe or stir a cup of cocoa should be weaker, due to the lack of congruent experience, compared to that of a five-year-old or a twenty-five-year-old. In support of this conjecture, see Iacoboni et al. (2005) for evidence of MMI in adults when observing another person grasping a mug. With music, the same should apply with regard to differences of expertise in playing various instruments and in singing (principle 16).

Goal-Oriented Imitation of Actions and Sounds

Grèzes, Costes, and Decety (1998) had participants observe two kinds of hand-arm actions under two conditions. The kinds of actions were defined as meaningful (goal-directed actions upon objects, such as opening a bottle or sewing a button) and meaningless (similar gestures without objects), and the conditions were (1) simple observation and (2) observation with the goal of subsequently imitating the observed actions. Notably, the experimenters found activation of motor-related brain areas under the first condition, when the participants were simply observing both kinds of actions.[19] Such activation is a representation of the observed action in motor imagery and is, therefore, a form of MMI. The pertinent implication is that simply attending to goal-directed actions can activate MMI.

Perhaps not surprisingly, the same experimenters also found activation in motor-related brain areas under the second condition; this would reflect motor planning in connection with the goal of subsequently imitating what was observed, analogous to a music student observing a teacher's demonstration. In such a context there are then two kinds of goals: the goal-directed actions of the performer, and the observer's goal of creating a motor representation of the observed action and subsequently imitating it.

Along these lines, Wohlschläger, Gattis, and Bekkering (2003) describe a fundamental feature of imitation, which is that we tend to focus more on the goal of the action (in music, the sounds produced) than on the specific movements
25
involved (principle 8). They also describe a hierarchy of goal movements, or a nesting of concatenated actions, which in music would correspond to the production of individual pitches and rhythms combining to make a phrase, and the combinations of phrases that make larger structures.

In listening to music one can focus on the sounds produced and/or on the actions that produce them, but in the aesthetic context of attending a concert or watching an audio-visual recording, part of the pleasure is in observing the combination of, or the relationship between, artistic actions and artistic sounds. Broadly speaking, greater attention to the specific actions fosters intra-modal MMI, and greater attention to the sounds produced—the pitches, rhythms, timbres and so forth—fosters cross-modal MMI.

Layered on top of the goal-oriented actions of performers is the sense of motion toward "musical goals," such as the cadences and climaxes of tonal music. In such moments we can expect to find that mimetic engagement is particularly strong.

3.1. MMI and Auditory Perception: Speech

In this section I begin with perception of nonvocal sounds, and then consider speech perception on the premise that speaking and singing share fundamental processes with respect to production and perception, so that studies of speech perception should have some relevance for song perception.

Since human speech and song are the sounds and sights of humans performing very specific motor actions, this physicality is perceived along with the auditory and visual information. Speech perception involves both nonmimetic and mimetic processes, but I will focus primarily on the mimetic. The evidence in each case comes from both psychological and neurological studies.[20]

MMI in Auditory Imagery for Nonvocal Sounds

Underlying all mimetic comprehension is the correlation between sounds and the actions or events that produce them. When we hear incidental human-made sounds, such as door-closing or footsteps, we infer the corresponding actions without needing to see them performed. Although MMI is likely minimal in such cases, consider the findings of Gazzola, Aziz-Zadeh, and Keysers (2006), who report fMRI experiments showing activation of motor areas in human brains both when performing hand actions and when only hearing such actions.[21] The actions in question involved reaching for and grasping a peanut or a piece of paper, breaking or ripping the object, and replacing the object. The activation of these populations of neurons represents an equivalence between goal-oriented heard and performed hand actions. This implies that my comprehension of such sounds involves simulation of the actions (in MMI) that I infer are likely to have created the sound. In musical contexts this would include the goal-oriented (sound-producing) hand actions of musical performers: the sound of the piano, for example, is the sound of actions performed on a piano, the sound of hand drums

26

is the sound of drumming actions, and so forth. One of the implications, then, is that whenever we give our attention to such musical sounds, normally we do not simply hear the sounds, but we also feel something of what it would be like to perform the sound-producing actions. (That is, we mimetically represent the sound-producing hand actions, to some degree of fidelity, and such representations have an affective dimension, in what it feels like to perform such actions.)

MMI and Speech Perception

The central proposition here is that part of how we comprehend heard speech is by covertly and nonconsciously imitating the speaker. With this in mind, the first and most relevant thing to note about speech is that it is the sound of very specific muscle movements. As Ulrich Neisser describes it, to speak

> is to make finely controlled movements in certain parts of your body, with the result that information about these movements is broadcast to the environment. For this reason the movements of speech are sometimes called articulatory gestures. A person who perceives speech, then, is picking up information about a certain class of real, physical, tangible … events (1976, 156).[22]

The pertinent question then is how these events are perceived and comprehended, and I am suggesting that the process involves MMI. To help see how this might be, try to imagine how one learns to speak one's first language (or languages). Whatever else may be involved, normally this includes learning to imitate the vocal sounds of those around us. This imitation is simultaneously of the sounds and of the muscle movements that produce the sounds. For the visible portion of the relevant muscle movements the imitation is by eye, and for the invisible portion (the sounds) the imitation is by ear. With practice each of us eventually acquires our own repertoire of words and phrases, and imitation in the comprehension of heard speech seems gradually to disappear, much as it does in learning a new language as an adult: at first one feels the effort of trying to mimetically represent and reproduce the sounds heard, and this effort and attention gradually fade as one acquires fluency. The pertinent result is that the sense that one is imitating gradually disappears as one acquires fluency in a language. However, the studies of speech perception discussed below indicate that speech-related imitation only seems to disappear, and we can identify two reasons for why it fades from awareness.

One reason involves the acquisition of fluency, as just discussed. Another rea-son is that normally one's focus is on the words spoken and the things to which they refer—the goals of the speech actions. Poetry can bring this physicality closer to the surface, and Vladimir Nabokov draws explicit attention to this physicality in his alliterative description of the name of one of his title characters:

> Lolita, light of my life, fire of my loins. My sin, my soul. Lo-lee-ta: the tip of the tongue taking a trip of three steps down the palette to tap, at three, on the teeth. Lo. Lee. Ta.[23]

27

The available meaning of "Lolita" here is not confined to its reference to the character Lolita but also includes the feeling of speaking this name. To put it another way, although it may be possible to read this paragraph while paying little attention to the sensuality that Nabokov highlights, such a reading would be impoverished relative to a reading that included the salient feeling of making these tongue articulations. Nabokov explicitly invites us to understand the narrator's (Humbert's) predatory obsession via mimetic comprehension of Humbert's demonstration of this lingual sensuality.

Among the clinical evidence is a study by Wilson et al. (2004) who found that speech perception involves activation of motor areas that serve speech production. Fadiga et al. (2002) found that listening to both spoken words and spoken nonwords produced excitation of tongue muscles, with the actual words (the more linguistically meaningful sounds) producing the strongest response. They note that participants were not given the explicit goal of imitating the words and nonwords afterward, which means that this mimetic motor response was automatic. Nishitani and Hari (2002) found activation in motor areas not only when observing lip movements but even in observing still photos of lips. In a similar study, Watkins, Strafella, and Paus (2003) tested for lip-muscle activation under four conditions: (1) speech only (listening to continuous spoken prose while viewing visual noise); (2) nonverbal sounds only (listening to nonverbal sounds, such as glass breaking, bells ringing, and gunfire while viewing visual noise); (3) vision only: lips (viewing speech-related lip movements while listening to white noise); and (4) vision only: eyes (viewing eye and brow movements while listening to white noise). As predicted, conditions 1 and 3 (listening to speech, and viewing lips) produced the strongest responses.[24] Condition 4 (viewing eyes) produced no significant response, but condition 2 (listening to nonverbal sounds) produced an unexpectedly significant effect, for which the authors offer a possible explanation in terms of the imitability of these sounds. Most pertinent here, however, is the activation of lip muscles in listening to speech without seeing the speaker (condition 1) and the significance of goal-oriented actions.

Watkins et al. note that the finding of lip muscle activation is not consistent with that of Sundara, Namasivayam, and Chen (2001), who used a similar measuring technique; however, Fadiga, Craighero, and Oliver (2005) note that Sundara et al. measured activation of facial muscles that are not directly involved in speech production. If participants were particularly interested in such actions, which are superfluous to speech comprehension, we might find activation of mimetic responses in the corresponding muscles of the participant-observers; but in listening to speech, one's attention is on the sounds and one's nonconscious attention is on the relevant actions, including those of the lips, in which Watkins et al. in fact found mimetic activity. Watkins et al. also point out that Sundara et al. used isolated syllables (such as "ba" and "ta") as their stimuli as opposed to continuous spoken prose. With this in mind, we should consider further the significance of goal-oriented actions as methods of testing for mimetic comprehension of speech and music.

Isolated syllables are not a statistically common part of everyday experience, where syllables normally occur within whole words in more or less continuous prose. Although isolated syllables are of course imitable, without context there is little or no communicative goal and correspondingly little motivation to mimetically represent the sounds. (If in doubting this claim one were to respond with "Bah," this vocal reaction would be in a communicative context, as a dismissive response to a proposed idea, and would be performed with a prosody that is not commonly found in the uninflected stimuli used in speech perception studies). The situation with musical studies should be similar: we can expect to find MMI activated only minimally in response to isolated abstract sounds, such as a single interval presented with no musical context, whether in a laboratory or a classroom. The farther a stimulus is removed from normal contexts, the weaker the mimetic response is likely to be, whether in speech, music, or otherwise.

3.2. MMI and Auditory Perception: Vocal and Instrumental Music

Only a relatively small proportion of the many brain imaging studies involving music focuses on matters related to mimetic motor imagery.[25] Leman (2008) has applied some of the same evidence under consideration here in making related arguments from similar premises, particularly in his fourth and fifth chapters. Godøy's 2003 essay, "Motor-Mimetic Music Cognition," takes a closely related approach, and Molnar-Szakacs and Overy (2006) offer a sketch of how a mirror system might play a role in creating emotional responses to music. More broadly, Zatorre, Chen, and Penhune (2007) summarize findings related to auditory-motor interactions in music perception and production, including some consideration of apparently mimetic processes.

Imagery for Speech and Song: Mimetic Subvocalization

The auditory system proper and its corresponding portions of the brain are anatomically distinct from the musculoskeletal system, its corresponding portions of the brain, and mimetic processes. However, since auditory imagery includes representations of sounds that we have performed or might perform, it is related to motor imagery. A subset of such imagery includes planned, recalled, or otherwise imagined speech actions, or subvocalization, which is covert vocalization that includes the sound and feel of covert speech (as in silently practicing what one might say aloud), and the sound and feel of one's covert voice when reading. Subvocalization thus integrates auditory imagery and motor imagery, or imagined voice-related sounds and actions.

A special form of subvocalization is mimetic subvocalization, which includes covert imitation of someone else's spoken words and/or singing. Mimetic subvocalization is thus also a form of MMI, and the mimetic hypothesis holds that it extends from imitation of vocal sounds to imitation of musical sounds in general. The importance of music-related subvocalization has been discussed philosophically by, among others, Spencer (1951 [1857]), Barthes (1977), Lidov (1987),

and Cumming (1997, 2000), and yet it remains largely peripheral to our general understanding of how music works.

The significance of mimetic subvocalization derives from the significance of overt and covert vocalization, both linguistic and nonlinguistic. For most of us the voice is our first and daily means of communication via sound; it is normally integral to how we express ourselves and how we comprehend the vocal expressions of others. Unlike manufactured musical instruments, the voice is completely embodied within the flesh; we do not need to pick it up or move to its location before we can use it to make sounds. The voice is also capable of mimetically representing most if not all sounds to a greater degree of fidelity than most instruments, and it does not require special training; it gives us a way of mimetically comprehending music without having any direct experience playing the various musical instruments or singing like Maria Callas.

In some contexts subvocalization is taken to include low-level activation of the voice-related muscles (in the abdomen, throat, and/or mouth), and such activation would constitute a liminal form between MMI and MMA. Mimetic Subvocalization subsumes all forms of covert, liminal, and overt mimetic representations.

Evidence of Voice-Related Mimetic Subvocalization

Zattore et al. (1996) investigated auditory imagery for song, using “Jingle Bells,” “Battle Hymn of the Republic,” and “Joy to the World” as stimuli, and they found, perhaps not surprisingly, that perceiving a melody and imagining a melody involve overlapping neural systems. More notably, they also found activation of the supplementary motor area (SMA) in both tasks. Since the perceptual task and the imagery task both involved comparing the relative pitch height of two words in the lyrics, the authors note the implication that the participants were likely subvocalizing the composite of melody+lyrics. In the perceiving task this would be mimetic subvocalization in real time. In the imagining task this would involve recall and/or possibly planning (which is not necessarily mimetic). In a theory of melodic recall that focused solely on hearing, activation of the SMA would be superfluous. By contrast, it would be consistent with a theory that understood recall as a combination of rehearing and reenacting—in this case, resinging (and/or possibly replaying). Such reenactment would be mimetic, insofar as songs such as “Jingle Bells” are normally learned via imitation, so that such reenactment involves reenacting the original stage of learning via mimetic participation. I return to this issue in the discussion of principle 9 in the next chapter.[26]

Consistent with the study above, Hickok et al. (2003) had participants listen to nonsense speech and music, followed by voluntary covert rehearsal, and found activation of shared auditory and motor brain areas in both perception and rehearsal.[27] In a related fMRI study with similar procedures (passive perception of lyrics spoken and sung, and voluntary covert rehearsal of both while visually presented), Callan et al. (2006) found activation in the same and additional areas, and they interpret their data as being consistent with that of Hickok et al.[28]

30

Callan et al. focused specifically on the shared and unshared features of speech and music, as represented in the same lyrics being spoken and sung, and they describe their results as pointing toward a motor theory of music perception. In particular, they note that perception and imagery for spoken and sung lyrics involved activation of a brain area known to represent the lips and tongue, and which has been shown to correlate with improved pitch memory (Gaab et al. 2003).[29]

In order to exclude the potential influence of specifically linguistic processes, Halpern and Zatorre (1999) studied auditory imagery for familiar melodies without lyrics. The examples included two six-second portions of classical music (Tchaikovsky’s Nutcracker Suite and Beethoven’s Fifth Symphony) and themes from film, TV, and elsewhere; familiarity with the music was confirmed by a pilot study. There were three conditions involving intentional imagery for three kinds of stimuli: (1) listening to the beginning of a familiar melody and then imagining the continuation, (2) a control task of listening to a novel and unfamiliar melody, and (3) listening to and then reimagining the novel melody. For the familiar melodies, one might expect that imagery for the continuations should have been a matter of nonmotor retrieval; however, PET scans indicated activation of SMA in imagery for the continuation of familiar melodies as well as in reimagining novel melodies. They note that SMA activation likely reflects motor planning associated with subvocal singing or humming, and they conclude by speculating that "the SMA is specifically involved in a motor process relevant for auditory image generation, irrespective of the familiarity of the imagined stimulus" (704).

For some, evidence such as the foregoing only specifies some of the details of what is already plainly evident: perception and recall involve mimetic subvocalization. For others, this kind of evidence points to the often hidden nature of the performative component of perception and recall. This hidden nature can foster a belief that music perception and recall do not involve mimetic motor processes, but clinical studies suggest otherwise.

Mimetic Subvocalization and Instrumental Timbres

Mimetic subvocalization for song is a relatively straightforward case of intramodal MMI, but the next issue is whether subvocalization occurs cross-modally, in imagery for instrumental music. If it does, then to the extent that it does it makes vocal experience relevant to our comprehension of music generally. The significance of such relevance will become clearer in subsequent chapters. For now, let me first theorize and then consider some evidence.

Recall Brahms's First Symphony and the Beethovenian first theme of the finale; or recall some other folk-like instrumental melody such as the Largo of Dvorak's "New World" Symphony. While recalling the Brahms, some string players understandably report MMI related to string playing, while other string players also report a measure of subvocalization, in line with reports of most nonstring players. (These reports are in nonclinical surveys.) If some form

31

of subvocalization in recalling such instrumental music is common, we should ask why this would be: Why would recall of an instrumental melody not simply involve rehearing that melody as played by violins or the English horn? The evidence in the previous section might seem to offer an answer, but remember that those studies involved songs and, thus, intramodal mimetic subvocalization. In considering that some people's recall of the Brahms involves cross-modal subvocalization, some might object that the Brahms example is biased in that it involves a particularly singable melody. This is true; this and similar melodies afford easy vocal and subvocal imitation, and so we should ask whether and to what extent subvocalization is activated in hearing and recalling less easily singable instrumental music, including music that some might not think of as being singable at all.

Recall that learning to speak involves two layers of imitation: (1) imitation of the sounds produced (product, or end), and (2) imitation of the relevant muscle movements (process, or means). In speech and song these remain coupled, but with instrumental music these become uncoupled, in that we can recreate the same pitches and rhythms in the voice without needing to imitate the muscle movements of particular instrumental performance. This leads to the question of whether and to what extent we can represent instrumental timbres without MMI: Is it possible to hear or imagine the sound of, say, a bassoon without subvocally imitating the timbre of that sound? Some authors (e.g., Crowder 1989) have proposed that nonmotor representations are not only possible but are the norm, but let us consider some of the complications to this view.[30]

Every sound with a unique name has a unique timbre; the bassoon, oboe, piccolo, piano, and so forth are all distinguishable by their particular timbres. In testing for representations of timbre in imagery one immediately encounters the challenge that most musical instruments, in the way they are most often played, do not produce a timbre without simultaneously producing a more or less definite pitch, and this makes it difficult to separate timbre imagery from pitch imagery. One study that confronts this is Halpern et al. (2004), who used fMRI to test for timbre imagery. Participants made similarity judgments involving various common musical instruments (flute, trumpet, etc.) in a perceptual task (hearing) and in an imagery task (imagining). In the perceptual task, participants heard a single note played by one instrument for 1.5 seconds, followed by 2 seconds of silence, and then 1.5 seconds of a different instrument. They then judged the similarity of the timbres on a scale of one to five. The imagery task followed the same pattern, except that participants were only shown the names of the instruments, without hearing them, and were asked to imagine the sounds of the instruments. One premise of the study was that timbre imagery likely does not involve subvocalization (based in part on Crowder 1989), and since the task was a matter of comparing the timbre of heard and imagined sounds, the authors predicted no activation of SMA, since this would seem to indicate subvocalization (or perhaps some other form of motor imagery). They found, however, that SMA was in fact activated, and they offer two possible explanations. One is that, "although

subvocalizing the timbre of an instrument is difficult, the timbre was accompanied by pitch, which itself is easily vocalizable" (1291), implying that the intentional comparison of *timbres* was accompanied by unintentional subvocalization of each note's *pitch*. The other possible explanation that they offer is that SMA activation might involve some other nonsubvocal motor-related activity. Although they do not offer additional thoughts on this, it is plausible that this nonvocal motor imagery would be related to the mechanics of playing these instruments. As it happens, in following a model from an earlier study, seven of the eight examples used were wind instruments and the one exception was violin. Since preparation for the test ensured familiarity with the correlations of timbre, instrument name, and, likely, the appearance of the instruments themselves and human performance upon them (participants had at least five years of formal musical training, and the instruments were common orchestral instruments), each example also included implicit information about the finger, arm, and torso exertions as well as the lip, tongue, and chest exertions. Both of the suggested explanations here imply MMI: in the first, mimetic subvocalization of pitch; in the second, mimetic representation of the other relevant exertions.

I suspect that both forms of MMI occurred, but I also suspect that the subvocalization was not only for pitch but also for timbre. I base this conjecture on the observation that, despite the limitations of the human voice, we do intentionally and overtly imitate timbres, to some degree of fidelity, as children and as adults. The fact that the fidelity of our imitations is in most cases rough or very rough—I cannot really make the sound of a bassoon or an electric guitar with my voice—is, in an important sense, irrelevant. What matters is the attempt to emulate the sounds, to feel something of what it would be like to make such sounds, and to thereby feel something of what it would be like to be an entity capable of making such sounds. When air guitarists sing, they recreate not simply the pitches but also something of the timbre, and when conductors demonstrate by singing, they commonly model the timbre that they are asking for. The continuum of vocal imitability, from vocal music to instrumental music, electronic music, and environmental sounds then becomes one of the factors that shapes the experience of different kinds of music.

Instrument-Specific MMI

Let us now focus on the more straightforward matter of intramodal (nonsub-vocal) imitation of the bowings, blowings, fingerings, and beatings specific to individual instruments and instrument families; I consider electronic sounds in chapters 2, 8, 9, and 10. According to Haueisen and Knösche (2001), pianists reported involuntary finger movements when listening to well-performed piano music, which is conscious, unintentional, and intramodal MMA. We should not expect that the same would necessarily hold for nonpianists because of the difference in expertise (as in the dance study by Calvo-Merino et al. 2005), and this difference was borne out in the study.

Haueisen and Knösche also studied neural activity in pianists and nonpianists while listening to piano music, and indeed found

33

greater neural activity (in contralateral primary motor cortex) in pianists than in nonpianists. In a psychological study, Repp and Knoblich (2009) report related findings. Both sets of findings are consistent with the premise that experience strengthens MMI and MMA according to specific physical modalities. In another study, Repp and Knoblich (2004) tested the more specific case of pianists comparing recordings of their own performances to recordings of the same works performed by other pianists, and found that pianists were better at recognizing their own performances. Based on this, the authors hypothesize that an observer's motor system is most strongly activated when perceiving one's own actions. The strength of modally specific MMI for pianists, then, in general should be greatest in listening to their own performances, somewhat lesser in listening to those of other pianists, and then still lesser in listening to performances on other instruments. A similar continuum likely applies to each instrument and vocal type (soprano, alto, etc.), as the next two studies indicate.

Drost, Rieger, and Prinz (2007) had pianists and guitarists silently play a series of isolated major or minor chords while they simultaneously heard either a major or a minor chord that was either congruent (minor/minor, major/major) or incongruent with the chord they were about to play.[31] The chords that participants heard were presented as recordings in five timbres: piano, organ, guitar, flute, and voice. For pianists, significant interference effects occurred, but only in hearing the piano and organ chords (keyboard instruments) and not the others. Analogously for guitarists, significant interference effects occurred only with the guitar chords. The interference effect in both groups was manifest in the greater time it took to play the indicated chord. While it is likely that a portion of the interference involves nonmotor processes, the authors describe the heard chord's distracting effect as a "potential action effect," which can be also explained mimetically: the heard chord "invites" a mimetic performance, and when this conflicts with the chord to be actually played, negotiation of the conflict results in a delay. The fact that significant delays occurred only when the heard chords were in the timbre of the participants' instrument of expertise suggests, but does not necessarily indicate, that MMI was involved; however, given the other studies cited in this chapter, a mimetic approach offers a relatively straightforward explanation for at least part of the interference effect.

In a related fMRI study, involving expert violinists and flutists listening to musical works performed either on their instrument or on a different instrument (J. S. Bach partitas for solo violin and for solo flute), Margulis et al. (2009) found neurological activity consistent with the studies cited above: motor imagery was plainly activated when participants listened to music played on their own instrument, but only minimally activated when listening to the other instrument. We know from Münte, Altenmüller, and Jäncke (2002) that musical training shapes neurophysiology; however, that study compared musicians and nonmusicians. Margulis et al. were interested in whether the same thing might apply to specific kinds of musical training, such as flute playing versus violin playing. As they reason,

34

> If acoustic differences between the two sets of stimuli had been the primary relevant factor, results would have shown selective responses to violin and flute music, regardless of instrument of expertise. Instead, results show selective responses to music played on the instrument of expertise (violin for violinists and flute for flutists).… (271)

As they reason further, this selectivity of response (in syntax-related BA44, timbre-related auditory association cortex, motor-related precentral gyrus, and posture-related globus pallidus) indicates that structural and syntactic features alone cannot account for all of the neurological responses. The finding of activity in motor-related areas is particularly notable since the task involved passive listening—passive in that participants were not asked to perform any motor tasks or to generate any motor imagery, but were instead told only that they would be asked questions about the performance afterward.

Rhythm and MMI

Mimetic rhythmic engagement is perhaps the most obvious form of mimetic behavior, as in the common overt behaviors of toe-tapping, swaying, and dancing to music. A great deal of music—past and present, in the West and globally—invites such mimetic behaviors in offering listeners a regular beat, and these forms of overt mimetic participation hardly require clinical evidence. However, since it is common and perhaps sometimes even preferable to enjoy such music without engaging in overt mimetic behavior, we should ask whether it is possible to listen to such music without activation of MMI (covert mimetic participation). We should also ask a different question that brings us to the same issue: Is it possible to comprehend rhythm and meter in the absence of MMI? If the answer seems obvious, bear in mind that for some scholars the obvious answer is “yes” and for others it is “no.”

In the first of two fMRI experiments, Chen, Penhune, and Zatorre (2008) had participants listen to six-second examples played on a woodblock, with the instruction that afterward they would be asked to tap along with the same rhythms. During the initial passive but anticipatory listening phase, motor-related brain areas were activated, which is not surprising: this involves deliberately generating a mimetic motor plan.[32] The second experiment, however, is more telling. A separate group of volunteers listened to the same examples but were not told that they would be asked to tap along with a second run of the examples. Nevertheless, the same motor areas were activated in this condition as well (with less robust cerebellar activity corresponding to the different condition of not explicitly planning to imitate). This is MMI because it is a motor representation that mimics the pattern of the stimulus. Still more notably, this response occurred even though the stimuli were only minimally musical in any normative sense: six seconds of irregular (nonrepeating) patterns played without inflection on a wood block. With actual music, and with some kinds of music more so than others, we should expect to find even stronger activation of MMI.

35

Concluding Remarks

There is a great deal of additional supporting evidence that I have not considered, but I hope that the foregoing survey makes it plain enough that various forms of mimetic motor imagery (MMI) and mimetic motor action (MMA) appear to be integral to music perception, comprehension, and recall. I also hope that it is plain that mimetic comprehension of music appears to be a special case of mimetic comprehension generally; that it would be strange if we generally comprehended the actions of others via imitation and yet did not do so in the case of musical actions. Finally, I hope that the suggestion is clear that much or most of mimetic comprehension occurs without our awareness, some of which is available to awareness but is ignored, and some of which is unavailable to awareness. The next task is to specify more precisely how these mimetic processes actually play out in music.

Notes

1. I am avoiding the noun form of *mimesis* because, although Aristotle's writings on the matter include audience members' comprehension of performance via mimesis, the predominant interpretation today is, roughly speaking, in terms of art imitating life, which is the opposite of what I intend here. See Max Paddison (2010) for a potent discussion of the various meanings of the term and its role in the construction of musical meaning. The hypothesis and my exploration of its implications overlap with Merlin Donald's account of the role of imitation and mimesis in the evolution of human cognition and culture (Donald 1991), except that he defines *mimetic* more strictly than I do here.

2. There is a definition of *nonconscious* that applies only to processes that are never available to consciousness; however, with mimetic processes this appears to vary among individuals, as well as for a given individual across one's life. In addition, unconscious bears connotations that may confuse my intention, so I am using this broader definition of nonconscious.

3. A partial list includes: Spencer (1951 [1857]), Clarke (1993), Iyer (2002), Godøy (2003), Godøy et al. (2006), Jackendoff and Lerdahl (2006), Molnar-Szakacs and Overy (2006), Phillips-Silver, Aktipis, and Bryant (2010), and Toiviainen, Luck, and Thompson (2010). More broadly, it also overlaps with the unpublished writings on aesthetics of the philosopher Adam Smith (Malek 1972), and with ideas from James (1890), Barthes (1977), Walton (1990, 1993, and 1994), and Dissanayake (1992).

4. Cusick's essay originally appeared in 1994 in the first edition of *Queering the Pitch*. While the notion of invited participation occurred to me prior to reading Cusick's essay, the related portions of the mimetic hypothesis and its implications are in effect an exploration of her idea.

5. Among the extensive literature, helpful collections include Meltzoff and Moore (1983), Papoušek et al. (1992), Nadel and Butterworth (1999), Meltzoff and Prinz (2002), and Hurley and Chater (2005), but among the most relevant for the mimetic hypothesis and its implications are Trevarthen, Kokkinaki, and Fiamenghi, Jr., (1999) and Meltzoff (2002).

6. For a philosophical application of infant studies in a book-length essay, see Gallagher (2005), who offers a theory of the how the adult mind is shaped by experience, including mimetic experience. For a different interpretation of the data on infant imitation, see Heyes (2001).

7. See Gallese and Goldman (1998) for a theory of how this relates to inferring the intentions of others in adult-adult interactions.

8. Meltzoff and Moore found that imitative behavior for newborns can occur after a delay of twenty-four hours. If a mimetic response is delayed, then there must be some form of representation of the modeled behavior that makes this delayed imitation possible. I argue below that this representation includes mimetic motor imagery, which in adults persists indefinitely.

9. Exertions schemas are analogs of the *image schemas* of conceptual metaphor theory, which we will consider in chapter 3.

10. We could refer to this as *entrainment*, just as when we move in time to music as adults (London 2012; Phillips-Silver, Aktipis, and Bryant 2010), but entrainment is perhaps best understood as a special case of mimetic engagement, as will become clear by the end of chapter 2.

11. One could also think of abdominal exertions as modally multivalent. The term *supramodal* might also be adapted for this purpose.

12. Although abdominal exertions anchor limb and vocal exertions, in exercises such as I have described you might also notice exertions in your leg muscles, revealing the role of the abdomen to be more of a fulcrum.

13. See Hatten (2004) for an exploration of musical gestures and their role in musical semiotics. In particular, Hatten's chapter 5 overlaps with the present discussion.

14. The song lip-synched by Brolsma is *Dragostea din tei*, performed by the Moldovan pop group O-Zone. A search for "Numa Numa" and "Chinese Backstreet Boys" should immediately produce links to these videos. The number of views does not compare to current popular videos, but the numbers are proportionally comparable when adjusted for the overall number of visitors in 2005 when YouTube was new.

15. See Luck, Toiviainen, and Thompson (2010) on the perception of conductors' expression. Per the dominant tradition, perception in this study is taken to be nonmimetic.

16. It is possible that these finger movements approach imitation of the finger movements of the pianists in those examples involving piano repertoire, and this might extend to the fingerings of the violinists in other repertoire. As always, top-down forces likely contribute as well, such as a notion of Mozart's music as "refined" influencing both the recorded performances and the intensity and "shape" of participants' motor responses.

17. The specialization of neurons within the category of mirror neurons was recognized as early as Keysers et al. 2003. See also Gallese and Lakoff 2005.

18. Specific brain areas are conventionally referred to without the definite article; thus, "activation of premotor cortex" instead of "activation of the premotor cortex."

19. Right cerebellum and bilateral activation in the dorsal pathway reaching premotor cortex. Meaningful actions also elicited bilateral activations in the supplementary motor area and in orbitofrontal cortex.

20. There is an overlap here with the motor theory of speech perception, but I am making only the claims discussed here and basing these claims on the evidence and arguments provided here.
21. Bilateral mid-temporal gyrus (MTG), left inferior parietal lobule, and left premotor cortex (BA 44/6). This is consistent with early findings in macaques in similar contexts. I should note that the strength of data based on human brain imaging has been called into question by Kriegeskorte et al. (2009), owing to an apparently common failure in statistical analysis of the data; however, Kriegeskorte et al. explain that they have no way of assessing the severity or significance of the distortions. This is relevant for some details of the mimetic hypothesis but not for its basic principles.
22. This passage is cited in Armstrong, Stokoe, and Wilcox (1995, 43).
23. Vladimir Nabokov, *Lolita*, from the beginning of chapter 1. Nabokov's elaboration on the experience of pronouncing this name extends to include the saturation of the entire opening paragraph with the phonemes /t/ and /l/.
24. The effect was lateralized to the left hemisphere; stimulation of the right hemisphere did not produce comparable results.
25. For a discussion of music-related brain organization beyond the focus here, see Peretz and Zatorre (2005).
26. The central implication for the development of musical imagery in musicianship courses (aural skills, ear training) is relatively straightforward: musical imagery combines not only auditory and visual imagery, but also motor imagery. The extent to which the motor component is mimetic is a finer detail.
27. Premotor cortex, superior temporal sulcus, and the parietal-temporal boundary within the Sylvian fissure.
28. The areas of activation in Callan et al.: "Left planum temporale/superior temporal parietal region, as well as left and right premotor cortex, lateral aspect of the VI lobule of posterior cerebellum, anterior superior temporal gyrus, and planum polare" (1327).
29. The VI lobule of the posterior cerebellum, in which they found bilateral activation.
30. In his conclusion, Crowder writes that "a strictly motoric representational mode (such as singing to oneself) may be ruled out. Humans are utterly incapable of reproducing physically any but the grossest dynamic or spectral features of timbre" (478).
31. The prompt for the imperative task was staff notation for pianists and tablature for guitarists, each projected onto a screen for three seconds. Pianists played a silenced electronic keyboard; guitarists were instructed to finger but not play the chords.
32. The areas in question were SMA, mid-premotor cortex, and cerebellum lobule VI. Ventral premotor cortex was active only during the conditions of the first experiment: action-coupled perception (anticipation of action) and action.