We have a very vivid experience of music we hear for the first time. However, a growing body of convergent research shows that our memory for it varies within the first minute after our experience, and that the memories of nonmusicians, moderately trained musicians, and professional musicians differ. All three groups hear familiar melodies in terms of scale-step values. However, only listeners with musical training automatically hear the pitches of unfamiliar melodies as scale steps, and that encoding takes time of the order of 10 s after first hearing. When tested 4-5 s after hearing a melody followed by continuing music, both nonmusicians and moderate musicians confuse targets with lures in which the melody has been shifted to a different place on the scale (with changes in pitch-to-pitch intervals). If that test is delayed another 10 s, both groups reject the same-contour lures better, and that performance is especially good with the moderate musicians. This result converges with the failure of both groups to notice out-of-key wrong notes in unfamiliar melodies; even if they were encoding scale-step values, that process is too slow to provide a timely response time to an out-of-key note. Furthermore, noticing a tonal modulation supposes the tracking of relative strength of scale-step values moment to moment. When listeners track modulations in both Western and South Indian melodies using the continuous-probe-tone paradigm, nonmusicians’ tonal profiles generally fail to show the effects of the modulations. Moderate musicians show them, and professionals show them even more strongly. But even for professionals, the shift in the tonal profiles continues to grow even after the first 15 s of the modulation. Taken together, this evidence suggests that even professionals will remember hearing a piece with firmly grounded tonal structure, even though their initial immediate experience of the piece lacked that structure.