Chapter Six: Conclusions

All four experiments in this dissertation showed evidence that trained musicians employ verbal encoding strategies when attempting to remember musical stimuli. Labels were shown to have an overall positive impact on musical memory, although some detrimental effects were also observed. In general, correctly named items were more likely to be recognized as targets and correctly rejected as lures than were incorrectly named items. Items that were not correctly named were especially likely to produce false alarms when presented as lures. Evidence of verbal encoding was observed among both men and women, but experiments 1, 2, and 4 demonstrated slightly stronger and more consistent verbal encoding effects among women, suggesting that women may be more inclined to adopt verbal encoding strategies.\footnote{The only comparison I noticed in which men appeared to be employing verbal strategies to a greater degree than women was in experiment 3. Men were significantly more likely to reject lures when they were able to identify at least four pitches in the standard series. However, women were also significantly more likely to reject lures when they were able to identify at least six pitches in the standard series. Furthermore, as mentioned in chapter 5, gender and solfège ability were unexpectedly confounded in this experiment, so it is impossible to ascribe these observations to gender alone.}

Comparison of experimental results with previous research

The results of experiment 1 are strikingly consistent with Bower & Holyoak (1973), Bartlett (1977), and Lawrence (1979). Instrumental timbres, like “environmental sounds,” appear to be retained as verbal interpretations, and the use of labels seems to improve recognition performance. Consistently named sounds were most likely to be recognized, and this was true regardless of whether the labels were provided by the listener or by the experimenter. Inaccurate labels were associated with chance (or even below-chance)
performance. The results of experiment 1 may also be seen as consistent with Ferrara, Puff, Gioia, & Richards (1978), who found that meaningful interpretations of sounds were more beneficial to memory than were evaluations of their physical properties. Ferrara et al. asked listeners either to judge the pleasantness of their associations with a sound or to determine in which ear the sound was louder. Listeners who subjectively interpreted the sounds had better recall than listeners who objectively evaluated their loudness.\(^2\)

As mentioned in experiment 1, I noticed that objective descriptions of unfamiliar timbres were apparently less effective as memory aids than were evocative — even whimsical — names that suggested particularly vivid subjective interpretations.\(^3\) It may be that feature-based descriptions of unfamiliar timbres were not sufficiently specific to distinguish between similar stimuli, and thus were more likely to produce verbal overshadowing.\(^4\)

Experiment 2 differed substantially from experiment 1 in two important respects: it addressed short-term memory rather than long-term memory, and it did not explicitly encourage verbal encoding. Despite the differences in design, the results were consistent with those of experiment 1: a large proportion of namable timbres improved performance, and the namability of the replaced timbre was a good predictor of a lure’s correct rejection. Thus, although this experiment did nothing to promote verbal encoding, it seems that participants preferred a verbal strategy.

\(^2\) In fact, listeners who interpreted the sounds performed similarly to a third group of listeners who had been explicitly instructed to memorize the sounds.

\(^3\) See chapter 2, footnote 11.

\(^4\) This would be consistent with research involving verbal labels for visual stimuli. See, for instance, Bartlett, Till, & Levy (1980).
Experiments 3 and 4 were closely related to Mikumo (1992) in experimental design, most notably in their use of different interference patterns (a feature shared by experiment 2). Both experiments 3 and 4 demonstrated a positive relationship between general solfège ability and overall performance, although we cannot assume causation (perhaps participants with better discrimination ability also tend to excel at solfège). In addition, experiments 3 and 4 showed a positive relationship between the specific ability to identify the changed notes in lures and correct rejection rates. Mikumo’s experiment did not investigate note-naming abilities specifically, but an interference pattern of spoken note names was shown to diminish music majors’ ability to recognize tonal melodies. However, almost no significant differences were observed among the various interference patterns used in my experiments 2-4.5

Interference patterns have been reported to produce a variety of effects. As discussed in chapter 1, Deutsch (1970) found that a series of pitches had a significant detrimental effect on pitch memory, whereas a series of spoken numbers had no such effect (even when participants were asked to recall lists of numbers). Pechmann & Mohr (1992) observed that a series of pitches and a series of spoken words both produced poor pitch memory performance among nonmusicians, but that musicians only experienced interference from the series of pitches. Mikumo found that spoken note names produced the lowest recognition rates among trained musicians, but no significant differences emerged for nonmusicians.

5 In experiment 1, scores obtained with the neutral words interference pattern were higher than those obtained with the random acoustic timbre interference pattern. No other main effects were observed in any of the three relevant experiments.
In light of these three previous experiments, why did the experiments in this dissertation find almost no significant differences involving interference patterns? Perhaps remembering a single pitch (as in Deutsch, 1970) typically involves different strategies from those used to remember a more complex stimulus like a melody. However, this does not explain why my results differ substantially from Mikumo’s. One possibility is that my neutral words interference pattern was not equivalent to Mikumo’s nonsense syllable interference pattern. Meaningful words may impact memory to a greater extent than do meaningless words, even if the meaningful words are not related to the stimulus itself. Because Mikumo did not indicate whether her participants had absolute pitch, it is impossible to determine the extent to which her note-name interference pattern was equivalent to my solfège interference pattern.\(^6\)

It is also possible that my non-verbal interference patterns triggered an unexpected verbal response. I believe that this is particularly likely in experiment 2. Participants who adopted the strategy of naming individual timbres in the standard and comparison series might also have tended to identify timbres in the interference pattern. This is one possible explanation why lower scores were obtained with the random acoustic timbre interference pattern than with the random electronic timbre interference pattern, since electronic sounds (without readily apparent labels) are less likely than easily namable sounds to trigger a labeling response.

Another possible reason that I did not find significant differences among the different interference patterns is that my pitch stimuli were less

\(^6\) Since Mikumo’s experiment was conducted in Japan, the note names were, in fact, spoken solfège syllables.
typically melodic than Mikumo’s, primarily because my stimuli did not imply a
traditional underlying harmonic progression, whereas her stimuli did, as the
examples below illustrate.\footnote{Mikumo describes her tonal stimuli has “high in tonal melodic structure,” and the examples
she supplies (which are presumably representative) strongly suggest a general tonic-dominant
alternation. (See Mikumo [1992], p. 75.)}

\begin{figure}
\centering
\includegraphics[width=0.7\textwidth]{mikumo_rogers.png}
\caption{Mikumo (left) vs. Rogers (right).}
\end{figure}

I chose to ignore harmonic considerations in my stimuli for several reasons. First, I believed that it would be advantageous to focus on relatively few specific kinds of verbal encoding (in this case, solfège and interval names). The addition of harmonic information would have compelled me to add a spoken harmony interference pattern (e.g., “tonic, dominant …”), significantly lengthening the experiment. Second, and more importantly, I felt that my lures would not be of equivalent difficulty were I to substitute stepwise motion for a leap or vice versa. Given the other constraints I imposed on my stimuli (e.g., a rising-then-falling contour, including the tonic in the first, fourth, and seventh positions), accommodating typical harmonic progressions and their associated voice-leading tendencies would have been impossible. When Mikumo used atonal stimuli under otherwise identical conditions, differences among the verbal and pitch interference patterns disappeared. Perhaps musicians are more likely to rely on sensory storage and non-verbal forms of encoding for stimuli that do not conform to traditional melodic expectations.
A final consideration is that Mikumo’s participants may have paid more attention to the interference patterns. Pilot studies for the experiments in this dissertation suggested that listeners routinely attempted to ignore interference patterns even when specifically instructed not to. The task of tallying repeated items in the interference pattern was added in order to encourage participants to pay attention, but this task may not have been sufficient. Mikumo indicates that participants in her experiment were asked to “shadow” the verbal interference patterns, thereby compelling them to pay attention. However, Mikumo did not ask participants to “shadow” the pitch interference pattern. This difference in instructions might explain why moderate scores were associated with the pitch interference pattern and why particularly low scores were obtained with the note-name interference pattern.

Experiment 1 demonstrated some effects that could be interpreted as similar to the verbal overshadowing observed by Schooler & Engstler-Schooler (1990) and Melcher & Schooler (1996) for non-auditory tasks. Data suggest that participants in the present experiment were reacting to their memories of verbal labels rather than to their memories of the timbres themselves. As long as these labels were accurate and/or consistent, they appeared to be helpful, but inaccurate labels led to poor performance (worse than providing no label whatsoever). Thus, although verbal labels seemed generally beneficial to performance in all four experiments, we should beware of the negative effects of verbal encoding of music even among trained musicians.

---

8 See Mikumo’s (1992) description of her experimental procedure, p. 76.

9 Recall that Melcher & Schooler (1996) concluded that verbal overshadowing seems to have an adverse effect on participants of moderate expertise but does not affect experts. Of course, one can
Suggestions for further research

The experiments in this dissertation demonstrate some benefits of verbal labels, and this is consistent with previous research in auditory memory. However, more research is needed to explore the role of verbal overshadowing in musical memory. Ideally, such experiments should involve participants of varying expertise. The inclusion of training sessions (similar to the training incorporated in experiment 1) would help to isolate verbal expertise as a controlled variable.

As discussed earlier, there are curious discrepancies among musical experiments involving interference patterns. These differences should be addressed systematically in an attempt to determine the conditions under which various types of interference patterns impact musical memory. One major challenge is compelling participants to pay attention to the interference patterns without biasing the results — that is, without treating the individual types of interference patterns differently (as Mikumo did).

One type of interference pattern that remains untested is a series of incorrect solfège syllables (or scale degrees) — solfège syllables sung to incorrect pitches.\(^\text{10}\) It would be appropriate to compare performance with incorrectly sung solfège to performance with correctly sung solfège, notes sung without solfège, and spoken solfège. For listeners with significant solfège training, I believe that incorrect solfège would be the most distracting of these four possibilities, both because it could be disorienting and because such an interference pattern would, in some sense, contain twice as many

\(^{10}\) For instance, in the key of C, E might be sung as sol, D as mi, B as fa, etc.
items (e.g., the actual pitches sung as well as the expected pitches that correspond to the solfège syllables).

The gender differences observed in this dissertation were intriguing and deserve further study. Female participants seem to have been more inclined than male participants to adopt a primarily verbal strategy. Does this truly reflect a gender difference in the general population, or were these listeners not adequately representative? It might be wise to ask participants to describe their primary strategy at the end of each experiment and, like Warden (1924), classify them as verbalizers, visualizers, etc. This would enable us both to look for differences among these groups and also to look for correlations between gender and preferred strategy.

Finally, of course, this dissertation only addresses verbal encoding for timbre and simple mostly diatonic pitch patterns. This research should be extended to a variety of different areas, such as more typical tonal melodies, atonal melodies, short polyphonic excerpts (representing a variety of familiar and unfamiliar musical styles), and rhythm. For example, an experimenter might:

• Train one group of listeners to name collections that are important in some twentieth-century music (e.g., whole-tone or octatonic scales). See how this group performs (compared to a control group that did not receive this preliminary training) on recognition, dictation, or sight-singing tests that include melodies based on the relevant collections.

• Play moderate-length harmonic progressions five times each, asking three groups of listeners to notate the progressions using different strategies. One group would produce standard musical notation after each repetition; a second group would identify features of the progression using words for the first two repetitions, then switch to standard musical notation; and the final group would write nothing at all for the first two repetitions, then use standard musical notation.
Implications of Results for Music Theory Pedagogy

The question of how music is represented in memory is a critical one for music pedagogy, and this dissertation grew out of an unusual pedagogical experience. I was in charge of a freshman aural skills program with nearly a hundred students, all of whom were simultaneously enrolled in an introductory written theory class taught by one of three other instructors. The various sections of written theory were not coordinated, and they differed substantially: topics were introduced at different times and in a different order, and homework assignments varied dramatically (for instance, some students wrote philosophical essays, some composed, and some did a lot of drill work).

After a couple of months, I noticed that students who were acquiring music-analytical vocabulary in their theory classes appeared to be making better progress in my aural skills class. This stood out to me not because I was particularly interested at the time in their developing vocabulary, but because these were not the students I had expected to be the most successful in my class. Theory classes were tracked, and these students had been placed into the middle sections, which were taught by someone who introduced terms whenever possible.\textsuperscript{11} This apparent correlation between vocabulary training and aural skills — particularly dictation skills — reminded me of experiments I had read suggesting that color codability impacts memory.\textsuperscript{12} It occurred to me that some of my students might be remembering the words for sounds in order

---

\textsuperscript{11} In many respects, this instructor and I used different pedagogical approaches, and I have no reason to believe that his class was generally better coordinated with mine than were the other theory classes.

\textsuperscript{12} Lenneberg & Roberts (1953) and Brown & Lenneberg (1954); see chapter 1.
to facilitate their memory of the sounds themselves. The idea seemed plausible, especially since verbal encoding can be very efficient and resistant to forgetting.

I didn’t know whether my aural skills students’ knowledge of terms was improving the efficiency of their musical encoding, or whether they were taking advantage of the relative longevity of verbal memory, or whether they were simply more conscious of musical patterns they could name and were therefore more likely to notice them in the first place. I just knew that these students were doing well while my students who still hadn’t been introduced to music-analytical terms were falling further and further behind. I began dedicating substantial class time to defining common musical terms and playing musical examples to illustrate each concept. My students who had been struggling showed noticeable improvement within a week. The sudden turnaround didn’t seem coincidental to me, but clearly further research and controlled experimentation was in order.

Those of us who teach music in the classroom necessarily rely on a great deal of verbal communication, so we need to be particularly aware of the effects of verbal encoding. As outlined in chapter 1, there is considerable evidence that verbal labels can enhance auditory memory (as well as visual, olfactory, and gustatory memory), at least under certain conditions. I therefore advocate reinforcing verbal and aural connections whenever possible. Suggestions include:

- Have students name conspicuous compositional features (e.g., cadence types, sequences, imitation, mode mixture) by ear
• Consistently use syllable systems (e.g., solfège or scale-degree numbers) for sight-singing.

• Before sight-singing a melody, perform a quick analysis (e.g., locate and name prominent patterns from tonic and dominant harmonies)

• Encourage students to describe progressions they hear in addition to writing them in standard musical notation (e.g., deceptive cadence in the middle, pre-dominant chord at the end of measure 2, chromaticism at the beginning)

When multiple labels are commonly applied to the same item, teachers should consider that labels appear to be most beneficial when they are meaningful and/or “catchy.” In some situations, it is advantageous to adopt both a formal and an informal term for the same feature — for instance, associating the word “amen” with “plagal cadence,” or “question” and “answer” with “antecedent” and “consequent.”

As observed in experiments 3 and 4, many listeners have a tendency to “correct” melodies that don’t conform to their expectations. From my experience, this is particularly common among weaker aural skills students. Although this problem could simply be caused by poor memory skills, I believe that it often stems from a deficiency in musical categories. For instance, students who are unaware of half cadences sometimes alter melodies so that every phrase leads to do, when in fact many phrases lead to re or ti.

Introducing the expression “half cadence” and pointing out where half cadences are likely to occur can greatly reduce such errors because students become conscious that re or ti can be made to sound like a strong melodic

---

13 Instructors should recognize that different systems reinforce different musical features and are likely to lead to the formation of different perceptual categories. For instance, scale-degree numbers equate the sixth scale degree in the major and minor modes, whereas movable-do solfège suggests that they are similar but not identical (la vs. le). Fixed-do solfège, as it is typically practiced, equates all pitches with the same letter name (e.g., D, D♯, and D♭). Movable systems foster the development of relative pitch, while fixed systems foster the development of absolute pitch.
goal. Similarly, I have found the expression “dependent transition” beneficial for students who have difficulty recognizing transitions that grow out of the first theme in sonata-form movements. Once students are aware of a musical pattern, they are better able to remember it and less likely to distort it into some other common musical pattern. An awareness of patterns is one of the great advantages of a well-developed musical vocabulary. The recognition of namable patterns enables listeners to chunk music effectively as well as offering an additional encoding possibility.

It would be inappropriate, however, to overwhelm students with an overabundance of terms all at once, since verbal encoding can also degrade memory. As described in chapter 1, Melcher and Schooler (1996) found that verbal descriptions tend to overshadow sensory memory when there is a marked discrepancy between perceptual expertise and verbal expertise. From my experience, most undergraduate music majors demonstrate exactly this kind of discrepancy between what they can hear and what they can express verbally. Introducing a large number of unfamiliar terms could therefore be detrimental to learning, particularly if these terms are not clearly defined or are not used consistently. As we have seen, inaccurate or inadequate labels are not helpful — and, indeed, possibly even harmful — to memory. If we do not clearly define the vocabulary we introduce, explicitly addressing both the connections and the distinctions among terms, we may be providing our students with vague labels that will lead to confusion rather than comprehension, and oversight rather than insight.
Summary

The experimental results in this dissertation suggest that trained musicians employ verbal strategies for remembering music, at least under certain circumstances. This is specifically relevant to researchers investigating verbal learning behavior and encoding abilities and inclinations. It is also of general interest to experimenters addressing sensory memory (especially auditory memory) because participants were shown to utilize verbal encoding even when not encouraged to do so. This suggests that covert verbalization may be a significant factor in experiments that never address the issue.

The experiments in this dissertation also have strong pedagogical implications, particularly regarding the relationship between traditional written theory classes and aural skills classes. Vocabulary training, it seems, may be an important, but often overlooked, component of an effective ear training program. This reminds theory teachers of the importance of coordinating written and aural skills — an issue that involves much more than similar content and pacing. To me, truly coordinating written theory and aural skills means carefully and deliberately linking concepts and techniques in order to promote the development of effective cognitive strategies. We must choose a syllable system that supports our concept of musical building blocks, develop vocabulary that leads to helpful perceptual categories, engage in activities that promote the clear understanding and habitual use of these categories, and demonstrate how students can continue to refine and build upon this work long after they have left our classrooms.