

VERY LONG TERM MEMORY FOR WORDS AND MELODY: AN EXPERT SINGER'S WRITTEN AND SUNG RECALL OVER SIX YEARS

Jane Ginsborg¹ and Roger Chaffin²

¹Royal Northern College of Music, Manchester, UK; ²University of Connecticut, USA

ABSTRACT

This study forms part of a longitudinal case study of a singer's preparation for performance and long-term recall of a piece of music, the first *Ricercar* from Stravinsky's *Cantata*. At the inaugural ICoMCS we reported written recall over three years; we now report findings based on data from two further written recalls and a sung recall. Accuracy of written recall declined steadily over time from 97% just before the public performance in December 2003 to 66% in November 2007, before improving slightly a year later to 68%. Sung recall was carried out at the end of July 2009 and was 89% accurate without, and 84% with accompaniment. The evidence suggests that while the decrease in accuracy of written recall is more gradual than the typical J-shaped forgetting curve, very-long term recall of music by a performer follows the same trajectory of decline as other kinds of material. The findings are discussed in relation to features noted by the singer during practice and rehearsal, and those that became performance cues, on which she drew when recalling the words and melody from memory.

1. INTRODUCTION

In the present study, we examined a singer and conductor preparing Stravinsky's *Ricercar 1* for soprano and small instrumental ensemble for performance. We have shown that the singer attended to certain *features* of the music during practice and performance, that a subset of these features, or *performance cues*, functioned as retrieval cues when she performed from memory (Ginsborg et al., 2006) and that these cues influenced her memory of the piece when she wrote it out from memory (Ginsborg & Chaffin, 2007). Writing out the score from memory was a normal practice activity for the singer. For the purposes of the study, she wrote it out repeatedly at intervals averaging 8.3 months over a 5-year period, providing an opportunity to observe the effects of features and performance cues on long-term retention. In the sixth year she provided two sung recalls, the first without, and the second immediately afterwards with accompaniment, for comparison with the written recalls.

Recall from long-term memory has been studied by psychologists ever since Ebbinghaus (1885/1993) pioneered its investigation using himself as his experimental subject. Retention has normally been studied, however, over relatively short periods – measured in minutes, hours or days. Relatively few studies have examined retention over months or years, although exceptions include Bahrck (1994), Rubin and Wenzel (1996) and Stevens et al. (2009). In the study that provides

the model for the present study of long-term musical memory (Chaffin et al., 2002), the pianist Gabriela Imreh prepared Bach's *Italian Concerto* for performance in 33 hours of practice. Two years later she wrote down as much as she could remember of the first page of the score, recalling around 65% of the notes. Her free recall was better at section boundaries and at expressive performance cues and declined with each successive bar. Basic performance cues (relating to fingering and technique, for example), in contrast, produced a serial position effect in the opposite direction: recall was *poorest* at basic cues and improved with distance from the cue. Similar serial position effects were also found in a study with a cellist (Chaffin, Logan & Begosh, 2009). Thus some retrieval cues can be thought of as *landmarks*, providing the musicians with direct, content addressable access to their memory of the music. Others, e.g. where the musician had to pay particular attention to some aspect of the sensori-motor context, such as basic performance cues, can be thought of as *lacunae*, since recall improved thereafter.

The aim of the present study was to describe the singer's declining accuracy of recall over a period of nearly six years, to compare written recalls in the first five years with sung recall in the sixth and to relate written recall to features and performance cues serving as landmarks and lacunae.

2. METHOD

Participants: Jane Ginsborg, the first author, is a former professional singer; she has worked with the pianist and conductor George Nicholson for more than 30 years, performing as a duo and as members of a variety of ensembles.

Materials: Stravinsky's *Cantata* for two solo singers, women's choir and small instrumental ensemble includes one movement for solo soprano and ensemble, *Ricercar 1* (circa 4 minutes in length). The present study investigated the singer's preparation and performance of the *Ricercar* only.

Procedure: From mid-November to mid-December 2003, the singer undertook five individual practice sessions lasting 4 hours 13 minutes in all. She carried out four joint rehearsals lasting 2 hours 47 minutes with the conductor. These nine practice sessions and rehearsals (although not three ensemble rehearsals lasting 57 minutes) were recorded and analysed. A public performance of the complete *Cantata*, conducted by George Nicholson, with the first author as solo soprano, was given on 16 December 2003. The features she noted during practice and rehearsal, and the subset serving as performance cues, were identified from annotations made on multiple copies of the score

immediately after the performance. The singer wrote out one free recall during the preparation period, as was her custom, to check that her memory would be secure. She gave four complete performances from memory, and made several further written free recalls for research purposes. The first (FR0) was made between the last two rehearsal sessions, five days before the performance. In the course of the penultimate rehearsal the singer and conductor, playing the piano, gave two uninterrupted performances of the piece. The singer made one error in one performance and two errors in the other. The final rehearsal included an uninterrupted performance in which the singer accommodated to two errors made by the conductor. The public performance was accurate in all respects.

The singer made eight further free recalls after the public performance, writing down what she could remember of the words and melody, notating rhythms above each word, and humming, beating a pulse and conducting as necessary until she had worked through the whole song from start to end. Her recalls at the end of January 2004 and the end of February, 2004 yielded only one or two trivial errors (these data are not reported). The first time after the performance that the singer made a substantial number of errors was 12 months later, when she recalled the piece in February 2005 (FR1); five further recalls were made in June 2005, August 2006, June 2007, November 2007 and November 2008 (FR2-6). Each was made after a period of months of not thinking about the piece, before resuming work on the study. Apart from FR0, recalls occurred 12, 18, 32, 42, 47 and 59 months after the public performance. The time intervals since last consulting the score were 10, 4, 10, 6, 5, and 4 months respectively for FR1, FR2, FR3, FR4, FR5 and FR6. While earlier recalls involved some element of reconstruction, in FR3-6 the singer did not attempt to reconstruct the piece but simply worked through it once from start to finish. FR7 and 8 were sung recalls made on the same day in July 2009, first without and then with piano accompaniment.

3. RESULTS

Each quaver beat was scored for accuracy of recall of word, pitch (FR1-8; pitches were not recorded in FR0), rhythm/duration and omission. Omitted beats were scored 0; perfectly recalled beats were scored 1, with scores in between reflecting the number of errors (1-4). Whole-beat rests were scored 1 (in the written recalls) if notated, as 0 if omitted. Accuracy of written recall declined steadily over time from 97% (FR1) to a low of 66% 47 months later (FR5; see Figure 1). It rose to 89% in the sung recall without accompaniment 67 months after the public performance (FR7) and 84% with accompaniment (FR8).

The decrease is more gradual than the typical J-shaped forgetting curve (Ebbinghaus, 1885/1993). Two factors may account for this. First, FR0 occurred as part of the singer's preparation for performance. Sung recall had not yet reached asymptote, which presumably occurred five days later when the singer performed with perfect accuracy from memory. Second, the singer's intermittent work on the research undoubtedly slowed forgetting. Despite these factors, written recall followed close to the expected course, levelling off at FR5, almost four years after the performance. To our knowledge, this is the first evidence that very-long term recall of music by a performer follows

the same trajectory of decline as other kinds of material (Bahrick, 1994).

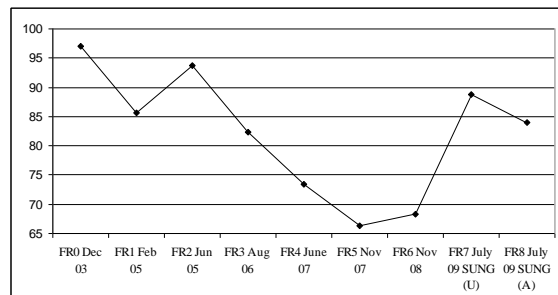


Figure 1: Accuracy of written and sung recall over time

Figure 2 shows the locations of errors in each of the seven written free recalls. The X-axes represent beats of the piece from 1 to 250. The Y-axes represent the degrees of accuracy with 1 representing complete accuracy and 0 representing complete omission.

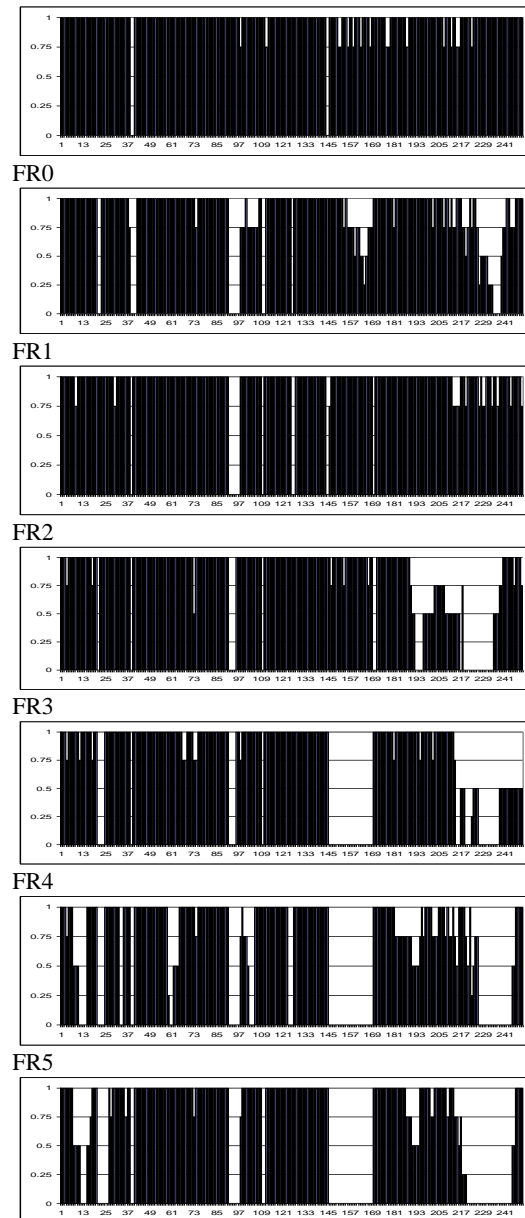


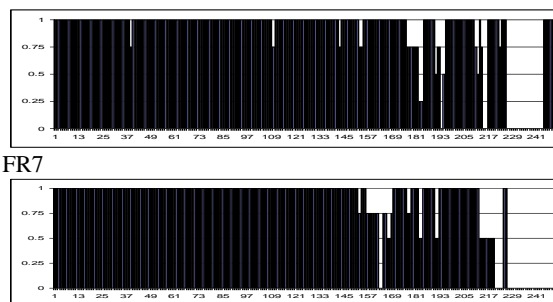
Figure 2: Locations of errors in written recalls

As shown in the top panel of Figure 2, the singer was able to notate the words and rhythms with 97% accuracy five days before the performance (FR0) and there were no gaps where it was forgotten entirely. The majority of errors concerned the recall and notation of rhythms. Errors involving the words included reversal and substitution. Twelve months later (FR1) there was a substantial drop in accuracy, to 86%. Most errors involved rhythm and duration, although the rests after the second appearance of the refrain were omitted (beats 92-97), as was the phrase *A place e-* (beats 235-238). After four months, 18 months after the performance (FR2), accuracy increased to 94%. Again, rhythm/duration errors predominated and the rests before the third appearance of the refrain were omitted. There was a further substantial drop in accuracy to 82% after 14 months, 32 months after the performance (FR3). As well as the omitted rests (beats 92-97), the passage *and eke vic-* (beats 193-196) was forgotten. The words *Vertuous and benign* (beats 203-208) were forgotten, although the melody was recalled at first; rhythms were forgotten at *Let us, lett us pray all, all to* (beats 209-216) before all was omitted between beats 219-234 (*Eternal Which is the heavenly King After ther liff grant them*). Recall continued to deteriorate. In FR4, ten months later and 42 months after the performance, accuracy dropped to 74% as the whole of the section following the third appearance of the refrain was forgotten: *For to report it now were tedious: We will therfor now sing no more Of the games joyus* (beats 145-172). The majority of errors were durations, forgetting the melody but preserving the words or omitting words and melody simultaneously. There was one pitch error that had not occurred before (*All*, beat 214), after which the next three words were forgotten (preserving the melody), and then the subsequent melody (preserving the words) before both words and melody were forgotten altogether, as before: *After ther liff grant them A place* (beats 227-237). The words *eternally to sing* (beats 238-244) were forgotten, although not the melody, and the final *Amen* was recalled only with rhythmic errors.

The next recall (FR5), five months later and 47 months after the performance, was substantially poorer. Ten passages were omitted entirely and the remainder was recalled with only 66% accuracy. Portions of the very first line were omitted along with the section following the third appearance of the refrain (as in FR4) and the words of *Our quen princis* (beats 182-189), as well as the closing section starting *After ther liff* (beats 227-244). The majority of errors involved forgetting either words or melody while the other was preserved. Many of the same passages were forgotten in FR6, a year later and 59 months after the performance. Accuracy of recall improved slightly, to 68%, including recall for two previously forgotten phrases, *And through the glass window shines the sone* (beats 98-109) and *Our quen princis* (beats 182-189).

Figure 3 shows the errors in the two sung recalls, made on the same occasion in July 2009, first without and then with piano accompaniment. Recall was much more accurate when sung than when written: 89% without and 84% with accompaniment, although in the unaccompanied sung recall the beats 145-172 were omitted altogether and the final section following the pitch error first made at beat 214 in FR4 was not retrievable, other than the closing *Amen*. The accompanied sung recall was even more confident and accurate than the unaccompanied sung recall until the singer reached beats 191-192, where she found it impossible to retrieve the correct underlay for *Prepotent*.

While she was able to resume after a break and continue successfully as far as *Vertuous and benign*, the attempt to recall the remainder of the piece had to be abandoned when the barrier presented by the pitch error at beat 214 proved insurmountable.



FR8
Figure 3: Locations of errors in the sung recalls

Landmarks were indicated by serial position effects in the probability of accurate recall (see Chaffin, Ginsborg & Dixon, 2009). Here we report the effects across FR1 to 6 (FR0 was not included because the small number of errors precluded the possibility of systematic effects). There were no clear lacunae. Starts of sections and phrases functioned as typical landmarks, while performance cues relating to technical issues and words (pronunciation, meaning, stress) were also landmarks, but produced slightly different patterns of recall probability.

Starts of sections; starts of phrases: the probability of recall increased to its highest level as each section and phrase boundary was reached and then declined in each subsequent bar to its lowest level at serial position 5. This final serial position also included all subsequent serial positions up to the beginning of the next phrase, thus maintaining a similar number of observations at each serial position. Mixed hierarchical regression analysis showed that the linear effect was reliable (sections: estimate = -0.049, $z = -3.41$, $p = .001$; phrases: estimate = -0.14, $z = -2.79$, $p = .005$).

Technical performance cues: As at starts of sections and phrases, the probability of recall was highest at the cue itself, but dipped and improved both beforehand and immediately afterwards, producing a W-shaped curve (estimate = 0.027, $z = 1.93$, $p = .05$).

Word performance cues: Again, the probability of recall was highest at the cue itself, but there was a double dip similar to that shown by technical performance cues (estimate = 0.028, $z = 2.09$, $p = .04$).

4. DISCUSSION

Like other experienced soloists who have been studied, the singer engaged in extended practice to ensure that recall occurred with the rapidity and fluency needed for performance. The singer had the piece memorized by Session 3, in which she sang 93% of the practice segments from memory. She then continued to practise mainly from memory. In the process, she developed the automaticity and reliability of memory retrieval needed for a secure performance. When she checked her memory by writing out the score for the first time (FR0), five days before the performance, there were no gaps; her memory was 97% accurate. As she progressively forgot the piece in the

months and years that followed – despite regularly revisiting it in the course of undertaking this research – the increasing gaps in subsequent recalls revealed how the music was organized in her memory.

Sung recall was much more accurate than written recall, even after a year-long gap. Similar findings are reported by Lisboa, Chaffin and Logan (2009) who attribute them to the role of auditory and motor memory when playing. In the present study, we would argue that the Stravinsky piece presents particular notational difficulties, with bars of irregular length, frequent use of demi-semi- and hemi-demi-semiquavers. The use of accompaniment for the second sung recall enhanced accuracy and fluency (bearing in mind that the singer had just rehearsed without accompaniment) since it provided auditory cues that had otherwise to be imagined, whether in terms of melody or mere beats to be counted; these cues, however, could not help the singer remember the closing passages since they had been comprehensively forgotten.

Serial position effects in the analysis of written recalls (sung recalls remain to be analysed) identify the location of landmarks in the singer's mental map of the piece. Probability of recall increased as she approached section and phrase boundaries. It declined with distance from the boundary, and at locations where she had addressed technical issues, and the pronunciation and meaning of, and emphasis on words to the extent that during performance she paid them conscious attention. This contrasts with the findings from the studies with the pianist and cellist, since the singer's technical performance cues (equivalent to some aspects of the instrumentalists' basic performance cues) did not function as lacunae. In the past we have suggested that landmarks are recalled better because they can be directly accessed by their content when retrieval falters. Serial cuing of subsequent bars then produces the characteristic negative serial position effect as the probabilities of forgetting at each beat accumulate with increasing distance from the cue. Another explanation is that they receive more attention during practice. In this case, recall should decrease with distance both before and after the cue, as we have seen in the present study.

It may be, however, that both explanations are valid. To perform reliably from memory requires that memory retrieval be practised. Musicians sometimes find themselves in the position of having to perform from memory without adequate preparation. In these cases, they must rely on serial cuing. Often they get away with it. But most musicians prefer to have a "safety net". Performance cues providing content addressable access to memory provide a backup, just in case. If things go wrong, the musician can jump forward to the next cue and avoid the ignominy of having to go back and start over. The singer in the present study, following her normal practice, tested her safety net by writing out the score from memory five days before the performance. Although most performers may not test their memories so thoroughly, we believe that most do set up performance cues and practise their use.

Performance cues differ, of course, as a function of the piece, musical style, instrument, experience of the musician, and demands of the particular occasion. Musicians may also differ in the extent to which they establish content addressable cues and/or rely on serial cuing. However, the studies of professional performers to

date suggest that every musician makes some use of both. This generalization is also suggested by the consistency of musicians' use of musical structure and performance cues, as evidenced in case studies, with the principles of memory developed from the study of experts in other fields, and the general population (Ericsson & Kintsch, 1995). There is good reason to expect, therefore, that the same principles generalize to other experienced performers.

5. REFERENCES

1. Bahrick, H. P. (1994). Extending the lifespan of knowledge. In L. Penner, G. Batsche, H. Khoff, & D. Nelson (Eds.), *The challenge in mathematics and science education: Psychology's response*. Washington, DC: American Psychological Association.
2. Chaffin, R., & Imreh, G. (2002). Practicing perfection: Piano performance as expert memory, *Psychological Science*, 13, 342-349.
3. Chaffin, R., Ginsborg, J. & Dixon, J. (2009). Serial position effects in a singer's long term recall identify landmarks and lacunae in memory. In A. Williamon et al. (Eds.), *Proceedings of the International Symposium on Performance Science*, Association of European Conservatoires, Auckland, December.
4. Chaffin, R., Logan, T.R., & Begosh, K.T. (2009). Performing from memory. In S. Hallam, I. Cross, & M. Thaut (Eds.), *The Oxford Handbook of Music Psychology* (pp. 352-363). Oxford: OUP.
5. Ebbinghaus, H. (1885, tr. 1913). *Memory: A contribution to experimental psychology*. New York: Teachers College, Columbia University.
6. Ericsson, K. A., & Kintsch, W. (1995). Long-term working memory. *Psychological Review*, 102(2), 211-245.
7. Ginsborg, J. and Chaffin, R. (2007). The effect of retrieval cues developed during practice and rehearsal on an expert singer's long-term recall for words and melody. In A. Williamon and D. Coimbra (Eds.), *Proceedings of the International Symposium on Performance Science*, Association of European Conservatoires, Porto, November.
8. Ginsborg, J., Chaffin, R. and Nicholson, G. (2006). Shared performance cues: Predictors of expert individual practice and ensemble rehearsal. In M. Baroni et al. (Eds.) *Proceedings of the 9th International Conference on Music Perception and Cognition*, Bologna, Aug 22-26, 2006.
9. Lisboa, T., Chaffin, R. and Logan, T. Memory for music performance: Comparing played and written recall. In A. Williamon et al. (Eds.), *Proceedings of the International Symposium on Performance Science*, Association of European Conservatoires, Auckland, December.
10. Rubin, D. C., & Wenzel, A. E. (1996). One hundred years of forgetting: A quantitative description of retention. *Psychological Review*, 103 (4), 734-760.

11. Stevens, C., Ginsborg, J., & Lester, G. (2009). Moving backwards and forwards in time: Recalling dance movement from long-term memory. In A. Williamon et al. (Eds.), *Proceedings of the International Symposium on Performance Science*, Association of European Conservatoires, Auckland, December.