DURMS
Durham Undergraduate Research in Music & Science

Gregorian Chant and Verbal Learning

Emma Bonnard
Durham University

ABSTRACT

The purpose of this study was to look at the effect of music on verbal learning, and more specifically, at whether people are more likely to remember a set of words when it is spoken or when it is sung. It was hypothesized that music would enhance recall. An experiment was conducted with 23 university-aged native English speakers, who were made to listen to short recordings of verses by Yeats, both in sung and spoken form, and then to write down, in order, what words they remembered. The results did not support the hypothesis, as participants did not recall the sung versions better, regardless of familiarity.

1. INTRODUCTION

It is generally thought that music can make it easier to learn new linguistic content; however, it is still unclear to what extent this is true. In this study, I looked at the effect of music on verbal learning, and more specifically, at whether people are more likely to remember a text when it is spoken or when it is sung. This is relevant to the issue of liturgy in religious services, where setting a text to music is often expected to enhance understanding and recall.

In the context of my research, I looked at four key studies by Rainey and Larson (2002), Thaut et al. (2008), Moussard et al. (2012), and Tamminen et al. (2017).

Rainey and Larson (2002). The first study by Rainey and Larson (2002) aimed to test whether music would enable students to 1) learn lists of unconnected texts in fewer trials, and then 2) retain that information better in long-term memory.

There were two iterations of the experiment, as results in the first one were skewed due to a few names being confusing to participants on the list used. There were 102 participants in the second iteration of the experiment, tested individually. They learned a set of fourteen nonsense names (derived from the names of dwarves in Tolkien's *The Lord of the Rings*) by hearing them being either sung or spoken, and were then asked to recall them aloud in the original order, repeating this process until they made no mistakes. What was measured in this case was the number of trials necessary to reach perfect recall. Then, a week later, they were made to relearn the lists using the same process.

There was no support for the first hypothesis (that the song version would require less trials to learn); however, there was consistent support for the second hypothesis, as participants in the sung condition recalled "the list of names better a week later than those in the spoken condition, as measured by the number of trials required to relearn the lists." (Rainey and Larson, 2002, p. 183).

This study had the advantage of a relatively large sample size, and it is good that the experiment was repeated in an improved form. However, results might have been affected by the fact that it was not possible to control whether the participants rehearsed the list over the week in between the two phases (with the sung group more likely to find themselves rehearsing the list since it was set to a familiar tune). It is also only relevant to learning lists of unrelated words, as opposed to texts where the words come together to create meaning, such as in psalms.

Thaut et al. (2008). The aim of this second study was to investigate whether music would improve learning and memory of lists of ordered words in patients with multiple sclerosis.

Participants were 20 volunteers with relapsing-remitting multiple sclerosis, who were randomly assigned to one of two conditions: 1) with a musical template providing temporal structure, and 2) without any such template. The experiment exclusively used recordings and kept participants isolated in a soundproofed booth, to make the test procedures as consistent as possible. The same list of 15 semantically-unrelated words was used over the ten trials, during which participants were asked, after each list presentation, to free recall as many words as they could. Recall of the original list was tested after participants were first made to recall a distractor list, without hearing the original list again, and then again after performing a 20-minute nonverbal distractor task (Thaut et al., 2008, p. 326). Participants presented with the sung version were asked to recall it by singing it back in the original order, as in the study by Rainey and Larson (2002).

Researchers found a trend for better recall in music compared to spoken learning, with musical verbal learning most improving word order recall in the early and late phases of learning, while spoken verbal learning was more effective during the middle phase of learning. The study suggests that musical verbal learning may act by accessing compensatory pathways, as well as "confer[ing] a neurophysiological advantage through the stronger synchronization of the same neuronal cell assemblies underlying conventional verbal learning and memory." (Thaut et al., 2008, p. 328). Such considerations, however, fall outside the reach of the present study.

Moussard et al. (2012). This is a case study of a 68-year-old woman suffering from Alzheimer's, taking place over a few months, and its goal was to assess "whether new lyrics are

better learned and memorized when presented in a spoken or sung form" (Moussard et al., 2012, p. 521).

The participant was made to learn four lyric excerpts, taken from the repertoire of a popular French-Canadian folk-singer and songwriter, in four conditions (spoken, sung in an unfamiliar melody, sung in an unfamiliar melody previously learned, and sung in a familiar melody) and then recall them after a 10-minute forgetting phase. There was then a relearning phase which took place in 10 sessions over the next few weeks, in which the participant had to relearn the excerpts.

In the relearning phase and in delayed recall, the participant recalled the lyrics better in the sung version, and so the experimenter concluded that "music is an aid for verbal long-term retention" (Moussard et al., 2012, p. 528), suggesting that this might be due to the emotional power of the medium.

This study's literature review points out that performance appears to be related to melodic characteristics and that it has been shown that "the melody associated with the lyrics to be memorized has to be simple, symmetric, and repeated across the successive lines of the song". (Moussard et al., 2012, p. 522) This was a point I kept in mind when deciding which specific Gregorian tone to use in my experiment.

Tamminen et al. (2017). This final study was particularly relevant to the present study in that it used healthy participants. It was composed of three experiments. The first experiment required participants to learn new spoken words, created for the purposes of the experiment, and served to establish that the typical pattern of recall could be obtained through the task which the experimenters had developed. The second experiment was identical to the first except that words were presented in the sung modality, using an unfamiliar hymn tune; the results showed no evidence that music accelerated lexical integration. The third experiment used a familiar melody, but once again, results did not show any significant improvement in recall; hence, the hypothesis, based on the findings of previous studies, that music would enhance recall, especially when the music was already familiar, was not supported by the study's results. However, the discussion section acknowledges that the study's ecological validity was limited due to the tasks' lack of similarity to the way songs are normally learned, and suggests that incorporating dimensions like rhythm and phrasing might have made a difference. This was taken into consideration in the present study.

Hypothesis. Based on the findings of the above studies, I hypothesized that participants would do slightly better in the sung condition, and that this would be most noticeable in participants already familiar with the melody.

2. METHOD

Design. The experiment was conducted in the form of an online survey, using Bristol Surveys.

Participants. Participants were all university-aged native English speakers; this was so that age or language ability would not be factors in the experiment. They were recruited from acquaintances, and contacted through social media. In total 23 participants took part in the study, and were divided into two groups (see Table 1); Group A consisted of 11 participants, with 12 allocated to Group B.

Materials/Stimuli. The text used consisted of the following 8 lines from Yeats's poem 'Dialogue of Mind and Soul':

"*I am content to follow to its source*
Every event in action or in thought;
Measure the lot; forgive myself the lot!
When such as I cast out remorse /
So great a sweetness flows into the breast
We must laugh and we must sing,
We are blest by everything,
Everything we look upon is blest."

The lines were split into two units of four lines each. The section from "I am" to "remorse" was labelled "Stanza 1", and the remaining four lines were labelled "Stanza 2".

This text was chosen for its similarity to a psalm in terms of line length and semantic range, as well as for its relative obscurity. These specific lines were chosen for their lack of any words that might fall outside the normal vocabulary of the average university student, since the purpose of the experiment was to test recall, not vocabulary. I avoided choosing a standard Christian text in case this might skew results in favour of Christian participants, who would be more familiar with the vocabulary used, and to keep the experiment more emotionally-neutral.

Both Sections were recorded both in sung and spoken versions. In the spoken version, the text was set to the ninth tone, or "*tonus peregrinus*", because it was simple and fit the text well (Figure 1).

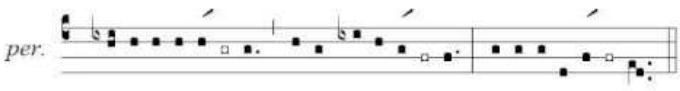

Figure 1. Tone 9

I recruited the cantor of the local Catholic church to record both the sung and spoken versions, so that the Gregorian chanting in the sung versions might be idiomatic, and so that the results would not be affected by using different speakers.

Procedure. The participants first had to read the information sheet (see Appendix), and provide their consent. Then, they were asked to listen to a recording of Stanza 1 "**three times** in close succession, while trying to memorise as much of the text as possible" (The words 'three times' were bolded so that there would be no risk of accidentally overlooking it). The recording was a google link. They were then instructed: "Write down as much of the text as you can remember in the

box below". These steps were then repeated with the second stanza. Group A listened to Stanza 1 spoken and Stanza 2 sung, while Group B listened to Stanza 1 sung, and Stanza 2 spoken (Table 1).

Table 1. Repartition of Recordings

	Stanza 1	Stanza 2
Group A	Spoken	Sung
Group B	Sung	Spoken

This was to minimize the effect of any disparity in the memorability of the two stanzas. Both groups listened to the stanzas in the same order (i.e. the same order as in the poem).

After they had finished the listening tasks, participants were asked to answer two follow-up questions:

1. Were you already familiar with the text used?
2. Were you already familiar with the melody used in the sung version?

The purpose of the first question was to enable the exclusion of any participants who were already familiar with the text. The purpose of the second question was to consider the claims of previous studies about the positive effect of familiarity with the music on recall, while looking at results.

3. RESULTS

Various methods of looking at and assessing the results were considered. One of them was 'Diff', which is a tool frequently used in computer science to compare text files. It would have been possible to count the number of differences either per-word, per-character, per-line or according to some more advanced algorithm (which, for instance, considered entirely omitted lines). Equally, a specialized piece of software could have been written in a general-purpose scripting language like Python or R (both commonly used in statistical studies). However, neither of these would have allowed for misspellings or misheard words, which is why they were ultimately not used.

Instead, responses were assessed by counting the number of words correctly recalled. Words were assessed independently (the correctness of one word was not dependent on that of the word before); however, the word order still had to be correct. For example, when the line "every event in action or in thought" was recalled as "every event and every thought", the words "every", "event" and "thought" were marked as correct. However, if it had been recalled as "every thought and every event", only the word "every" would have been marked as correct, due to "thought" and "event" being in the wrong order.

After looking over the results and noticing a few recurring cases of "misheard words", it was decided to allow for them so long as they were plausible and occurred in the right places. The most common one was "laugh" being misheard several times as "love"; due to context and the speaker's accent, the two words were not easily distinguishable, and since the purpose of the experiment was to test memory, not understanding, it would not have been helpful to mark it as a mistake. This does not mean, however, that the text was not intelligible from the recordings, as two participants recalled both stanzas perfectly.

The number of correct words was divided by the total number of words, and made into a percentage (see Table 2 and Figure 2). This was so that the results for both stanzas could be assessed side by side, despite the stanzas not being of the same word length.

Table 2. Range, Mean, Median, Variance and Standard Deviation of the Data (% = % correct)

	Range (%)	Mean (%)	Median (%)	Variance	Standard deviation
Stanza 1 (read)	21-100	53	48	413.97	20.35
Stanza 1 (sung)	38-93	55	52	247.74	15.74
Stanza 2 (read)	50-100	73	67	305.22	17.47
Stanza 2 (sung)	31-100	58	62	311.87	17.66

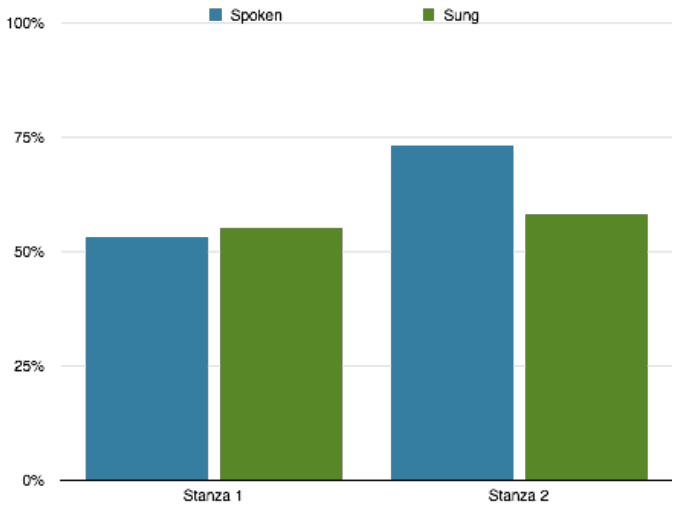

Figure 2. Mean percentage of text correctly recalled

Since the two groups were unequal, an unequal Variance t-test (Welch's t-test) was used to compare their performance:

Stanza 1

t = 0.26, p = .79

Stanza 2

t = -2.05, p = .03

Qualitative observations. Five participants who responded 'yes' to 'whether they were familiar with the melody used' also misremembered the word "Lord" as being part of the text, with one participant recalling the third line of Stanza 1 as "Lord something something". Due to the placement of the word in all cases and the clear difference in vowels, this does not appear to have been the result of a confusion with the word "lot".

Table 3. Individual Performance (with green indicating participants already familiar with the melody used)

Participant	Spoken (%)	Sung (%)	Difference (%)
1	100	100	0
2	21	31	+10
3	55	58	+3
4	48	46	-2
5	48	38	-10
6	41	46	+5
7	55	62	+7
8	48	62	+14
9	62	69	+7
10	31	65	+34
11	76	65	-11
12	100	59	-41
13	100	93	-7
14	69	52	-17
15	50	38	-12
16	62	55	-7
17	81	48	-33
18	58	41	-17
19	77	48	-29
20	54	52	-2
21	65	62	-3
22	65	38	-27
23	100	79	-21

Note: Participants 1-11 belong to Group A, and 12-23 belong to Group B.

4. DISCUSSION

There was no significant correlation between method of delivery and successful memorization, nor between familiarity with the music and recall in the sung version. Out of the 23 participants, 9 were already familiar with the Gregorian tone used; their average successful recall of the sung stanza (55%) did not significantly differ from the average recall of the group which was unfamiliar with the tone (58%). This might have been because the sample size was too small, and the variation in individual performance too great, to yield significant results. However, even the individual participants who were familiar with the melody did not recall the text better in the sung condition; indeed, some of the lowest scores in the sung condition were by participants familiar with the melody (See Table 3). The lack of correlation is consistent with the findings of previous studies, in which music only enhanced recall in the long-term, but not the short-term. A follow-up study could take place over multiple weeks, and look at whether Gregorian chant enhanced recall in the long-term. This, however, would make it only relevant to those parts of the liturgy which remain the same from week to week, but no longer, for example, to the Sunday Mass psalm.

It was interesting to note that not only did recall ability greatly vary between participants, but so did methods of recall. Some participants seemed to chiefly remember contextual meaning, so that in recall they replaced words by plausible synonyms or parts of the text by equivalent structures; one participant from Group B, on the other hand, appeared to be taking a purely phonetic approach, writing down "source" as "sauce", and "so" as "saw", either due to mishearing or misspelling.

The fact that some of the participants who were familiar with the Gregorian tone misremembered the text, when sung, as including the word "lord", is interesting as it indicates that because of the associations the melody carried for them, they found themselves attempting to conform it to the paradigm of 'Christian psalm' when listening and attempting to memorize the text. This is illustrative of the effect of background and expectations on memory.

If this experiment were to be repeated and expanded, it would be helpful to use a larger participant sample, not only because it was small in this case, but because the level of variation between participants is liable to be high in this specific experiment. It would also be preferable to administer the experiment in a more controlled environment, in person, to ensure that participants were able to give it their undivided attention, as a limitation of online surveys is that participants might perform poorly due to a distracting environment, or due to participating late at night when they are not in full possession of their faculties.

The fact that participants generally recalled Stanza 2 better, regardless of the method of delivery (see Table 2), appears to indicate that, overall, participants found Stanza 2 "easier". This is likely because Stanza 2, although similar in length and lexical range to Stanza 1, had more repetitions and parallel structures (e. g. "we must laugh and we must sing"), which

might have made it easier to memorize. If the experiment were to be repeated, the text used should not include any such repetitions, to avoid this kind of disparity in memorability.

In general, the combination of more long-term approaches and the exclusive use of participants familiar with the music would seem likely to yield results closer along the lines of those of some previous studies. Still, the link between music and memory would appear to be more complex than is perhaps commonly believed.

REFERENCES

Moussard, Aline, Emmanuel Bigand, Sylvie Belleville & Isabelle Peretz. 2012. 'Music as an Aid to Learn New Verbal Information in Alzheimer's Disease', *Music Perception* 29/5: 521-531.

Rainey, David W. & Janet D. Larsen. 2002. 'The Effect of Familiar Melodies on Initial Learning and Long-term Memory for Unconnected Text', *Music Perception* 20/2: 173-186.

Tamminen, J., K. Rastle, J. Darby, R. Lucas, & V. J. Williamson. 2017. 'The impact of music on learning and consolidation of novel words', *Memory*, 25/1: 107-121.

Thaut, Michael H., David A. Peterson, Kimberly Sena Moore & Gerald C. McIntosh. 2008. 'Musical structure facilitates verbal learning in multiple sclerosis', *Music Perception* 25/4: 325-330.