

Representation and Execution of Vocal Motor Programs for Expert Singing of Tonal Melodies

Eileen L. Zurbriggen
University of California, Santa Cruz

Dwight L. Fontenot and David E. Meyer
University of Michigan

Three experiments were conducted to study motor programs used by expert singers to produce short tonal melodies. Each experiment involved a response-priming procedure in which singers prepared to sing a primary melody but on 50% of trials had to switch and sing a different (secondary) melody instead. In Experiment 1, secondary melodies in the same key as the primary melody were easier to produce than secondary melodies in a different key. Experiment 2 showed that it was the initial note rather than key per se that affected production of secondary melodies. In Experiment 3, secondary melodies involving exact transpositions were easier to sing than secondary melodies with a different contour than the primary melody. Also, switches between the keys of C and G were easier than those between C and E. Taken together, these results suggest that the initial note of a melody may be the most important element in the motor program, that key is represented in a hierarchical form, and that melodic contour is represented as a series of exact semitone offsets.

Keywords: singing, music production, motor programs, response-priming paradigm, musical key and contour

One of the important goals of human performance research has been to understand the psychological mechanisms that underlie and control physical movements (Rosenbaum, 1991). Researchers interested in motor control have most commonly studied the manual, ocular, and vocal movement systems (e.g., see Abrams, Meyer, & Kornblum, 1990; Gordon & Meyer, 1987; Rosenbaum, 1980; Rosenbaum, Kenny, & Derr, 1983; Sternberg, Monsell, Knoll, & Wright, 1978). With respect to vocal motor control, nearly all studies have focused on the production of verbal units such as words, syllables, or phrases (e.g., Dell, 1986; Gordon & Meyer, 1987; Rosenbaum, Gordon, Stillings, & Feinstein, 1987;

Sternberg et al., 1978). These studies have advanced understanding of both language production specifically and vocal motor control more generally. Yet this focus on language production has bypassed an entire realm of performance that is psychologically important and culturally universal. That realm is singing.

There have been virtually no studies of the motor programs that enable singing in humans, even though such research would be interesting for a number of reasons. Singing is a commonplace activity and also one in which performance can reach high levels of virtuosity. Unlike language production, in which an adequate level of competence is achieved by nearly everyone, there are large individual differences in singing ability. Singing provides researchers with an opportunity to study both novice and expert performance as well as the developmental and learning processes whereby expertise is gained. Moreover, whereas language perception and production are mainly lateralized in the left hemisphere (Lenneberg, 1967), music perception and production also involve significant right hemisphere activation (Cadalbert, Landis, Regard, & Graves, 1994; O'Boyle & Sanford, 1988; Patel, Gibson, Ratner, Besson, & Holcomb, 1998; Tramo & Bharucha, 1991; Zatorre, 1984). Thus, important complementary insights about the neurological basis of motor control may potentially be discovered through the study of vocal processes (like singing) that are in some ways similar to language production but are also distinct neurologically.

By studying singing and other forms of musical performance, researchers can also advance and refine theories in the psychology of music that were originally developed through the study of music perception (e.g., Deutsch, 1969; Dowling, 1978; Krumhansl, 1990; Krumhansl & Kessler, 1982). In particular, it is important to determine which aspects of these theories generalize from perception to production. Studies of singing may likewise expand the scope of previous music production research, which has focused mostly on instrumentalists (especially pianists) and on rhythm and

Eileen L. Zurbriggen, Department of Psychology, University of California, Santa Cruz; Dwight L. Fontenot and David E. Meyer, Department of Psychology, University of Michigan.

This article is based in part on a doctoral dissertation submitted by Dwight L. Fontenot and a master's thesis submitted by Eileen L. Zurbriggen to the Rackham School of Graduate Studies at the University of Michigan. Portions of this research were presented at the Second International Conference on Music Perception and Cognition, Los Angeles, California, February 1992, and at the meeting of the Society for Music Perception and Cognition, Philadelphia, Pennsylvania, June 1993.

We thank Stuart Klapp and Ralph Krampe for their valuable suggestions, comments, and criticisms concerning an earlier version of this article. In addition, we thank David Rosenbaum for thoughtful feedback. Janine Anderson, Elizabeth Morgan, and Megan Yost assisted in creating the figures, and Peter Bovair and Chris Coon provided computer programming support; we gratefully acknowledge their help. We are especially indebted to Richmond H. Browne, Sylvan Kornblum, and William C. Stebbins, who served on Dwight L. Fontenot's doctoral dissertation committee.

Correspondence concerning this article should be addressed to Eileen L. Zurbriggen, Department of Psychology, Room 277 Social Sciences 2, University of California, Santa Cruz, Santa Cruz, CA 95064. E-mail: zurbrigg@ucsc.edu

timing (e.g., Palmer, 1989; Todd, 1985; see Palmer, 1997, for a review).

The present experiments investigated the nature, structure, and content of the motor programs that mediate skilled singing of short tonal melodies. The literatures on both music perception and music production have guided our theorizing about which features or attributes of melodies are likely to be represented as elements in motor programs. Because we used a response-priming paradigm originally developed to study speech production and manual movements, we also review the literatures relevant to them.

Music Perception and Production

Much of the psychological research on music has dealt with the perception of simple and complex tones and with the ability of participants—both experts and novices—to recognize certain aspects of melodies (for reviews, see Krumhansl, 1991, 2000). This research has led to a number of theories about melody perception and recognition that may or may not generalize to melody production. The most important theories include one proposed by Krumhansl and others about hierarchical pitch perception and one proposed by Dowling about the roles of musical key and contour in music perception.

Krumhansl and colleagues (Krumhansl, 1979, 1990; Krumhansl & Kessler, 1982; Krumhansl & Shepard, 1979) have demonstrated that each musical key is associated with its own tonal hierarchy (see Figure 1). The tonic note (“do”) is consistently rated as “fitting most closely with” a particular key. The tones that are part of the major triad (“mi” and “so” [for the key of C major, the notes E and G]) receive the next highest ratings, followed by the other remaining notes of the scale (“re,” “fa,” “la,” and “ti” [for the key of C major, the notes D, F, A, and B]). The tones judged to fit least well within the tonal hierarchy are those not among the seven notes of the scale (for the key of C major, the notes C#/D♭, D#/E♭, F#/G♭, G#/A♭, and A#/B♭). Thus, for each key, there is a four-tiered hierarchy, with the tonic at the first level, the other two notes of the major triad in the second level, the remaining four notes of the scale in the third level, and the five notes that are outside the scale in the last level.

Krumhansl (1985) also compared her hypothesized tonal hierarchy to existing analyses of the frequency and duration of notes in tonal compositions by Schubert, Mozart, and other Western composers. The same four-tiered structure was found in these compo-

sitions. Regardless of the key in which a composition was scored, the tonic was included more frequently and for longer durations, followed by the other two notes of the major triad, and then the remaining four notes of the scale. The five notes that were not a part of the scale were typically used only as ornamentation and were thus included only infrequently and had brief durations.

A second influential theory is one proposed by Dowling (1978) regarding musical memory. According to this theory, two components of melodies are important in musical recognition—melodic contour (the pattern of upward and downward steps between notes in a musical passage) and musical scale or key. Dowling postulated the use of two schemata—one for the melodic contour and one for the underlying key. This theory has been supported by studies demonstrating that contour can be remembered separately from pitch (Dowling, 1978; Dowling & Fujitani, 1971) and that familiar melodies can be recognized even when the exact pattern of intervals has been modified as long as the basic shape of the contour is preserved (Dowling & Fujitani, 1971).

On the basis of these results, we hypothesized that musical key and contour may provide two main organizing principles of motor programs for melodic singing. Accordingly, our experiments were designed to determine the role and importance of key and contour in singers’ motor programs. To appreciate this design, further consideration of previous research on motor programming for manual movements and speech is warranted.

Motor Programming for Manual Movements and Speech

Some especially relevant empirical and theoretical research focuses on manual keypressing and reaching. For example, using a movement-precuing procedure, Rosenbaum (1980) found that manual reaching movements may be programmed in terms of distinct features corresponding to the effector, direction, and extent of each movement. Under some circumstances, it is conceivable that such features are organized hierarchically within the movements’ motor programs. Consistent with this possibility, sequences of keypresses that alternate systematically between homologous and nonhomologous fingers of the two hands have response latencies whose pattern supports the hypothesis of a “tree-traversal process” in manual movement production (Rosenbaum et al., 1983). Some evidence suggests that such a process could underlie speech production as well (Rosenbaum, Weber, Hazelett, & Hindorff, 1986).

In our three experiments, we used an adaptation of a response-priming procedure developed by Gordon and Meyer (1987; see also Meyer & Gordon, 1985; Yaniv, Meyer, Gordon, Huff, & Sevald, 1990). This procedure requires participants to make a rapid and accurate shift from anticipating a prepared (primary) response to producing a new but related (secondary) response. Because participants attempt to produce secondary responses as rapidly as possible, it is assumed that in doing so they do not simply discard their original motor program for the primary response. Rather, they presumably modify the primary response program so as to construct a program for the secondary response. By retaining as much of the original primary motor program as possible, participants perhaps can minimize the time needed to construct the commands that will enable them to execute the new secondary response. Thus, by varying the relationship between primary melodies and second-

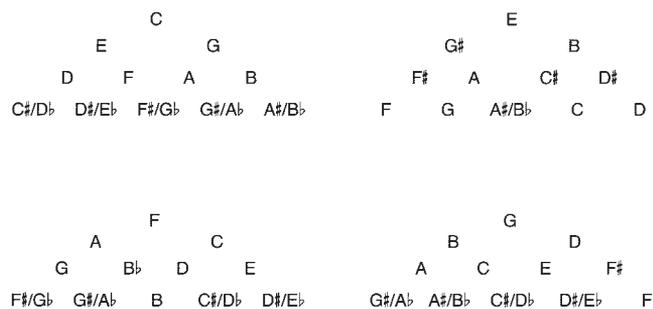


Figure 1. Hypothesized hierarchical representations (see Krumhansl, 1990) for the keys of C (upper left), E (upper right), F (lower left), and G (lower right).

ary melodies, one may deduce the form of the representation used in the motor program.

For example, Gordon and Meyer (1987) presented participants with a primary syllable sequence and a hierarchically congruent or hierarchically incongruent secondary sequence of four syllables. One such primary sequence contained the four syllables “Bee, Bay, Bah, Boo.” It was paired with either the hierarchically congruent secondary sequence “Bah, Boo, Bee, Bay” or the hierarchically incongruent sequence “Bah, Bay, Boo, Bee.” Participants produced the “Bah, Boo, Bee, Bay” secondary sequence more quickly, supporting the hypothesis that the motor program for producing the primary sequence was hierarchically rather than linearly organized.

Representation and Execution of Motor Programs for Singing

Because virtually no research on the motor programs for singing has been conducted, very few logically possible ways of representing and executing them can be ruled out at this time. We therefore provide a brief discussion of the many possibilities here. In the General Discussion, we provide a more complete outline of our tentative preferred theoretical model.

Issues of Representation

There are several possibilities for the form in which the motor program to produce a melody might be represented in memory. Such melodies might be represented in an acoustic form (corresponding to the actual sounds made when the singer produces the notes) or even in a visual form (e.g., the notes as written in a musical score). However, because the task of singing is to accurately produce (rather than perceive) the notes, both of these types of representation seem relatively implausible. More likely are either articulatory (i.e., commands to move tongue, lips, diaphragm, and larynx) or abstract representations. If the representations are abstract, they might take one of several forms, such as exact frequencies, exact frequency offsets, or semitone offsets. A rough contour representation (cf. Dowling & Fujitani, 1971) is also possible. Although some of these possibilities seem more plausible than others, to our knowledge none have been tested experimentally for expert singing.

Representations Based on Exact Frequencies

Notes may be stored in memory as exact frequencies. For example, in the melody shown in Figure 2, the exact frequencies are F4, A4, D4, G4, and E4, or 349 Hz, 440 Hz, 294 Hz, 392 Hz, and 330 Hz. Alternatively, notes may be represented as exact frequency offsets from some other note (e.g., the initial note of the melody, the previous note in the melody, or the tonic of the key in



Figure 2. Sample five-note tonal melody: F4, A4, D4, G4, E4.

which the melody is scored). For the melody in Figure 2, exact frequency offsets from the initial note would be represented as 0 Hz, +91 Hz, -55 Hz, +43 Hz, and -19 Hz.

Although representations based on exact frequencies are logically possible, they are implausible for several reasons. First, based on such representations, all individual frequencies and sequences of frequencies would be equally easy to sing. In fact, however, Western singers have a very difficult time singing quarter-tone intervals or other sequences of frequencies that are outside of the Western octave. Second, a representation based on exact frequencies also implies that absolute or “perfect” pitch (the ability to identify or produce a note on the musical scale without any reference note) should be fairly common. However, this is a rare ability among adults (although perhaps not among infants; Saffran & Griepentrog, 2001; see also Deutsch, 2002). Third, a representation based on exact frequencies implies that it would be equally easy to switch between any two notes or melodies, but the subjective experience of singers contradicts this, as do data from the present experiments.

Representations Based on Semitone Offsets

Another possible set of representations for motor programs in singing involves semitone offsets. Because the semitone is the smallest pitch interval used in Western music and is the foundation of the Western musical scale,¹ this possibility seems psychologically plausible. As with exact frequency offsets, semitone offsets may be based on several alternative standards of comparison (i.e., the notes used to compute the offset). Each semitone offset might be computed from the initial note of the melody (e.g., for the melody in Figure 2, the offsets would be 0, +4, -3, +2, and -1). Alternatively, each semitone offset might be computed from the immediately preceding note (e.g., for that same melody, the offsets would be 0, +4, -7, +5, and -3). Or the semitone offsets might be computed from the tonic of the key in which the melody is scored (e.g., if one assumes middle C or C4 as the reference note, the offsets would be +5, +9, +2, +7, and +4).

In each of these hypothetical representations, the direction of pitch change (+ or - [to a higher pitch or to a lower pitch]) and magnitude of change (e.g., four semitones) are stored together. However, the direction and magnitude might instead be stored separately (e.g., for offsets from middle C, +, +, +, +, and + may be stored in one buffer and 5, 9, 2, 7, and 4 in another). This type of representation would enable a “rough” representation of contour, as in Dowling’s (1978) theory of music perception.

Assumptions About Execution

A complete theory of how expert singers produce short tonal melodies also should include an explicit account of the processes whereby melody representations are constructed, modified, and executed. Our use of the response-priming paradigm requires two assumptions. First, we assume that before each trial begins, the singer has fully prepared the motor program for the primary melody, and it is ready to be executed. Second, we assume that if

¹ In Western music, there are two semitones embedded within the major key: the interval between “mi” and “fa” (E and F in the C-major key) and the interval between “ti” and “do” (B and C in the C-major key).

the singer has to switch and sing a secondary melody, then she or he will modify the already-prepared motor program as efficiently as possible. In other words, insofar as possible, the singer will reuse elements of the motor program for the primary melody that is already in memory. Thus, the aspects of execution that are most relevant under our experimental paradigm involve checking for differences between the primary and secondary melodies and modifying the contents of storage buffers to accommodate these differences.

There are a number of questions related to storage modification that a complete theory of singing would address. For example, are elements of the storage buffer overwritten or merely adjusted? Are elements of the storage buffer modified simultaneously or sequentially? Can execution of some elements begin while modification of other elements is still proceeding? Obviously, some mechanism must be available for overlapping the construction and execution of the motor programs in melody production. Otherwise, an entire aria would have to be fully prepared before the first note is sung. However, for short melodies such as those used in our experiments, a strategy of full preparation before execution might be preferred. Although our data do not directly answer all of these questions, we mention them for the sake of completeness, and the present experiments do yield evidence that favors some possibilities over others.

Another aspect of storage modification concerns the process whereby a new note is accessed. One possibility is that every note may be equally accessible (e.g., perhaps it is as easy to replace C4 with A4 as it is to replace C4 with C3). A second possibility is that the new note may be more or less accessible on the basis of its physical distance from some standard (e.g., the tonic, the initial note of the melody or the note that is being replaced). This would fit well with a model of storage modification that involves analog adjustments. That is, if C4 is replaced with A4 (a difference of +9 semitones) by (metaphorically or physically) sliding up through all the intervening pitches, then we would expect that this switch might take longer than, for example, a switch from C4 to D4 (a difference of +2 semitones). A third possibility is that the new note may be more or less accessible on the basis of its function in the musical key of the primary melody. Data from our three experiments provide evidence regarding the veracity of these three theoretical possibilities.

Overview of Experiments

A modified version of the response-priming procedure developed by Gordon and Meyer (1987) was used in our three experiments. Expert singers were trained to prepare a particular primary response melody and then to produce it as quickly as possible after a response signal. On certain trials, a secondary response melody was required instead. The participants were directed to always be ready to produce the primary melody, but when the secondary melody was required, they had to switch and produce it as quickly as possible. We assumed that participants had prepared and "loaded" a motor program for the primary response. The quickest way to switch to a secondary response melody would be to use parts of the motor program for the primary response melody rather than constructing an entirely new motor program from scratch. Differences in speed and accuracy of production in different con-

ditions could thus allow us to make inferences about the structure of the motor program for the primary response melody.

Only highly trained singers who were also expert at reading music and at sight singing served as participants. Our selection of expert singers was necessitated by the complexity of the required behavior (producing a series of five sequential, speeded responses) and the stringency of the performance criteria (near-perfect production of each pitch). Even so, the task was such that we did not expect flawless performance.²

From the music perception literature (especially the work of Dowling, 1978), we hypothesized that key and contour are likely to be salient aspects of the primary response melody. Requiring either of them to be changed to produce the secondary response melody should therefore yield relatively long secondary response latencies. Experiment 1 was designed to test this hypothesis. In this experiment, participants sang secondary melodies that had a similar or very different contour compared with the primary; these melodies also were scored in the same key as or a different key than the primary. All melodies in Experiment 1 began with the tonic; thus, a secondary melody in a different key than the primary also had a different initial note than did the primary. Experiment 2 was designed to unconfound these two factors (i.e., differences in initial notes and in keys). In Experiment 3, the difference between singing an exact transposition and switching to a different contour was explored. Also, hypotheses about how singers represent initial notes in motor programs for singing were tested.

Experiment 1

The purpose of Experiment 1 was to initially investigate the nature of the motor programs used for melodic singing. In particular, this experiment was designed to test whether musical key and "rough" melodic contour constitute important aspects of the motor programs for singing short tonal melodies. Secondary response melodies differed from primary response melodies along two dimensions. They were scored in either the same or different keys, and their contours were either very similar or quite different.

Method

Participants

One male and 3 female participants were recruited from graduate students at the University of Michigan School of Music. All were studying either vocal performance or vocal music education. They had 14.5 years of music education on average, and they considered themselves excellent sight readers and singers. This high level of expertise was necessary because of the difficulty of the experimental task. One female singer reported having absolute pitch. However, on the basis of her subsequent performance in our experiments, this appeared to be very nearly, but not strictly, true.

² In typical reaction time tasks, one expects error rates of 5%–10% per basic response. However, in our experiments, a 10% error rate for each individual note could result in an overall error rate as high as 40% (i.e., if one assumes a .90 probability of a correct response for each of the five notes of a melody, and if one assumes statistical independence of these five events, this would result in .90⁵ or 60% completely correct response melodies).

Participants were tested individually under laboratory conditions. Each was paid \$10 per hour and also received a bonus based on quality of performance. Ample time was provided to learn the procedures and practice with the apparatus.

Apparatus

An AST Premium 386 personal digital computer was used to control data collection. Visual messages, printed melodies, and response feedback were presented to participants on a Zenith (Lincolnshire, IL) ZVM-1200 monochrome display monitor. Warning signals and response signals were presented over Sennheiser (Old Lyme, CT) HMD 224 headphones with an attached microphone. Participants' vocal responses triggered an MED Associates (St. Albans, VT) voice-activated switch (ANL-923), which signaled the computer that a vocal response had occurred. A software voicekey (Meyer & Gordon, 1985) was used to measure response latencies and durations and to store the digitized vocal responses in computer memory. All vocal responses were recorded on both disk and audiotape.

Response Melodies

The melodies used in Experiment 1 appear in Figure 3. Eight 5-note melodies were constructed: four in the key of C and four identical melodies in the key of F. Two of the C melodies and two of the F melodies had a strictly rising contour. The other four melodies had a "bumpy" contour with two direction reversals. To eliminate timing artifacts, each melody had the same tempo markings and included only quarter notes. Participants were instructed to sing each note in the melody with an equal duration and accent.

Each of the melodies in Figure 3 served as a primary response melody. For each primary response melody, any one of four other melodies from Figure 3 served as the accompanying secondary response melody. For example, for Rising Melody 2 in the key of C (C4, E4, F4, B4, C5), the following four melodies could serve as the accompanying secondary response melody: Rising Melody 1 in the key of C (C4, E4, G4, B4, C5), Bumpy Melody 2 in the key of C (C4, F4, B3, F4, E4), Rising Melody 1 in the key of F (F4, A4, C5, E5, F5), and Bumpy Melody 2 in the key of F (F4, B4, E4, B4, A4). The associated secondary response melodies differed from the primary melodies along two dimensions: key and contour. They were scored either in the identical key (IK) as the primary response melody or in a different key (DK), and they had either a similar contour (SC; i.e., both melodies were bumpy or both were strictly rising) or a different contour (DC; i.e., one melody was bumpy and one was strictly rising). The design was thus a 2×2 within-subject factorial design,



Figure 3. Melodies used in Experiment 1. R₁C = Rising Melody 1, key of C; R₂C = Rising Melody 2, key of C; B₁C = Bumpy Melody 1, key of C; B₂C = Bumpy Melody 2, key of C; R₁F = Rising Melody 1, key of F; R₂F = Rising Melody 2, key of F; B₁F = Bumpy Melody 1, key of F; B₂F = Bumpy Melody 2, key of F.

with the four cells being IKSC, IKDC, DKSC, and DKDC. All four of these primary–secondary melody relations occurred equally often throughout the experiment.

Procedure

Each participant was tested individually during a series of test trials. On each trial, there was an initial practice phase followed by a test phase. During the practice phase, the participant was introduced to a primary melody, which was presented in a musical staff on the computer display screen. He or she attempted to sing the primary melody for practice, and the experimenter provided feedback about the accuracy of the response.³ If it was incorrect, then the correct response was played over the participant's headphones, and he or she attempted again to sing the primary melody accurately. Throughout the practice phase, the visual image of the primary melody was continuously present on the computer display monitor, and auditory presentation of it was repeated as many times as necessary to enable correct performance. When the experimenter and participant agreed that the primary melody was being sung properly, the next test phase of the trial began.

At the start of the test phase, the computer display screen was cleared. Next, a series of three warning signals (400-Hz tones, each lasting 100 ms) was presented via the participant's headphones. These were followed by the simultaneous presentation of auditory and visual response signals. On trials that required participants to make the primary response, a high-pitched (500-Hz tone) response signal was presented along with a blank musical staff on the display screen. This cued the participant to quickly and accurately sing the primary melody, which had already been rehearsed and stored in memory during the prior practice phase. Alternatively, on trials whose test phase required participants to make the secondary response, a low-pitched (100-Hz tone) response signal was presented instead, along with a musical staff showing the secondary response melody on the display screen. Participants had to sing the secondary response melody as quickly and accurately as possible.

Participants did not know beforehand which response melody would have to be sung, but they were instructed always to be prepared to make the primary response. The primary response melody was required on half of the trials, and the secondary response melody was required on the other half.

Participants' responses were timed online. The timing of a trial began with the presentation of the response go-signal and ended with the completion of the response melody. Onset and offset of the response were computed by the software voicekey; however, these computer-generated markings could be overridden by the experimenter (e.g., when the response was preceded by a cough, which caused the voicekey to mark the response onset incorrectly). In this way, the latency and duration associated with the production of each response melody were measured. An experimenter rating (correct or incorrect) and a computer-generated point score for the trial were also recorded following each trial and were presented to the participant on the display monitor.

Participants earned points for fast and accurate singing of both primary and secondary melodies. Correct primary responses that were completed within 1,200 ms received a score of 30 points; if the correct response was given but the completion time (latency plus duration) was greater than 1,200 ms, a score of 0 points was given. For incorrect responses, 30 points were subtracted. For secondary responses, participants were given a full 3 s to start and finish singing. They received a base score of 30 points when their secondary response was correct. This base score was then adjusted for the speed of their response: 1 point was subtracted for each 100 ms of

³E.L.Z. and D.L.F. served as experimenters. Both are experienced singers and were competent to judge the quality of the participants' performance.

Table 1

Mean Primary and Secondary Response Latencies (in Milliseconds), Durations (in Milliseconds), and Error Rates (in Percentages) by Primary and Secondary Melody Characteristics in Experiment 1

Prim-sec relationship	Melody key		Primary response			Secondary response		
	Primary	Secondary	Latency	Duration	Error	Latency	Duration	Error
IKSC	C	C	368	828	7.8	508	909	16.1
	F	F	365	838	2.1	534	928	29.0
IKDC	C	C	369	831	2.1	499	876	15.6
	F	F	373	831	4.2	521	876	31.8
DKSC	C	F	379	826	1.6	584	918	30.7
	F	C	368	836	2.1	538	895	33.2
DKDC	C	F	374	823	6.3	592	903	45.3
	F	C	376	831	5.7	561	913	25.5

Note. Latencies and durations are computed from correct responses only. IKSC = identical key, similar contour; IKDC = identical key, different contour; DKSC = different key, similar contour; DKDC = different key, different contour.

completion time. Thus, a correct response that took the entire allotted time of 3 s would result in a score of 0. For incorrect secondary responses, 30 points were subtracted from the participant's score.

Feedback was provided after each trial, both visually (on the display monitor) and verbally (from the experimenter). If an error was made, the participant was given the opportunity to hear and see the primary and secondary response melodies and to hear the response that he or she actually made so as to help him or her improve performance on subsequent trials.

Design and Data Analysis

Each participant was tested on 3 separate days within a 7-day period. The 1st day was considered practice. Only the data from the 2nd and 3rd days were used in the following analyses. Each daily session included up to eight blocks of trials: two practice blocks and six test blocks. Melodies in the practice blocks were not used in the test blocks, and data from the practice blocks were not included in the analyses. Each test block included 32 trials, with 16 primary and 16 secondary responses. Each primary response melody was paired twice with each of its four secondary response melodies. One of these pairings required the primary response, and the other required the secondary response. Trials were randomized within each block, and blocks were counterbalanced across participants. Data collected on each trial included response latency (time to start singing), response duration (time from when singing started to when it stopped), the experimenter's subjective evaluation of response accuracy (correct or incorrect), and error type (e.g., missed notes, weak response, singing the primary melody instead of the secondary melody). At the end of a trial block, the participant was shown his or her mean completion times, latencies, and response durations for both primary and secondary responses. A listing of the participant's total point score was also presented at the end of each block of trials. Participants completed at least three test blocks per session.

Results

The main results are shown in Table 1, which contains mean latencies, durations, and error rates for the four types of relationships between primary and secondary response melodies. For all analyses, latencies, durations, and error rates were averaged across sessions and blocks for each individual participant; repeated measures analyses of variance (ANOVAs) were then performed on these means. For latencies and durations, only correct responses were included in the computation of the participants' means. For each participant, these means were based on approximately 28 trials for primary responses and 22 trials for secondary responses;

exact *ns* depended on the number of blocks completed by each participant as well as the error rate for each condition.

Primary Responses

Mean latencies for primary responses were about as short (on the order of 370 ms) as they have been in experiments involving speech production (e.g., Gordon & Meyer, 1987; Sternberg et al., 1978). There was no effect of primary melody key, $t(3) = 0.38$, $p = .73$. Primary melodies in the key of F ($M = 371$ ms) were produced about as quickly as primary melodies in the key of C ($M = 373$ ms). Error rates were low, typically less than 5%. There was no effect of primary melody key on error rates, $t(3) = 0.69$, $p = .54$. Primary melodies in the key of F (3.5% errors) were produced about as accurately as primary melodies in the key of C (4.4% errors). Durations were relatively short—on the order of 830 ms (~165 ms per note). There was no effect of primary melody key on durations, $t(3) = 0.63$, $p = .57$. These findings lead us to believe that participants were essentially fully prepared to produce the primary melody, with the motor program for singing that melody loaded into motor working memory.

This assumption was further supported by data showing that participants were much faster and more accurate when singing the primary melody than when singing the secondary melody. Mean response latency for correct primary responses was 372 ms; the mean for secondary responses was 538 ms, $t(3) = 2.85$, $p = .03$ (one-tailed).⁴ Mean durations were also shorter for primary responses ($M = 831$ ms) than for secondary responses ($M = 902$ ms), $t(3) = 2.70$, $p = .04$ (one-tailed). Finally, participants made fewer errors when singing primary melodies (4.0%) than when

⁴ We restricted the use of one-tailed testing to situations in which a directional hypothesis was clearly appropriate. Here, the directional hypothesis that primary responses would be faster and more accurate than secondary responses sprang directly from the experimental design. Participants prepared fully to sing the primary melodies; this was not the case for the secondary melodies.

singing secondary melodies (28.4%), $t(3) = 2.63$, $p = .04$ (one-tailed).⁵

Secondary Responses

The overall error rate for singing secondary response melodies was relatively high. This was not unexpected, given that the participants' task was extremely challenging, even for well-trained and highly motivated singers. Although all participants considered themselves excellent sight singers and the response melodies were tonal, the present task differed from the typical experience of sight singing in several ways. First, responses were speeded, with participants trying to minimize the latency of each response. Second, abrupt key switches were required. Finally, some of the response melodies were similar to each other, with several notes shifted by only a half or whole step. Most errors involved singing one or more wrong notes in a melody (e.g., the third note of the scale when the fourth note was required); often the wrong note was an intrusion from the primary or one of the other secondary melodies.

The present task also differed in two important ways from typical choice-reaction tasks. Rather than requiring only one physical response, it entailed singing a series of five separate notes. The accuracy required for successful completion of each note is much greater than that required by, for example, a keypress task, in which one must ensure that the key is fully pressed but the placement of one's finger on the key is irrelevant. In contrast, when one sings a musical note, slight changes in larynx position, tongue placement, or lip shape can change the pitch. Much greater precision is therefore required to produce a correct response.

Even though error rates were relatively high here, the secondary response latencies and durations are still meaningful for several reasons. First, the results are unlikely to have stemmed from speed-accuracy trade-offs; factor effects on mean latencies were generally complemented by corresponding effects on mean error rates. In no case was a reliable effect on latency or duration accompanied by a reliable opposite effect on error rate. The high error rates are also unlikely to have stemmed from lackadaisical performance by the participants, because error rates for producing primary melodies were typically quite low (<5%). Also, as mentioned before, reaction times associated with the production of primary responses were typical of those produced by trained participants in previous experiments on speech production (e.g., Gordon & Meyer, 1987; Sternberg et al., 1978). Furthermore, the observed rate of errors on secondary response melodies suggests relatively "good" performance in at least one respect; each individual note of these melodies had a probability on the order of .90 or greater of being sung correctly when secondary melodies were produced (cf. Footnote 2).

Effects of Primary-Secondary Response Relationships: Melodic Contour and Musical Key

To determine the effect of primary-secondary relationship on secondary response performance, we performed three separate repeated measures ANOVAs (for latency, duration, and error rate), with melodic contour and musical key as the two independent variables. For the latency data, there was a reliable main effect of musical key, with shorter latencies to sing secondary melodies that were in the identical key as the primary ($M = 513$ ms for identical-

key trials; $M = 568$ ms for different-key trials), $t(3) = 4.73$, $p = .02$. There was no reliable effect of contour, $t(3) = 0.16$, $p = .89$, and no reliable interaction between key and contour, $t(3) = 1.05$, $p = .37$. These results are depicted graphically in Figure 4.

The results for error rates were similar (see Figure 5). There were more errors when the secondary melody was in a different key than the primary (33.7% vs. 23.1%); this result was marginally reliable, $t(3) = 2.60$, $p = .08$. Thus, the effect of key on secondary response latencies was not a result of a speed-accuracy trade-off. If one were to use latency and error data to compute the rate of information transmission in bits per second (Pachella, 1974), this effect would be highly reliable. There was no reliable main effect of melodic contour, $t(3) = 0.69$, $p = .54$, and no reliable interaction, $t(3) = 0.33$, $p = .77$. There were no reliable differences in the duration data (all $ps > .10$).

Additional Findings

We also obtained some additional important findings. Participants were faster at singing secondary melodies in the key of C than in the key of F. The mean latency for secondary responses in the key of C (regardless of primary-secondary relationship) was 526 ms, and the mean for secondary responses in the key of F was 559 ms, $t(3) = 2.80$, $p = .07$. There was no reliable difference in duration for the different keys, $t(3) = 0.70$, $p = .53$. However, secondary melodies in the key of C were also sung more accurately (22.6% error rate) than those in the key of F (34.2% error rate), $t(3) = 2.90$, $p = .06$. Although each result was marginally reliable, the composite significance level ($p = .004$) was highly reliable. Thus, on balance, it appears that secondary melodies in the key of C were easier for participants to sing.

Accompanying this main effect of musical key, there was a difference in secondary response latency depending on the direction of the key switch. When participants had to switch keys, they were faster at switching from a primary melody in F to a secondary melody in C (550 ms) than they were at switching from C to F (589 ms). In contrast, when the secondary response was in the same key as the primary response, average latencies were similar (503 ms for melodies in the key of C; 529 ms for melodies in the key of F). This switching asymmetry (i.e., the interaction between secondary key and whether a key switch was required) was reliable, $t(3) = 5.28$, $p = .01$.

One basis for this key-switch effect may be that the key of C is used as a general reference key and so is always at least partially available in memory. Furthermore, musical key information may be represented hierarchically in a production schema. In numerous studies using a variety of methodologies, Krumhansl (1979, 1985, 1990; Krumhansl & Kessler, 1982; Krumhansl & Shepard, 1979;

⁵ The present analyses concern main effects and interactions that have 1 degree of freedom. Their reliability was quantified originally in terms of F values. However, for purposes of exposition, we have transformed them to t values. This lets us test unidirectional (one-tailed) as well as bidirectional (two-tailed) hypotheses involving a priori predictions about the signs (positive or negative) that particular differences between mean reaction times should have. By squaring the reported t values, they may be transformed back to F values whose numerators have 1 degree of freedom and whose denominators have the same degrees of freedom as the t values (Hays, 1994).

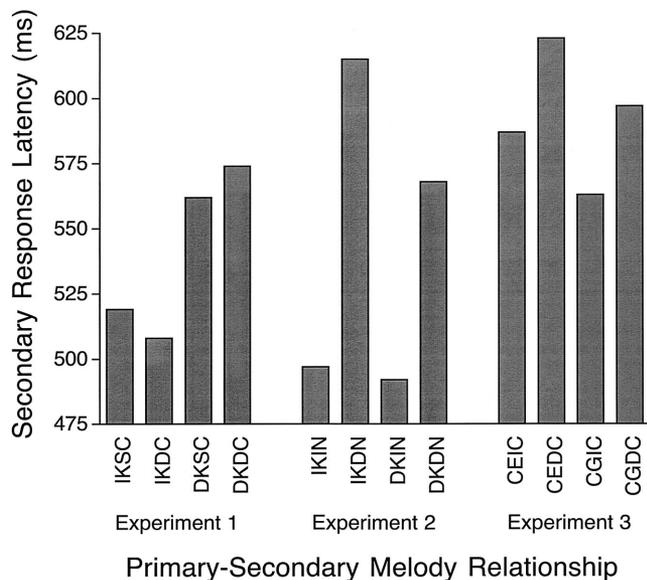


Figure 4. Mean latencies for correct singing of secondary melodies in Experiments 1–3. IKSC = identical key, similar contour; IKDC = identical key, different contour; DKSC = different key, similar contour; DKDC = different key, different contour; IKIN = identical key, identical note; IKDN = identical key, different note; DKIN = different key, identical note; DKDN = different key, different note; CEIC = switch between keys C and E, identical contour; CEDC = switch between keys C and E, different contour; CGIC = switch between keys C and G, identical contour; CGDC = switch between keys C and G, different contour.

Palmer & Krumhansl, 1987; see Krumhansl, 2000, for additional citations) has demonstrated the existence of a tonal hierarchy that is used in pitch and melody perception. Such a hierarchy may also be used in the construction of motor programs for pitch and melody production. In our experiment, participants were required to produce melodies in the key of C and in the key of F. Although the key of C includes six of the seven notes in the key of F, and the key of F includes (the same) six of seven notes in the key of C, the placement of these notes in Krumhansl's tonal hierarchy differs. When switching from producing a melody in C to producing a melody in F, perhaps one must retrieve the new tonic (F) from the third (and less accessible) level of the C-key hierarchy, whereas when one switches from F to C, the new tonic (C) may be retrieved from the second (and more easily accessible) level of the F-key hierarchy. Thus, the finding that switches from F to C are faster than switches from C to F is consistent with a hierarchical representation of key as found in Krumhansl's music perception studies.

Practice Effects

Because our participants were highly trained expert singers and accomplished sight readers, we expected that large practice effects would be less likely than if we were studying novice singers. To further minimize the possibility that increased practice with the procedure would affect performance during test trials, participants had an entire day of practice at the start of the experiment. Given these two elements of our experimental design, practice effects

should have been moderate at best during the second and third sessions of the experiment.

To test this expectation, we conducted three separate univariate ANOVAs with session, key, and contour as fixed factors; participant as a random factor; and response latencies, durations, and errors as the three dependent variables. Only responses from correct trials were included in the ANOVAs for latency and duration. For response latency, there was a relatively small main effect of session ($p = .13$) and no interactions with melodic key or contour (all $ps > .35$). For response duration, there was no main effect of session ($p = .22$); however, there was a marginally reliable ($p = .08$) interaction between session and contour, such that the difference in Session 2 between the mean similar-contour duration (943 ms) and the mean different-contour duration (907 ms) was larger than the difference between the corresponding two mean durations in Session 3 ($M = 891$ ms for similar contour; $M = 883$ ms for different contour). The main effect of contour was not reliable ($p = .21$). There were no other interactions with session (all $ps > .15$). For errors, there was a marginally significant ($p = .09$) main effect of session, such that participants had fewer errors in Session 3 (22.2%) than in Session 2 (34.6%). In predicting errors, there were no interactions with session (all $ps > .35$). Taken together, these results suggest that practice effects were minor. Participants' mode of performance was relatively consistent over time; no qualitatively different behavioral patterns or strategies emerged with increasing practice.

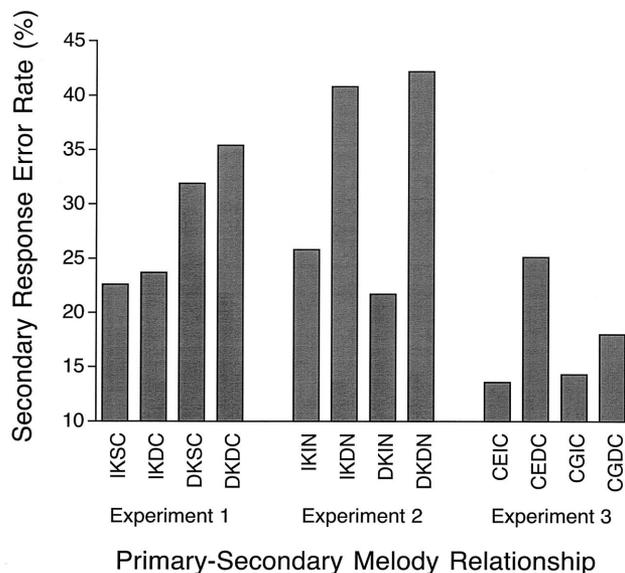


Figure 5. Mean error rates for singing secondary melodies in Experiments 1–3. IKSC = identical key, similar contour; IKDC = identical key, different contour; DKSC = different key, similar contour; DKDC = different key, different contour; IKIN = identical key, identical note; IKDN = identical key, different note; DKIN = different key, identical note; DKDN = different key, different note; CEIC = switch between keys C and E, identical contour; CEDC = switch between keys C and E, different contour; CGIC = switch between keys C and G, identical contour; CGDC = switch between keys C and G, different contour.

Discussion

The results from Experiment 1 suggest that musical key may play an important role in the representation of the motor program for singing a melody. When key information could be preserved while constructing a new motor program, secondary responses were faster and more accurate. However, when participants had to switch to a new key, they were slower and more prone to errors.

Contrary to Dowling's (1978; Dowling & Fujitani, 1971) findings about music perception, it made no difference whether participants had to switch to a secondary melody with a similar contour or to one with a different contour. This suggests that a "rough" representation of melodic contour (one that includes the direction but not the size of each tonal interval; i.e., *up, up, up, up* for a rising melody or *up, down, up, down* for a bumpy melody) is not a part of the motor programs for singing short tonal melodies, at least in expert singers.

Yet it seems unlikely that for the motor programs used by our participants, information about the musical notes of the primary and secondary melodies was stored merely as a series of specific articulatory motor commands or pitch targets (e.g., 262, 330, 392, 440, and 523 Hz in a C-key melody). If there were only such storage, then performance in singing the secondary melodies would depend only on how many identical motor commands or pitch targets were shared with the primary melodies at corresponding serial positions. Nevertheless, our results indicate that other aspects of the relationship between primary and secondary melodies affected participants' secondary response latencies and error rates. It appears, therefore, that the motor programs for singing these melodies incorporated some more abstract representation of this relationship. For example, one plausible hypothesis consistent with the present data is that singers' motor programs represent melodies in terms of their musical key plus a series of exact semitone offsets after the initial notes of the melodies.

Experiment 2

Although Experiment 1 provided some evidence that musical key may play a significant role in the representations of motor programs for singing primary and secondary melodies, this experiment covaried two major independent variables: the melodies' musical keys and initial notes. The melodies in the key of C always began with a C note, and the melodies in the key of F always began with an F note. Consequently, the effects on singing performance (i.e., response latency, duration, and error rate) that we have attributed tentatively to the musical-key factor are perhaps attributable instead to the identities of our melodies' initial notes.

Given this possibility, Experiment 2 was designed to separate these putative effects and to assess the relative importance of musical key versus initial note for speeded singing. Toward achieving this objective, the melodies for Experiment 2 were scored in the same keys (C and F) as those for Experiment 1. Also as for Experiment 1, all melodies had either the note C or the note F as their initial note. However, unlike for Experiment 1, initial note and key were not confounded for Experiment 2. Instead, they served as separate fully crossed (within-subject) factors.

Method

Participants

Three female and 2 male vocal performance or vocal music education graduate students served as paid participants. Four of them had participated in Experiment 1.

Response Melodies

The eight melodies used here appear in Figure 6. All had similar (bumpy) contours. Four of the melodies were in the key of C, and four were in the key of F. Of the melodies in C, two had C as the initial note, and two began with F. Similarly, two of the melodies in the key of F started with C, and two had F as the initial note. Each of the melodies in Figure 6 served as a primary response melody. For each primary response melody, any one of four other melodies from Figure 6 served as the accompanying secondary response melody. For example, for Melody 1 in the key of C (C4, F4, D4, G4, E4), the following four melodies could serve as the accompanying secondary response melody: Melody 2 in the key of C (C4, F4, E4, A4, G4), Melody 4 in the key of C (F4, A4, B3, G4, C4), Melody 2 in the key of F (F4, Bb4, G4, C5, A4), and Melody 4 in the key of F (C4, A4, E4, G4, F4).

Secondary response melodies were related to primary response melodies in one of four ways: (a) identical key and identical initial note (IKIN), (b) identical key and different initial note (IKDN), (c) different key and identical initial note (DKIN), or (d) different key and different initial note (DKDN). All four relationships occurred equally often throughout the experiment. For all four relationships, the contour of the secondary melody was similar to that of the primary melody (i.e., the pattern of intervals was *up, down, up, down*) without being an exact transposition.

Procedure

The apparatus, design, and procedure were identical to those of Experiment 1. Participants were tested on 3 separate days, and they performed at least three test blocks per session.

Results

The main results from Experiment 2 are presented in Table 2. As in Experiment 1, only correct responses were included in the computation of means latencies and duration. For each participant, these means were based on approximately 25 trials for primary responses and 18 trials for secondary responses; exact *ns*



Figure 6. Melodies used in Experiment 2. C₁ = Melody 1, key of C; C₂ = Melody 2, key of C; C₃ = Melody 3, key of C; C₄ = Melody 4, key of C; F₁ = Melody 2, key of F; F₂ = Melody 2, key of F; F₃ = Melody 3, key of F; F₄ = Melody 4, key of F.

Table 2

Mean Primary and Secondary Response Latencies (in Milliseconds), Durations (in Milliseconds), and Error Rates (in Percentages) by Primary and Secondary Melody Characteristics in Experiment 2

Primary melody		Secondary melody		Prim-sec relationship	Primary response			Secondary response		
Key	Initial note	Key	Initial note		Latency	Duration	Error	Latency	Duration	Error
C	C	C	C	IKIN	359	703	6.7	494	871	30.0
		C	F	IKDN	353	699	3.3	678	1,030	30.0
		F	F	DKDN	383	702	0.0	571	963	45.8
		F	C	DKIN	359	697	3.3	479	898	6.7
C	F	C	C	IKDN	370	717	3.3	582	1,004	36.7
		C	F	IKIN	361	730	3.3	555	933	23.3
		F	F	DKIN	358	716	10.0	493	922	30.0
		F	C	DKDN	376	729	3.3	590	812	40.2
F	F	C	C	DKDN	352	715	6.7	551	879	16.6
		C	F	DKIN	370	717	3.3	537	1,010	23.3
		F	F	IKIN	384	713	6.7	488	921	26.7
		F	C	IKDN	360	713	3.3	537	941	23.5
F	C	C	C	DKIN	360	716	3.3	459	932	26.7
		C	F	DKDN	355	717	0.0	565	843	26.9
		F	F	IKDN	373	708	3.3	621	1,078	33.5
		F	C	IKIN	358	716	10.0	463	1,045	23.3

Note. Latencies and durations are computed from correct responses only. IKIN = identical key, identical first note; IKDN = identical key, different first note; DKIN = different key, identical first note; DKDN = different key, different first note.

depended on the number of blocks completed by each participant as well as the error rate for each condition.

Primary Responses

As in Experiment 1, mean latencies for primary responses were short (on the order of 365 ms). There was no effect of primary melody key on mean latencies, $t(4) = 0.08$, $p = .94$. Primary melodies in the key of F ($M = 364$ ms) were produced as quickly as primary melodies in the key of C ($M = 365$ ms). Error rates were low, typically less than 5%. There was no effect of primary key on error rates, $t(4) = 0.19$, $p = .86$. Primary melodies in the key of F (4.6% errors) were produced about as accurately as primary melodies in the key of C (4.2% errors). Mean response durations were relatively short—on the order of 715 ms (~145 ms per note). There was no effect of primary key for them, $t(4) = 0.64$, $p = .56$. These findings lead us to believe that, as in Experiment 1, participants were essentially fully prepared to produce the primary melody.

Again as in Experiment 1, participants were faster and more accurate when singing the primary melody than when singing the secondary melody. Correct primary response latencies had a mean of 364 ms, and correct secondary response latencies had a mean of 528 ms, $t(4) = 6.98$, $p = .001$ (one-tailed). Error rates were lower for primary (4.4%) than for secondary (32.7%) responses, $t(4) = 2.59$, $p = .03$ (one-tailed). There was also a reliable difference in response durations; primary response duration averaged 713 ms, compared with 948 ms for secondary response duration, $t(4) = 2.13$, $p = .05$ (one-tailed). As in Experiment 1, participants were considerably more prepared to make the primary rather than the secondary response.

Effects of Primary–Secondary Response Relationships: Musical Key and Initial Note

There was no reliable effect of key on secondary response latencies. The mean response latency (556 ms) for singing second-

ary melodies in a key identical to the primary melody did not differ reliably from the mean response latency (530 ms) for singing secondary melodies in a different key, $t(4) = 1.32$, $p = .26$. However, responses for secondary melodies with the same initial note as the primary melody had reliably shorter mean latencies (494 ms) than responses for secondary melodies that started with a different note than the primary (591 ms), $t(4) = 3.43$, $p = .03$. The interaction between key and initial note was not reliable, $t(4) = 1.06$, $p = .35$. These results are depicted in Figure 4. There were no main effects or interactions for the mean response durations (all $ps > .10$).

As in Experiment 1, the results for secondary error rates were similar to those for latencies (see Figure 5). The mean error rate for secondary responses in a key identical to the primary melody was 33.3%; this did not differ reliably from the 32.0% error rate in the different-key conditions, $t(4) = 0.34$, $p = .75$. However, there was a marginally reliable effect of initial note, $t(4) = 2.39$, $p = .07$. Secondary melodies beginning with the same initial note as the primary had an error rate of 23.8% versus 41.6% for secondary melodies beginning with a different note. Thus, the effect of initial note on mean secondary response latencies was probably not attributable to a speed–accuracy trade-off. There was no interaction between the key and initial note factors, $t(4) = 1.01$, $p = .37$.

Additional Findings

In Experiment 1, our data suggested that it was easier to sing secondary melodies in the key of C. Response latencies for secondary melodies in the key of C were shorter and accuracy was greater than for secondary melodies in the key of F. For response latencies, there was also an interaction between key switching and secondary key. When a key switch was not required, latencies were similar regardless of whether the secondary key was C or F. However, when participants were required to switch to a different key, they were slower when switching from C to F than from F to C.

These preceding results did not generalize entirely to Experiment 2. Here, participants were not faster and more accurate when singing secondary melodies in the key of C. Instead, correct responses for secondary melodies in the key of F had shorter mean latencies (508 ms) than did those in the key of C (545 ms), $t(4) = 3.11, p = .04$. There were no reliable differences in durations (C melodies: $M = 946$ ms; F melodies: $M = 950$ ms) or error rates (29.2% for the key of C; 36.1% for the key of F; both p s $> .25$).

We did, however, replicate the previous interaction between key switch and secondary key. When a key switch was not required, error rates for secondary melodies in the key of C ($M = 30.0\%$) approximated error rates for secondary melodies in the key of F ($M = 26.8\%$). Yet when singers switched keys, error rates for secondary melodies in the key of C ($M = 23.4\%$) were lower than error rates for secondary melodies in the key of F ($M = 30.7\%$). This switching asymmetry (i.e., the interaction between secondary key and key switch) was marginally reliable, $t(4) = 2.15, p = .10$. Although there was no reliable interaction in the secondary response latency data, these exhibited the same qualitative pattern as the error rate data, thus suggesting that the error rate data were not attributable to a speed–accuracy trade-off.

Practice Effects

As in Experiment 1, we expected that practice effects would be moderate. We tested this expectation by conducting three separate univariate ANOVAs with session, key, and initial note as fixed factors; participant as a random factor; and latencies, durations, and errors as the dependent variables. There were no main effects of session, and session did not interact with key, initial note, or the Key \times Initial Note interaction (all p s $> .10$).

To further test for any effects of practice, we ran the main analyses separately, using only data from the 1 participant who had not been in Experiment 1. Substantive results for this new participant did not differ significantly from the overall results reported previously. In particular, there was no reliable effect of musical key on secondary response latencies ($p = .25$), durations ($p = .48$), or errors ($p = .25$). There were also no reliable interactions between key and initial note (all p s $> .25$). However, there was a reliable effect of initial note on mean secondary response latencies. They were shorter when the secondary melody's initial note was the same as that of the primary melody, $t(33) = 2.83, p = .008$. In addition, the accompanying error rates were lower, $t(33) = 4.84, p < .0001$.

Discussion

Viewed overall, the results of Experiment 2 suggest that for skilled singers, the initial note rather than the key of a melody provides one of the fundamental features of the motor program to begin melody production. When singers had to switch from a primary to a secondary melody that began with a different initial note, they were slower and less accurate. This happened whether or not they were also switching to a different key. Surprisingly, however, switching to a different key had no reliable effects on singers' latencies, durations, or error rates.

Our present findings therefore contrast with past ones in the perception and music-memory literatures. There, musical key has appeared to be an important feature of people's schematic percep-

tual representations for melodies. In our study of melody production, though, there was no reliable effect of key, only an effect of initial note. Given the differences between the tasks of singing and recognizing a melody, it is not surprising that initial note should play a fundamental role in motor programs for singing. The absence of a main effect of musical key is, however, quite surprising. Because key plays such an important role in music theory and performance, it is interesting that having to switch to a different key did not impair singing performance in Experiment 2.

Nonetheless, other aspects of our results suggest that musical key may play some significant implicit role in the structure of motor programs for melodic singing. The relative ease or difficulty of switching between primary and secondary melodies with different initial notes seemed to depend on the role that the notes played in the primary melody's key. It was not equally easy for singers to switch between any two initial notes. In both Experiments 1 and 2, singers were better at switching from F to C than vice versa.

These results lead us to several tentative conclusions about the structure of the motor programs for melodic singing. Our data suggest that rather than involving musical key as a principal basis, the abstract version of a skilled singer's motor program may contain a representation of initial note plus a series of exact semitone offsets. If so, then when a switch to a different initial note is required, singers may simply delete the representation of initial note from the motor programs for the primary melody (i.e., the note that they were prepared to sing) and replace it with the representation of the initial note for the secondary melody. However, perhaps not every possible different initial note can be accessed with equal ease; it may be easier for singers to access a representation of C when they are prepared to sing in the key of F than it is for them to access a representation of F when they are prepared to sing in the key of C.

Experiment 3

Experiment 3 was designed to evaluate these conclusions by extending the present research to musical keys other than C and F. A second purpose of Experiment 3 was to further test the effect of melodic contour on singers' performance. In Experiment 1, it made no difference whether singers were required to switch to a secondary melody with a contour that was similar to or different from the primary melody's contour. However, no trials there included switches to a secondary melody that had exactly the same contour (i.e., an exact transposition). So, to extend our results, Experiment 3 included trials on which the secondary melody was an exact transposition of the primary. This let us test whether the motor programs for singing short tonal melodies include information about both the initial note and the exact intervals between subsequent notes of a melody.

In Experiment 3, all melodies had the tonic as the initial note. Primary and secondary melodies were scored in one of three keys: C, E, or G. In addition, secondary melodies were either an exact transposition of the primary or they had a completely different contour. The design was thus a 2 (contour) \times 3 (key) within-subject manipulation.



Figure 7. Melodies used in Experiment 3. R₁C = Rising Melody 1, key of C; R₂C = Rising Melody 2, key of C; B₁C = Bumpy Melody 1, key of C; B₂C = Bumpy Melody 2, key of C; R₁E = Rising Melody 1, key of E; R₂E = Rising Melody 2, key of E; B₁E = Bumpy Melody 1, key of E; B₂E = Bumpy Melody 2, key of E; R₁G = Rising Melody 1, key of G; R₂G = Rising Melody 2, key of G; B₁G = Bumpy Melody 1, key of G; B₂G = Bumpy Melody 2, key of G.

Method

Participants

Participants were the same 5 singers as in Experiment 2.

Response Melodies

Two rising and two bumpy melodies were constructed for this experiment. Each melody was presented in three different keys (C, E, and G),

making a total of 12 different frequency patterns, as shown in Figure 7. All 12 melodies were used as both primary and secondary melodies, but not in every possible combination. For each primary response melody in the key of C, any one of four other melodies from Figure 7 served as the accompanying secondary response melody. For each primary response melody in the key of E or G, any one of two other melodies served as the accompanying secondary response melody. For example, for Rising Melody 1 in the key of C (C4, E4, G4, A4, C5), the following four melodies could serve as the accompanying secondary response melody: Rising Melody 1 in the key of E (E4, G#4, B4, C#5, E5), Bumpy Melody 1 in the key of E (E4, A4, G#4, C#5, B4), Rising Melody 1 in the key of G (G4, B4, D5, E5, G5), and Bumpy Melody 1 in the key of G (G4, C5, B4, E5, D5). For Bumpy Melody 2 in the key of G (G4, A4, F#4, C5, B4), the following two melodies could serve as the accompanying secondary response melody: Rising Melody 2 in the key of C (C4, D4, F4, B4, C5) and Bumpy Melody 2 in the key of C (C4, D4, B3, F4, E4).

Secondary melodies differed from primary melodies along two dimensions—contour and key distance. Each secondary melody had either the identical contour as the primary (i.e., the two melodies were exact transpositions [Condition IC]), or the secondary had a different contour than the primary (i.e., if the primary was a bumpy melody, then the secondary was a rising melody, or vice versa [Condition DC]). Along the dimension of key distance, the primary and secondary melodies were scored either in two keys that share 6 of 7 notes (C and G [Condition CG]) or in keys that share only 3 of 7 notes (C and E [Condition CE]). Thus, there were four basic conditions in the experiment: CEIC (switch between C and E or E and C, identical contour), CEDC (switch between C and E or E and C, different contours), CGIC (switch between C and G or G and C, identical contour), and CGDC (switch between C and G or G and C, different contours).

Procedure

The apparatus, design, and procedure were identical to those of Experiments 1 and 2. Participants were tested in three separate 1-hr sessions within a 7-day period. Each test block involved melodies in only two keys (either C and G or C and E).

Results

The main results from Experiment 3 are presented in Table 3; summaries of error rate and latency data are also presented in Figures 4 and 5. As for Experiments 1 and 2, only correct responses were included in the computation of mean latencies and durations. For each participant, these means were based on approximately 32 trials for primary responses and 28 trials for

Table 3
Mean Primary and Secondary Response Latencies (in Milliseconds), Durations (in Milliseconds), and Error Rates (in Percentages) by Primary and Secondary Melody Characteristics in Experiment 3

Contour relationship	Melody key		Primary response			Secondary response		
	Primary	Secondary	Latency	Duration	Error	Latency	Duration	Error
IC	C	G	342	665	1.0	645	712	16.8
	C	E	349	665	2.3	668	716	11.2
	G	C	353	685	8.0	464	713	11.8
	E	C	344	678	3.5	486	699	16.0
DC	C	G	343	663	4.5	714	724	21.5
	C	E	345	671	0.0	691	733	22.8
	G	C	348	681	4.5	471	705	14.5
	E	C	338	680	2.7	520	714	27.5

Note. Latencies and durations are computed from correct responses only. IC = identical contour; DC = different contour.

secondary responses; exact *ns* depended on the number of blocks completed by each participant as well as the error rate for each condition.

Primary Responses

As in Experiments 1 and 2, mean latencies for primary responses were short (on the order of 345 ms). There was no effect of primary key on mean latencies, $F(2, 8) = 0.91, p = .44$. Primary melodies in the key of G ($M = 351$ ms), in the key of E ($M = 341$ ms), and in the key of C ($M = 345$ ms) were produced about equally quickly. Error rates were low, typically less than 5%. There was no effect of primary key on error rates, $F(2, 8) = 1.63, p = .26$. Primary melodies in the key of C (2.0% errors) were not produced reliably more accurately than primary melodies in the key of G (6.3% errors) or primary melodies in the key of E (3.1% errors). Mean response durations were relatively short (on the order of 675 ms). There was an effect of primary key on them, $F(2, 8) = 4.78, p = .04$. Melodies in the key of C ($M = 666$ ms) had shorter durations than those in the key of G ($M = 683$ ms) or those in the key of E ($M = 679$ ms). As in Experiments 1 and 2, it appears that participants were essentially fully prepared to produce the primary melody.

Also as in our first two experiments, the mean difference between correct primary and secondary response latencies was reliable. Primary melodies had shorter latencies ($M = 346$ ms) than secondary melodies ($M = 591$ ms), $t(4) = 1.85, p = .07$ (one-tailed). Durations were shorter for primary melodies ($M = 673$ ms) than secondary melodies ($M = 714$ ms), $t(4) = 4.08, p = .008$ (one-tailed). Errors were also less frequent for primary responses (3.4%) than for secondary responses (17.7%), $t(4) = 2.24, p = .04$ (one-tailed).

Effects of Primary–Secondary Response Relationships: Key Distance and Melodic Contour

Unlike in Experiment 1, in which switching to a similar versus a different contour had no effect on mean latencies, durations, or error rates, participants in Experiment 3 were at least marginally faster and more accurate when singing exact transpositions than when switching to a different contour. The mean latency for producing secondary response melodies with a contour identical to the primary melody was 575 ms, whereas the mean latency for producing those secondary response melodies that had a different contour than the primary melody was 610 ms, $t(4) = 2.17, p = .09$. Exact transpositions had slightly shorter durations ($M = 710$ ms) than secondary melodies with a different contour ($M = 719$ ms), $t(4) = 1.96, p = .12$. Error rates were also lower for exact transpositions (13.9%) than for trials on which a different pattern of notes was sung (21.6%), $t(4) = 1.88, p = .13$. Although each of the preceding three contrasts is just marginally reliable at best, combining across them yielded a composite significance level that was highly reliable ($p = .001$), suggesting that melodic contour (exact transposition vs. arbitrary transition) was a crucial factor for secondary melody production. Thus, exact transpositions may be a special case of switching melodic contours.

The fact that secondary melodies with exact transpositions were easier to sing suggests that melodic contour is represented in the same form across keys. This rules out the possibility that contour

representations involve a series of exact frequency offsets (e.g., for the melody in Figure 2, exact frequency offsets from the initial note would be represented as 0 Hz, +91 Hz, –56 Hz, +43 Hz, and –20 Hz). Because the musical scale is logarithmic rather than linear, these offsets are the appropriate ones only for a melody whose initial note (349 Hz) is the F above middle C. In contrast, a representation of contour based on semitone offsets is invariant across key (e.g., for the melody in Figure 2, semitone offsets from the initial note are 0, +4, –3, +2, and –1). Our data therefore indicate that this latter representation is more plausible.

Furthermore, the “distance” between keys also mattered somewhat in Experiment 3. Switching between the keys of C and G was easier for most participants than switching between C and E. Mean response latencies and durations for these two types of key switching did not differ (both $ts < 0.80$, both $ps > .45$). However, error rates were lower when singers switched between C and G (16.1%) than when they switched between C and E (19.4%), $t(4) = 2.87, p = .05$. This finding must be interpreted in light of the results from Experiment 2, which suggest that the lower error rate for switching between the keys of C and G was actually a result of switching between initial notes rather than key per se. (Recall that in Experiment 3, as in Experiment 1, initial note and key were perfectly confounded).

It is interesting to note, as well, that the key-distance effect stemmed mostly from trials on which the secondary melody was in the key of C (see Table 4). When participants prepared to sing a primary melody in C and then had to switch to a secondary melody in E or G, the secondary key had little effect on their mean latencies, durations, or error rates (all $ts < 0.80$, all $ps > .45$). However, when they prepared to sing a primary melody in E or G and then had to switch instead to C, a key-switch effect occurred in the error rates, $t(4) = 3.63, p = .02$. Switching from a primary melody in G to a secondary melody in C yielded an error rate of 13.1%; in contrast, switching from a primary melody in E to a secondary melody in C yielded an error rate of 21.8%.

These data are consistent with the hierarchical organization of musical key demonstrated by Krumhansl (1979, 1990) and pre-

Table 4
Mean Secondary Response Latencies (in Milliseconds),
Durations (in Milliseconds), and Error Rates (in Percentages)
by Primary and Secondary Response Melody Keys in
Experiment 3

Primary melody key	Secondary melody key		
	C	G	E
C			
Latency		679	679
Duration		719	724
Error		19.1	17.0
G			
Latency	467		
Duration	709		
Error	13.1		
E			
Latency		499	
Duration		705	
Error		21.8	

Note. Latencies and durations are computed from correct responses only.

sented in Figure 1. When the primary melody is in the key of C, the notes E and G both may be represented at the second level of the motor-program hierarchy; thus, as a result, it should be equally easy to switch to either E or G. According to this account, however, the note C is not represented at the same hierarchy level in the key of G as in the key of E. For the key of G, the note C, although not part of the major triad, is a part of the scale; thus, it is represented at the third level of the motor-program hierarchy. In contrast, for the key of E, the note C is not a part of the scale (C# is); thus, C is represented at the lowest (fourth) level of the hierarchy. We would expect, therefore, that switching from E to C should take longer and/or be more likely to result in errors than switching from G to C; this expectation was supported by our data.⁶

Additional Findings

Responses for secondary melodies in the key of C had shorter durations ($M = 707$ ms) than responses for those in the keys of G ($M = 719$ ms) or E ($M = 724$ ms). This overall effect of secondary key was marginally reliable, $F(2, 8) = 3.46, p = .08$. The contrast comparing the key of C with the average of the keys G and E was also marginally reliable, $F(1, 4) = 5.47, p = .08$.

Practice Effects

As in Experiments 1 and 2, we tested for practice effects by running three univariate ANOVAs with session, key, and contour as fixed factors; participant as a random factor; and mean response latencies, durations, and errors as the three dependent variables. There were reliable main effects of session on both mean latencies ($p = .02$) and durations ($p = .01$). Mean response latencies were shorter in Session 3 (578 ms) than in Session 2 (613 ms); mean response durations were also shorter in Session 3 (702 ms) than in Session 2 (728 ms). There was no reliable effect of session on errors ($p = .20$). As in Experiments 1 and 2, practice effects did not interact with any other factor effects of theoretical interest. Increasing practice did not appear to change participants' performance strategies or patterns of behavior in any qualitative way.

Discussion

Previously, on the basis of results from Experiment 2, we showed that initial note is a fundamental coding feature of motor programs for singing. Supplementing this demonstration, the results of Experiment 3 reveal that musical key may also be important in a certain special way. The initial note of a melody is apparently stored or coded so as to indicate its role or position in a particular tonal context. For example, the note C in the key of F may play a different role than C in the key of E or the key of G. This follows because in Experiment 3, it was easier for singers to switch to the initial note C when they had prepared to sing a melody in the key of G than when they had prepared to sing a melody in E. Similarly, in both Experiments 1 and 2, it was easier to switch from F to C than from C to F. These similarities suggest that switching to a new initial note for a secondary melody involves more than simply deleting one parameter value from a prepared primary motor program and replacing it with another. Rather, the switch appears to take longer if the new initial note

occupies a less accessible position in the hierarchical representation of the key for the primary melody.

Experiment 3 also showed that switches to exact transposition melodies are easier than switches to melodies with a different contour. This result, along with those from Experiment 1, permits some inferences about the representation of melodic contour in motor programs for singing. Specifically, we infer that contour is not represented in a "rough" form (e.g., *up, up, up, up* vs. *up, down, up, down*). If it were, then switches to melodies with the same rough contour should be easier than switches to melodies with a different contour. However, the results of Experiment 1 demonstrated that these two types of switches are equally difficult. Instead, taking the results of Experiment 3 into account, it appears that melodic contour may be represented as a series of exact semitone offsets (e.g., *up one semitone, down four semitones*, etc.).

General Discussion

Our goal in conducting these experiments was to characterize the motor programs used by expert singers for producing short tonal melodies. We sought to discover whether the melodic elements important in music perception are also important in skilled singing or whether music production and perception differ because of the distinct cognitive representations and behaviors involved in them. In addition, we wanted to start developing a theoretical model of the cognitive representations and processes involved in accurately singing short tonal melodies.

The results from Experiments 1–3 provide a basis for this development. In Experiment 1, expert singers were equally fast when producing a secondary melody with a qualitatively similar (though nonidentical) contour and when producing a secondary melody with a contour qualitatively different from that of a previously prepared primary melody. Apparently, contrary to what one would predict from conclusions about music perception (e.g., Dowling, 1978; Dowling & Fujitani, 1971), a rough representation of melodic contour is not used by expert singers to produce short tonal melodies. However, musical key may have been important during Experiment 1; there, secondary melodies were produced more slowly when they required switching of musical keys. But because key was perfectly confounded with initial note in Experiment 1, a second experiment was necessary to disentangle the effects of initial note and musical key.

The results of Experiment 2 indicated that the "key" effect found in Experiment 1 was actually an effect of initial note. That is, in Experiment 2, participants were slower when switching to

⁶ One way in which our data did not fit this expectation is that switching from C to E or G was more difficult than switching from G to C. Under the hierarchy model, one would expect a switch to a key whose tonic is represented at the second level of the hierarchy (as E and G are for the key of C) to be accomplished more easily than switching to a key represented at the third level (as C is for the key of G). That we did not find this may be an artifact related specifically to the key of C. This key, with no sharps or flats, is the easiest key to play or sing, and it is highly familiar. In addition, during Experiment 3, melodies in the key of C were presented twice as often as were melodies in the key of E and melodies in the key of G. Thus, even within the context of this experiment, participants received extra practice producing melodies in the key of C. This may account for the relative facility with which they were able to switch to the key of C.

produce a secondary melody with a different initial note; however, they did not slow down when switching to produce a secondary melody with a different musical key. Key per se, therefore, seems not to be the principal organizing basis of singers' motor programs for melody production.

Furthermore, in Experiment 3, singing a secondary melody that was an exact transposition of the primary melody was easier than singing a secondary melody with a different contour than the primary. This suggests that notes in singers' motor programs are represented as semitone offsets from some standard (e.g., initial note of melody, tonic of key, or absolute standard like middle C) rather than as exact frequencies. Also, because switches between some pairs of keys were easier than switches between other pairs of keys, it seems likely that the notes of a musical key are

represented in motor programs as a hierarchy similar to that demonstrated by Krumhansl (1990) for music perception.

A Model of Motor Programming for Expert Singing

Given the present findings, we have formulated an initial, tentative theoretical model to describe the knowledge representations, memory stores, and cognitive processes whereby expert singers construct and execute vocal motor programs for singing short tonal melodies (see Figure 8). The following subsections outline our model's assumptions and explain how they account for our experiments' results while leading to further testable predictions about expert singing.

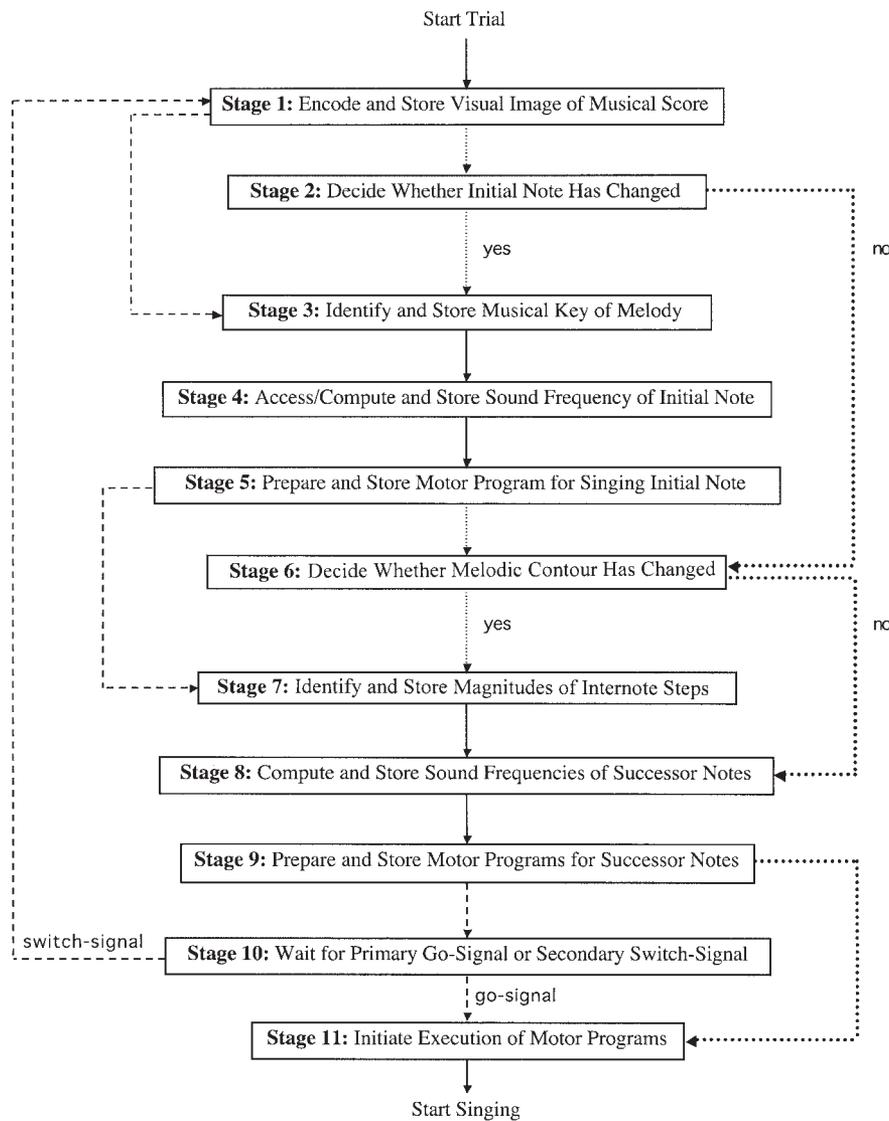


Figure 8. Proposed stages of processing for preparation of motor programs to sing primary and secondary melodies in the response-priming procedure of Experiments 1–3. Dashed lines denote paths between stages that are only taken in the case of primary melodies. Dotted lines denote paths that are only taken in the case of secondary melodies. Solid lines denote paths that are taken regardless of the melodies' status.

Knowledge Representations and Memory Stores in the Model

In our model, several distinct complementary types of knowledge representations and memory stores are used to construct motor programs for singing melodies. Some of these representations are abstract and presumably stored in various sectors of long-term memory. Other representations are presumably stored in various sectors of short-term (*working*) memory and may be more or less “concrete,” depending on their particular functions as part of expert singing. Complementing working-memory representations for visual images of musical scores and for sound frequencies of individual notes, our model includes the following memory stores.

Long-term memory for key-note information. Inspired by Krumhansl (1990), we assume that for each standard key, an expert singer has permanently stored a hierarchical structure (see Figure 1) that contains information about the 12 notes of the chromatic scale. This hierarchy specifies the role that each note plays in that particular key, the probable frequency of each note’s inclusion in standard Western compositions, and other relevant details (including, for singers with absolute pitch, information about exact sound frequencies). With such a structure, it would be more or less easy to access information about individual notes, depending on their relative depth in the hierarchy.

Working memory for musical-key identities. An important part of working memory in our model is devoted to storing the identity of a transcribed melody’s musical key. We assume that when the score for a melody is presented, an expert singer evaluates its visual image through perceptual recognition processes to identify the melody’s key, and the key identity is stored temporarily. The stored key identity serves as a pointer to long-term memory for accessing information about the properties of the melody’s individual notes.

Working memory for magnitudes of internote steps. Also stored in working memory are the magnitudes of the scalar *steps* (or offsets) between a melody’s successive notes. These magnitudes are relevant because for any melody, once its initial note is specified, the subsequent notes can be determined completely by how many upward (positive) or downward (negative) semitone steps separate each one from the next.

Working memory for motor programs. Finally, the ultimate form of procedural knowledge for present purposes is the motor program used to sing a melodic note. We assume that for each note of a melody, singers construct a motor program that enables vocal production of the note’s required physical characteristics (e.g., sound frequency, duration, amplitude). The parameters of this program and the manner in which it is executed presumably conform at least somewhat to those described previously for production of discrete phonemic units (e.g., see Rosenbaum, 1991).

Cognitive Processes in the Model

Figure 8 outlines our model’s overall set of processing stages, which use the aforementioned knowledge representations and memory stores for our response-priming procedure. The overall sequence of processing depends on various factors that together account qualitatively for many of the effects that we observed on response latencies, durations, and error rates.⁷

Stage 1: Encoding transcribed musical scores. During the first stage, an image of the melody depicted by the transcription is formed and stored in visual working memory. This provides a basis for some later stages of processing that use the visual image.

Stage 2: Deciding whether initial note of melody has changed. The second stage of processing, which occurs only for secondary melodies, involves deciding whether the initial note of a secondary melody differs from the initial note in the immediately prior primary melody. Under our model, this decision is made by comparing the stored visual images of the primary and secondary melodies for the trial. If the initial note is judged to have changed across these images, then Stage 3 commences. Alternatively, if the initial note is judged not to have changed, then processing skips to Stage 6, as described below.

Stage 3: Identifying and storing musical key of melody. The third stage is relevant for both primary and secondary melodies. It identifies and stores the musical key of the current melody in working memory. This presumably involves a perceptual recognition process that compares certain aspects of the stored visual image for the melody’s musical score against appropriate information in long-term memory.

Stage 4: Accessing/computing and storing sound frequency of initial note. The fourth stage uses two types of information: the identity of the musical key for the relevant melody and the concomitant visual image that specifies the melody’s initial note. On the basis of this information, the initial note’s required sound frequency is retrieved or computed, and the obtained value is stored in working memory. The sound frequency of the melody’s initial note provides an intermediate step toward preparing a motor program for singing it, which happens in Stage 5.

For singers with absolute pitch, the sound frequency of any note can be retrieved from long-term memory. This process presumably involves accessing a hierarchical key-note structure like those shown in Figure 1. For other singers, the sound frequency of the initial note of a primary melody may be computed from the auditory input. When producing a secondary melody, however, they must obtain the sound frequency of the new initial note through online computations. We assume that such computations use information about the relative distances between notes in that musical key (stored in a hierarchical form in long-term memory) as well as information about the sound frequency of the primary melody’s initial note (stored in working memory). The retrieval or computation of a new initial note’s sound frequency may be more or less enhanced, depending on how high or low it is within the hierarchical structure for the musical key of the currently relevant melody (cf. Krumhansl, 1990).

Stage 5: Preparing and storing motor program for singing initial note. Using the stored sound frequency of the initial note for the currently relevant melody, the fifth stage prepares and

⁷ At first glance, our model may seem too complex. However, such complexity is required here because skilled melodic singing involves an enormously rich mixture of cognitive processing, perceptual-motor coordination, and physical action, with exquisite requirements for both performance speed and accuracy during the production of an extended sequence of individual, challenging articulatory movements. Hence, it is unlikely that one could encompass all of the fundamental aspects of this performance with a simpler model.

stores a motor program that, when executed subsequently, enables this note to be sung at the start of physical melody production. The prepared initial-note motor program, like those for the successive notes in the melody, may be stored in a working-memory buffer that contains modularized specifications for a variety of crucial parameters. These parameters could include the notes' melodic serial positions, sound frequencies, amplitudes, durations, and so forth (Rosenbaum, 1991). The separation of Stage 5 from other subsequent stages enables previously stored information about different parts of a primary melody to be reused efficiently when a more or less similar secondary melody has to be sung instead. After the fifth stage, processing will skip to Stage 7 if the currently relevant melody is primary but otherwise proceed directly to Stage 6 if it is secondary.

Stage 6: Deciding whether melodic contour has changed. The sixth stage decides whether the notes of a current secondary melody form a contour that is identical to or different from the one in an immediately prior primary melody. This decision presumably involves comparing the visual images of these melodies to check whether they have quantitatively different up-and-down spatial patterns of notes. If the decision is affirmative, then processing proceeds directly to Stage 7. Alternatively, if the decision is negative (i.e., the secondary melody is an exact transposition), then processing skips to Stage 8.

Stage 7: Identifying and storing magnitudes of internote steps. When the current melody is primary, or it is secondary but has a contour different from the immediately prior primary melody, the seventh stage identifies and stores the magnitudes of the current melody's successive internote steps. For example, in dealing with a melody like the one illustrated by Figure 2, Stage 7 would yield a stored sequence of internote steps whose magnitudes are +4, -7, +5, and -3 following the initial note. Such magnitudes provide the input to the eighth stage.

Stage 8: Computing and storing sound frequencies of successive notes. The eighth stage computes the sound frequencies of a melody's successive (postinitial) notes on the basis of previously stored values of the initial note's sound frequency and successive internote steps. We assume that this involves repeatedly applying a simple equation that relates the frequencies of adjacent notes and the magnitudes of the internote steps between them.⁸ The computation provides inputs to the ninth stage.

Stage 9: Preparing and storing motor programs for successive notes. The ninth stage prepares motor programs for singing a melody's successive notes in the same way as Stage 5 does for the initial note. With a primary melody, processing then proceeds directly to the tenth stage. Alternatively, with a secondary melody, processing skips to the eleventh stage.

Stage 10: Waiting for primary go-signal to start singing. The tenth stage involves waiting for the go-signal to begin primary melody production. If and when this signal occurs, processing proceeds immediately to Stage 11. Alternatively, if a switch-signal (i.e., another transcribed musical score) is given to sing a secondary melody instead, then processing reverts back to Stage 1, and a subset of the preceding stages is repeated.

Stage 11: Execution of motor programs. The eleventh stage initiates execution for all of the motor programs that have been prepared previously to sing the current melody's notes.

Theoretical Account of Experimental Results

Across Experiments 1–3, we obtained several major results that provide empirical tests of our model's assumptions and can be explained in terms of them.

Superior performance for primary melodies. In singing primary melodies, participants' mean response latencies and durations were always shorter, and their error rates were lower, than they were in singing secondary melodies. Of course, this is not surprising from the perspective of our model. Except for Stages 10 and 11, all of the model's stages are presumably completed ahead of time—before the ultimate signal to respond—in the case of primary melodies. Thus, Stages 1–9 and any factor effects on them would not contribute to observed primary response latencies. Also, during preparation to sing a primary melody, the products of these stages may be checked and refined, enabling the observed primary response durations to be shorter and error rates to be lower.

Better performance in singing secondary melodies with identical initial notes. For secondary melodies, we found that singers performed better when their initial notes were identical to those of the primary melody. According to our model, this occurred because when a secondary melody has an initial note identical to that of the prior primary melody, a singer bypasses identification of the secondary melody's musical key (Stage 3), accessing the frequency of its initial note (Stage 4) and preparing a new motor program to sing the initial note (Stage 5). Thus, these three early stages would not contribute to the secondary response latencies or error rates when an identical initial note is involved.

Null effects of identical musical keys on singing secondary melodies. Such considerations also explain, at least partly, why performance for secondary melodies during Experiment 2 did not depend much, if at all, on whether they had the same musical key as the immediately prior primary melodies. Instead, what mainly affected secondary response latencies and error rates there were identical initial notes (see Figures 4 and 5). This supports our model's assumption that Stage 3 is skipped when a secondary melody has the same initial note as the accompanying primary melody.

Asymmetries in singing secondary melodies. Nevertheless, throughout all three experiments, we did find that certain asymmetries occurred in singing secondary melodies. These asymmetries can be explained at least tentatively in terms of what happens under our model for preparing to sing secondary melodies when their initial notes and musical keys differ from those of the accompanying primary melodies. Specifically, in this case, Stage 4—identifying a melody's musical key—plays a major role. Its duration may depend on where the tonic note of the secondary melody's key is located in the hierarchical structure of notes for the primary melody's key (cf. Krumhansl, 1990). If this note is located at a relatively high position in the primary-key hierarchy, then less time may be required to retrieve information from long-term memory about the secondary melody's key and initial note. In turn, this would shorten the duration of Stage 4 and the response

⁸ Let f_i denote the frequency of note i . Let s_i denote the magnitude of step i , with positive values representing upward steps and negative values representing downward steps. Then the equation needed for the above computation is $f_{i+1} = f_i * (1.05946)^{s_i}$; $1 \leq i < n$, where n denotes the number of notes in the melody.

latencies to which it contributes. Thus, because the tonic note for the key of C is located at a relatively high position in the hierarchical structure for the key of F (see Figure 1), it should be easier to switch from a primary melody in the key of F to a secondary melody in the key of C—rather than vice versa—as Experiments 1 and 2 showed. By analogy, a similar explanation would account for why, in Experiment 3, switching from the key of G to the key of C was easier than switching from the key of E to the key of C.

Better performance in singing secondary melodies with identical contours. Another significant result of Experiment 3 was that singers produced secondary melodies with shorter latencies and fewer errors when they had contours identical to those of the accompanying primary melodies (i.e., the primary and secondary melodies were exact transpositions). From Figure 8, it can be seen that this result is likewise explained by our model. According to it, Stage 7, which identifies the magnitude of the secondary melody's internote steps, would be skipped when the secondary melody has a contour identical to that of the primary melody. Thus, under these circumstances, this stage would not contribute any additional time or errors to the overall preparation process for singing the secondary melodies, thereby yielding improved performance through their exact transpositions.

Explanation of typical errors. When a secondary melody's initial note was different from the primary melody's initial note, singers sometimes sang the right melody but in the wrong key. In our model, errors of this type could stem from inaccurate identification of the correct secondary key (Stage 3); flaws in computing, accessing, or storing the sound frequency of the initial note (Stage 4); or incorrect specification of the motor program for singing the initial note (Stage 5).

On other occasions, singers produced secondary melodies with the correct key; however, one or more of the tonal intervals was incorrect. Often, an interval from the primary melody or from one of the other secondary melodies was substituted for the correct interval. Under our model, errors of this type could stem from faulty computations of the internote steps (Stage 7), faulty identification or storage of sound frequencies (Stage 8), or incorrect specification of the motor programs for successive notes (Stage 9).

Alternative Models

Our tentative preferred theoretical model (see Figure 8) is consistent with the data obtained from all three experiments. Yet other theoretical alternatives can also explain at least some of the present data. One such model includes what has been variously described as *distributed planning* (Garcia-Colera & Semjen, 1988), *online amendment* (Inhoff, Rosenbaum, Gordon, & Campbell, 1984), or *online programming* (Ketelaars, Khan, & Franks, 1999; Smiley-Oyen & Worringham, 1996; see also Klapp & Wyatt, 1976). Under this latter model, participants in our response-priming procedure would begin singing the first note of a secondary melody before having fully prepared all five of the melody's notes. While production of the first note or two of the secondary melody was underway, other processes would modify the intermediate and terminal notes of the motor program. Models that include online programming processes such as these generally predict that factor effects related to intermediate and terminal notes of secondary melodies would be seen in response durations more so than in latencies.

As do models that include online programming processes, our preferred model (see Figure 8) gives a high priority to the motor program for the initial note of the melody. However, our model does not specifically allow for online programming—motor programs for all notes are prepared and stored in Stage 9, but execution of the motor programs does not begin until Stage 11. Our model thus predicts that factor effects related to intermediate and terminal notes of secondary melodies are generally seen in response latencies rather than in durations.

This prediction received some support. In Experiment 3, there was an exact transposition effect on response latencies for secondary melodies, suggesting that changes to intermediate and terminal notes were made before execution of the motor programs began. Thus, a radical version of the online-amendment hypothesis seems unlikely. Nevertheless, there was also a small exact-transposition effect on response durations, suggesting that some online amendment might have taken place. Further research will be necessary to ascertain the precise extent of online programming in the production of short melodies. Experimental designs that vary the serial position of the first discriminative response would be particularly useful (cf. Inhoff et al., 1984).

Another theoretical model that could explain some of our data is one that focuses on the perception and encoding of musical scores rather than the production of melodies. As in such a model, visual processing and encoding have been included in our model (see Figure 8, Stage 1). However, our model assumes that factor effects in all three experiments stemmed from varying amounts of time taken to modify existing motor programs rather than from differences in the amounts of time taken to read musical scores and encode the information contained within them. In our opinion, this is a reasonable assumption, especially for highly trained singers who are expert sight readers. Although novices might take longer to read and interpret a melody in a "difficult" key (e.g., the key of E), this seems much less likely to be true for experts, who will have extensive familiarity with singing in a variety of keys.

Yet to definitively rule out the possibility that our results are attributable to vocal production processes rather than to reading processes, further research is needed. For example, one possible set of studies might replicate those reported here while modifying the present procedure. In this extension, melodies would be associated with nonmusical symbols (e.g., geometric shapes) rather than musical scores. Participants would first memorize these symbol-melody associations, which would be counterbalanced across participants. Then, participants would be cued by the nonmusical symbols whenever secondary melody production was required. As a result, any obtained differences in latencies could be attributed unambiguously to production rather than to perceptual processes (cf. Rosenbaum, 1980).

Other Directions for Future Research

Although the theoretical model (see Figure 8) outlined here accounts for most of the results from the present three experiments, it also incorporates some assumptions that we have not explicitly tested. One untested assumption concerns the reference point for the working-memory representations of the magnitudes of internote steps computed in Stage 7. At least three different reference points are logically possible. A semitone offset (e.g., for the fourth note of the melody in Figure 2) could be computed by using as a

reference point either the initial note of the melody (e.g., G4 is offset from initial note F4 by two semitones), the nearest lower frequency tonic of the melody (e.g., G4 is offset from tonic C4 by seven semitones), or the previous note in the melody (e.g., G4 is offset from previous note D4 by five semitones). We have assumed (but not tested) that each semitone offset is based on the immediately previous note in the melody. In future work, this assumption could be explicitly tested versus other possibilities.

Further studies could also test more specific hypotheses related to the proposed hierarchical nature of long-term memory for key-note information. For example, a switch from an initial note of G to an initial note of E should be easier when one has prepared to sing in C (in which both G and E are on the second tier of the hierarchy) than when one has prepared to sing in F (in which G and E are on the third tier of the hierarchy).

In a related vein, we consistently found key-distance effects associated with switches to a different initial note. Key-distance effects resulting from switching other subsequent notes in a melody might also be present. For example, in Experiment 1, Rising Melody 1 in F shared all five notes with the primary key (C major). Rising Melody 2 in F and Rising Melody 3 in F (which contained one instance of the note B \flat) each shared four notes with the primary key, but Rising Melody 4 in F (which contained two instances of the note B \flat) shared only three notes with the primary key. Lack of sufficient statistical power prevented us from examining whether these differences in the amount of overlap between secondary melody notes and the primary key affected response latencies, durations, and error rates. Future research could specifically investigate this possibility carefully by adopting more powerful experimental designs.

An additional important research direction will be to study the acquisition of skilled singing. We do not yet know to what extent the motor programs of novice singers are similar to those of expert singers. As one becomes an expert singer and sight reader, perhaps the content and/or organization of one's articulatory motor programs change. For example, novice singers may include a rough representation of melodic contour in their motor programs when sight-reading music. Indeed, such a rough representation might be the only kind available to them, which would explain why they make many errors in producing a melody. However, with continued practice, they could progress toward using other types of representations that enable exact transpositions, which would let them produce various melodies more accurately. The amounts and kinds of practice required for these progressions in skill level need to be explored more fully.

Likewise worthy of more exploration are possible individual differences among various types of expert singers. For example, in the context of our response-priming procedure, different results could emerge from skilled jazz vocalists compared with those from our present classically trained singers. Jazz vocalists, like jazz instrumentalists, must be highly adept at musical improvisation. They have to respond quickly, flexibly, and appropriately to spontaneous musical decisions made by other members of a musical ensemble. With this in mind, one might expect them to be especially facile at modifying and executing articulatory motor programs with melodic changes that feature different keys, initial notes, and/or contours. Thus, it would be interesting and informative to examine the response latencies, durations, and accuracies of

these experts versus those whose training puts less emphasis on musical spontaneity.

References

- Abrams, R. A., Meyer, D. E., & Kornblum, S. (1990). Eye-hand coordination: Oculomotor control in rapid aimed limb movements. *Journal of Experimental Psychology: Human Perception and Performance*, *16*, 248–267.
- Cadalbert, A., Landis, T., Regard, M., & Graves, R. E. (1994). Singing with and without words: Hemispheric asymmetries in motor control. *Journal of Clinical and Experimental Neuropsychology*, *16*, 664–670.
- Dell, G. S. (1986). A spreading-activation theory of retrieval in speech production. *Psychological Review*, *93*, 283–321.
- Deutsch, D. (1969). Music recognition. *Psychological Review*, *76*, 300–307.
- Deutsch, D. (2002). The puzzle of absolute pitch. *Psychological Science*, *11*, 200–204.
- Dowling, W. J. (1978). Scale and contour: Two components of a theory of memory for melodies. *Psychological Review*, *85*, 341–354.
- Dowling, W. J., & Fujitani, D. S. (1971). Contour, interval, and pitch recognition in memory for melodies. *Journal of the Acoustical Society of America*, *49*, 524–531.
- Garcia-Colera, A., & Semjen, A. (1988). Distributed planning of movement sequences. *Journal of Motor Behavior*, *20*, 341–367.
- Gordon, P. C., & Meyer, D. E. (1987). Control of serial order in rapidly spoken syllable sequences. *Journal of Memory and Language*, *26*, 300–321.
- Hays, W. L. (1994). *Statistics* (5th ed.). Fort Worth, TX: Harcourt Brace.
- Inhoff, A. W., Rosenbaum, D. A., Gordon, A. M., & Campbell, J. A. (1984). Stimulus-response compatibility and motor programming of manual response sequences. *Journal of Experimental Psychology: Human Perception and Performance*, *10*, 724–733.
- Ketelaars, M. A. C., Khan, M. A., & Franks, I. M. (1999). Dual-task interferences as an indicator of on-line programming in simple movement sequences. *Journal of Experimental Psychology: Human Perception and Performance*, *25*, 1302–1315.
- Klapp, S. T., & Wyatt, E. P. (1976). Motor programming within a sequence of responses. *Journal of Motor Behavior*, *8*, 19–26.
- Krumhansl, C. L. (1979). The psychological representation of musical pitch in a tonal context. *Cognitive Psychology*, *11*, 346–374.
- Krumhansl, C. L. (1985). Perceiving tonal structure in music. *American Scientist*, *73*, 371–378.
- Krumhansl, C. L. (1990). *Cognitive foundations of musical pitch*. New York: Oxford University Press.
- Krumhansl, C. L. (1991). Music psychology: Tonal structures in perception and memory. *Annual Review of Psychology*, *42*, 277–303.
- Krumhansl, C. L. (2000). Rhythm and pitch in music cognition. *Psychological Bulletin*, *126*, 159–179.
- Krumhansl, C. L., & Kessler, E. J. (1982). Tracing the dynamic changes in perceived tonal organization in a spatial representation of musical keys. *Psychological Review*, *89*, 334–368.
- Krumhansl, C. L., & Shepard, R. N. (1979). Quantification of the hierarchy of tonal functions within a diatonic context. *Journal of Experimental Psychology: Human Perception and Performance*, *5*, 579–594.
- Lenneberg, E. H. (1967). *Biological foundations of language*. New York: Wiley.
- Meyer, D. E., & Gordon, P. C. (1985). Speech production: Motor programming of phonetic features. *Journal of Memory and Language*, *24*, 3–26.
- O'Boyle, M. W., & Sanford, M. (1988). Hemispheric asymmetry in the matching of melodies to rhythm sequences tapped in the right and left palms. *Cortex*, *24*, 211–221.
- Pachella, R. G. (1974). The interpretation of reaction time in information

- processing. In B. H. Kantowitz (Ed.), *Human information processing: Tutorials in performance and cognition* (pp. 41–82). Hillsdale, NJ: Erlbaum.
- Palmer, C. (1989). Mapping musical thought to performance. *Journal of Experimental Psychology: Human Perception and Performance*, *15*, 331–346.
- Palmer, C. (1997). Music performance. *Annual Review of Psychology*, *48*, 115–138.
- Palmer, C., & Krumhansl, C. L. (1987). Independent temporal and pitch structures in determination of musical phrases. *Journal of Experimental Psychology: Human Perception and Performance*, *13*, 116–126.
- Patel, A. D., Gibson, E., Ratner, J., Besson, M., & Holcomb, P. J. (1998). Processing syntactic relations in language and music: An event-related potential study. *Journal of Cognitive Neuroscience*, *10*, 717–733.
- Rosenbaum, D. A. (1980). Human movement initiation: Specification of arm, direction, and extent. *Journal of Experimental Psychology: General*, *109*, 444–474.
- Rosenbaum, D. A. (1991). *Human motor control*. New York: Academic Press.
- Rosenbaum, D. A., Gordon, A. M., Stillings, N. A., & Feinstein, M. H. (1987). Stimulus-response compatibility in the programming of speech. *Memory & Cognition*, *15*, 217–224.
- Rosenbaum, D. A., Kenny, S. B., & Derr, M. A. (1983). Hierarchical control of rapid movement sequences. *Journal of Experimental Psychology: Human Perception and Performance*, *9*, 86–102.
- Rosenbaum, D. A., Weber, R. J., Hazelett, W. M., & Hindorff, V. (1986). The parameter remapping effect in human performance: Evidence from tongue twisters and finger fumblers. *Journal of Memory and Language*, *25*, 710–725.
- Saffran, J. R., & Griepentrog, G. J. (2001). Absolute pitch in infant auditory learning: Evidence for developmental reorganization. *Developmental Psychology*, *37*, 74–85.
- Smiley-Oyen, A. L., & Worringham, C. J. (1996). Distribution of programming in a rapid aimed sequential movement. *Quarterly Journal of Experimental Psychology: Human Experimental Psychology*, *49(A)*, 379–397.
- Sternberg, S., Monsell, S., Knoll, R. L., & Wright, C. E. (1978). The latency and duration of rapid movement sequences: Comparisons of speech and typewriting. In G. E. Stelmach (Ed.), *Information processing in motor control and learning* (pp. 117–152). New York: Academic Press.
- Todd, N. (1985). A model of expressive timing in tonal music. *Music Perception*, *3*, 33–57.
- Tramo, M. J., & Bharucha, J. J. (1991). Musical priming by the right hemisphere post-callosotomy. *Neuropsychologia*, *29*, 313–325.
- Yaniv, I., Meyer, D. E., Gordon, P. C., Huff, C. A., & Sevald, C. A. (1990). Vowel similarity, connectionist models, and syllable structure in motor programming of speech. *Journal of Memory and Language*, *29*, 1–26.
- Zatorre, R. J. (1984). Musical perception and cerebral function: A critical review. *Music Perception*, *2*, 196–221.

Received June 2, 2003

Revision received September 1, 2004

Accepted August 2, 2005 ■