Healthy ageing effects on implicit auditory memory: from encoding to 6-month retention

Roberta Bianco[1,2], Edward T. R. Hall[3], Marcus. T. Pearce[3,4], Maria Chait[1]

[1] Ear Institute, University College London, London, UK

[2] Neuroscience of Perception and Action Laboratory, Italian Institute of Technology (IIT), Rome, Italy

[3] School of Electronic Engineering and Computer Science, Queen Mary University of London, London, UK

[4] Department of Clinical Medicine, Aarhus University, Aarhus, DK

Corresponding authors:

Roberta Bianco
roberta.bianco@iit.it
Neuroscience of Perception & Action Laboratory, Italian Institute of Technology (IIT)
Viale Regina Elena 291, 00161 Rome, Italy

Abstract

Any listening task, from sound recognition to sound-based communication, rests on auditory sensory memory which is known to decline in healthy ageing. However, whether this decline maps on to multiple components and stages of auditory memory remains poorly characterised. We tested ageing effects on implicit auditory memory for rapid tone-patterns in an online unsupervised longitudinal study (day 1, day 8 and 6-month sessions) including younger (aged 20-30) and older adults (aged 60-70). The test required participants to quickly respond to rapid regularly repeating patterns (REG) emerging from random sequences. Patterns were novel in most trials (REGn), but unbeknownst to the participants, a few distinct patterns reoccurred identically throughout the sessions (REGr). After correcting for processing speed, reaction times (RTs) to REGn were taken as a measure of the amount of information held in echoic and short-term memory before detecting the pattern; an RT advantage (RTA) to REGr vs REGn was expected to grow with exposure reflecting implicit long-term memory formation and retention. The results showed that older participants were slower than younger adults in detecting REGn and exhibited a smaller RTA to REGr. Computational simulations using a model of auditory sequence memory suggested that these effects reflect age-related limitations both in early and long-term memory stages. In contrast to ageing-related accelerated forgetting of verbal material, the older adults in the present experiment maintained stable memory traces (RTA) for REGr patterns up to 6 months after the first exposure. The results demonstrate that ageing is associated with poorer echoic and long-term memory formation for tone-patterns, but not forgetting, even over surprisingly long timescales.

Introduction

Memory loss is one of the most significant changes to cognitive processing experienced in healthy ageing [1–5]. The most pronounced memory deficits are related to direct recall of episodic memory [6,7], but evidence is increasingly revealing impairment in older listeners also in tasks that draw on automatic sensory memory processes [8–17].

Sensory memory is at the core of auditory processing [18–21]. The nature of the unfolding signal is such that any listening task, from sound recognition to sound-based communication, depends on the ability to store successive events in memory in order to derive a coherent representation [22–25]. Age related deficits in implicit auditory memory are increasingly being documented [10,26–29], but remain poorly characterised due to limited computational tractability and paucity of longitudinal research designs. Recently, emerging links between auditory processing and dementia are also making it urgent to quantify and understand these impairments. In particular, the brain networks that are thought to underlie auditory memory – involving auditory, frontal and hippocampal areas [30–36] – are the same networks that exhibit the earliest decline in Alzheimer disease [12,13]. This makes auditory memory decline a promising proximity marker of dementia that is worth investigating [37,38].

Traditionally, auditory memory is conceived as a sequence of stores [19]: an echoic buffer stores detailed unprocessed information for several hundred milliseconds to allow successive sounds to be linked to representation of sequences; then information passes to short-term memory for a few seconds and strengthens in long-term memory upon repeated presentations of the to-be-remembered sound [20]. Impairments can potentially arise at any of these different stages.

Older, compared to younger, adults exhibit diminished amplitude and longer latencies of the mismatch negativity (MMN) – an automatic brain response evoked by a rare deviant sound in a sequence of standard sounds [10,26,27,39–41]. The MMN is hypothesised to reflect the process of

comparing incoming inputs with sensory-memory traces [24,42–45]. Diminished MMN in older listeners suggests reduced echoic buffer and/or short-term memory [46]. Age-related performance decline is also reported in probabilistic sequence learning tests using artificial auditory material and hypothesised to reflect short- and long-term implicit memory deficits in older adults [47]. In such tests, target sequences of arbitrary syllables or tones (presentation rate ~2Hz) are structured according to certain probabilistic or deterministic statistics, and repetitively presented to listeners performing a decoy task [48,49]. Relative to novel sequences, a memory benefit for the target sequences is reflected by better immediate sequence reproduction or higher familiarity ratings after the exposure phase. This benefit declines with age and drastically at the age of around 65 [47,50,51].

Age-related decline has also been described in verbal memory tests asking participants to memorise a list of words or stories to a minimum required level of accuracy. Memory is then typically probed with free recall at delayed sessions. Older compared with younger adults exhibit a reduced primacy effect suggesting age-related impairment in short-term memory [52], as well as accelerated long-term forgetting [53–55] implicating impairments of long-term memory consolidation [56]. Notably, accelerated long-term forgetting has gained particular traction in the clinical field because of its potential as predictors of Alzheimer pathology [57].

Despite evidence of age-related deficits at several memory stages, a methodological challenge remains to longitudinally track memory dynamics – from memory formation to long-term retention – with consistent testing measures across stages. Moreover, a general confounding factor has been that previously used auditory materials and tasks involve direct engagement of the participant with the to-be-remembered information (e.g. requirement to recall, judge familiarity etc). Therefore, experience, attention-related or executive processing factors might conceal a core informational aspect of age-related memory decline. Indeed, age-differences in immediate or long-term recall of auditory material might reflect effects of attentional load [58], availability of feedback [51], or vocabulary knowledge [59], as well as rehearsal strategies or interference [53,55,57,60]. Additionally, complex material, such as words, limits cross linguistic and translational diagnostic potential of verbal tests and, critically, it is difficult to informationally quantify and model. As a consequence, whilst it is clear that ageing is associated with auditory sensory memory impairment, the properties of this mnemonic decline are poorly understood.

We tested younger (aged between 20 and 30) and older (aged between 60 and 70) participants with an online paradigm that allows us to quantify short-term memory, the dynamic of long-term memory formation and long-term retention of tone-patterns in an unsupervised manner. The "Auditory pattern Memory test" (ApMEM) [61] employs arbitrary rapid pure-tone sequences spanning the acoustic time scale of speech (20 Hz tone presentation rate, 1 Hz pattern rate) [62]. In 50% of the sequences, a pattern (REG; a repeating sequence of 20 tones) emerges partway, and listeners are required to detect it as quickly as possible (Fig. 1A). The rate at which successive tones are presented precludes deliberate tracking of the sequence structure, instead the REG patterns pop-out perceptually. REG pattern detection is hypothesised to arise from an automatic process that scans the unfolding sequence and maintains a certain portion of the just-heard pattern in memory for comparison with stored representation of the longer-term context. The reaction time (RT) associated with REG detection can therefore be used as a measure of the information (e.g. number of tones) required by listeners to detect the repeating pattern. In the task, the vast majority of REG patterns are novel on each trial (REGn). The associated RT can therefore be used as a measure of the combined contribution of ***echoic and short-term memory*** to pattern detection. Additionally, unbeknownst to the participants, a few different patterns reoccur sparsely every 2 minutes (REGr).

Memory for REGr strengthens through repetitive exposure, as reflected by the gradual emergence of a reaction time advantage (RTA) in REGr pattern detection compared to the REGn. This measure is used as a measure of ***long-term memory formation***. Previous results from young adults have shown that this effect is implicit, in that it is not driven by explicit familiarity [61]. To measure ***long-term memory retention*** of REGr patterns, RT to REGr is also measured 8 days and 6 months after the first exposure.

The advantage of using arbitrary stimuli is that they overcome linguistic barriers and, because the stimuli are unlikely to be encountered in real-life environments, they minimise mnemonic biases such as rehearsal strategies or interference with daily encountered auditory material. Importantly, they also facilitate the ability to systematically relate listeners' memory performance to the information-theoretic properties of the stimuli, and so to computationally quantify the contribution or limitations of mnemonic subcomponents. The Prediction by Partial Matching (PPM) model, has successfully predicted listeners' performance with a variety of discrete musical and artificial auditory sequences [63–68]. Here, we use its memory-constrained variant (PPM-decay, see [67], previously used to simulate performance in the ApMEM task in young listeners' [61] (Fig. 1C). The model encodes sequences by weighting sub-sequences (n-grams) of multiple orders on the basis of recency. Each auditory event is recorded as a single count with a certain weight which decays over time as determined by a customisable decay kernel. This non-linear decay profile simulates the contribution of three subcomponents to memory formation corresponding to echoic buffer, short- and long-term memory decay. Based on stored observations, the model estimates the information content (IC) of each new event (additionally adding customisable levels of prediction noise to replicate similar imperfections in human memory). When a random sequence transitions into a REG pattern, the IC drops reflecting the match between incoming information and past observations held in memory. For a previously encountered REG pattern, there will be a stronger match with information held in long-term memory, resulting in faster pattern discovery.

Fitting ApMEM task data from young and older adults with parameters associated with the decay kernel and memory weights, allowed us to identify the potential sources of impairment in the ageing cohort.

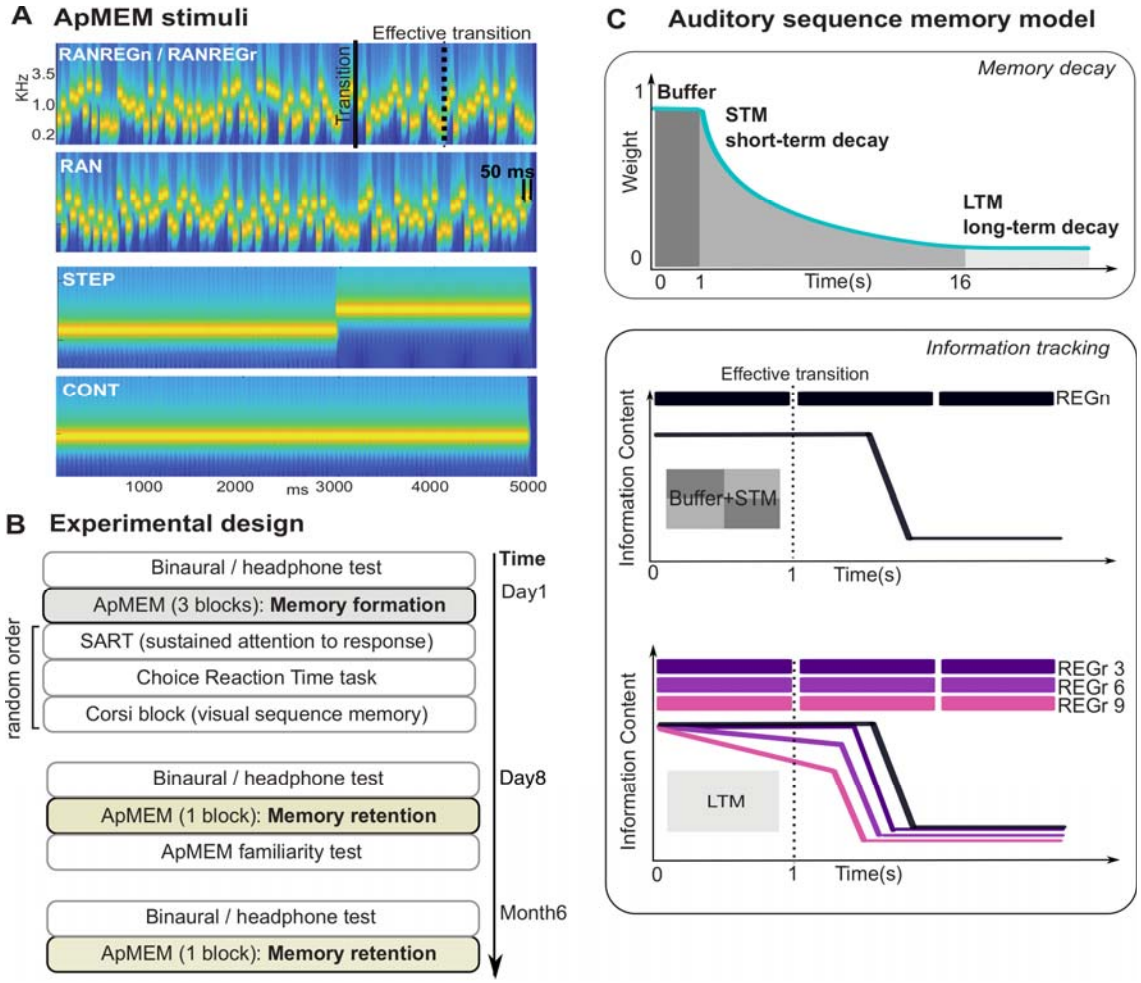


Figure 1. Stimuli, experimental design, and a model of auditory memory. A) Example spectrograms of the ApMEM stimuli*. RAN sequences contained a random arrangement of tone pips. RANREG sequences contained a transition from a random (RAN) to regularly repeating cycles of 20 tone-pips (REGn). The repetition of the REG pattern becomes detectable from the beginning of the second cycle ('Effective transition'). RAN and REGn sequences were generated anew on each trial. Three different regular patterns (REGr) were each presented identically trice within a block. Reoccurrences were spaced ~2 min apart. STEP stimuli, containing a step change in frequency, (and their 'no change' control, CONT) were also included in the stimulus set as a form of attentional checks, and to control for age-related differences in speed of processing reflected in RT-based measures. To distil the computation time required to detect the patterns RTs to RANREG were corrected by the median STEP RTs.* ***B) Experiment design and task order.*** *This was a 3-session study, conducted on day 1 (d1) to test memory formation over 3 blocks of ApMEM task, and on day 8 (d8) and month 6 (m6) to test memory retention with 1 block of ApMEM. Additional cognitive tests were included on d1.* ***C) Auditory sequence memory model.*** *A schematic representation of the parameters implemented in the model to simulate auditory memory.* ***Memory decay.*** *Sequence statistics are memorised through partitioning the unfolding sequence into events and sub-sequences of increasing order (n-grams) that are thereon stored in memory. The salience of these observations ('weight') decays over time through 3 phases indicated by the shaded areas: (1) the high-fidelity echoic buffer defined by a weight and a duration (no decay), (2) the short-term memory phase (STM, with a fast exponential*

decay from the weight of the buffer to the starting weight of the next phase over a given duration), and (3) the long-term memory phase (LTM, with a slow exponential-decay defined by a starting weight and a half-life). ***Information tracking:*** *The model uses these stored statistics to quantify the predictability (IC, where high IC corresponds to low predictability and low IC corresponds to high predictability) of incoming tones. A newly encountered pattern (REGn) is detected when incoming information matches the information held in the buffer and the STM resulting in a drop of IC. A change point detection algorithm detects this change in the IC corresponding to when the REG becomes distinguishable from the RAN. When a pattern is encountered the first time, its representation enters a slow LTM-decay phase. Pattern reoccurrence (REGr, illustrated with different shades of purple to indicate the same pattern after 3, 6 and 9 reoccurrences), leads to weight increase in memory as an index of increasing memory strength meaning that the fall in IC associated with pattern recognition occurs progressively earlier.*

Results

With an online version of ApMEM, we characterised implicit memory for rapid tone patterns in young and older participants over multiple time-scales, from early mnemonic stages required to detect novel patterns to long-term memory formation and retention at 1-week (d8) and 6 months (m6). With a computational model of auditory sequence processing, we then distilled and quantified how different memory parameters contribute to between-group differences in performance on day 1 (d1).

To account for age-related effects on general aspects of the ApMEM task (ApMEM task is RT-based, measures memory for sequences and requires focused attention) we also included tests of processing speed [69], spatial-visual sequence memory [70,71], and attention [72] (Fig. 1B). We explored whether these measures might explain response differences in ApMEM performance between the age groups.

No group difference in accuracy of pattern detection

Overall, the ability to detect the emergence of a pattern from a random sequence, as quantified with d', was consistently high across blocks (Fig. 2B). d' was similar between groups on d1 (W = 2524.5, p = .113, CI [-.006 .231], mean OLD: 3.3 ± .40, YOUNG: 3.17 ± .454), on d8 (W = 2036, p =.513, CI [-3.212e-01 7.035e-06], mean OLD: 3.34 ± .48, YOUNG: 3.36 ± .558), and on m6 (W = 1265, p = 0.082, CI [-1.997736e-05 4.965339e-01], mean OLD: 3.22 ± .55, YOUNG: 2.97 ± .67). This confirms high sensitivity to the presence of regularities and allows us to confidently interpret the between-group differences in RTs as a measure of memory strength.

Hit rate for novel (REGn) and reoccurring patterns (REGr) was computed for each session (d1, d8, m6). Hits were higher for REGr than REGn on d1 in both groups (REGr: 96.6±4.36, REGn: 94.1±7.7 in OLD; REGr: 97±4.57, REGn: 93.6±7.9 in YOUNG; one-sample wilcoxon test of REGr – REGn hit percent, OLD: V = 596, p = .006; YOUNG: V = 807, p < .001), indicating a memory effect for REGr. The hit advantage of REGr over REGn did not differ between the groups on d1 (W = 2439.5, p = .218, CI [-2.86e-05 3.70e+00]), on d8 (W = 2341.5, p= .39, CI [-1.38e-05 3.97e-05]), and m6 (W = 1118.5, p = .531, CI [-3.47e-05 3.44e-05]). Overall, both groups were highly accurate in detecting the patterns, and showed a memory advantage for REGr. Below we demonstrate that this effect, and associated between-group differences, is more sensitively captured when focusing on reaction times (RT).

Age related decline in early mnemonic stages

RTs to simple frequency changes (STEP) collapsed across d1, d8 and m6 showed substantial inter-individual variability and were also generally slower in the OLD than YOUNG group (Fig. 2C; W = 2766.5, p = .007, CI [15.9 92.9], mean OLD: 482 ± 115, YOUNG: 434 ± 108 ms). For each subject, the RT to pattern emergence (RANREGn and RANREGr) was corrected by the RT to the STEP (median per session). This was done to control for individual variability in simple response times, and thus isolate the computation time required to detect an emerging pattern.

We first analysed responses to REGn, as a measure of early mnemonic stages. We computed the median RT to REGn, collapsed across all sessions (Fig. 2D). The OLD group took longer than the YOUNG to detect the REGn patterns (RTs to REG: W = 2654.5, p = .03, CI [5.49 106.25]; mean OLD: 1807 ± .152; YOUNG: 1747 ± 137 ms), suggesting age-related decline of early mnemonic components (echoic / short-term memory) supporting pattern detection.

Age-related decline in long term memory formation

Fig. 2E shows the median RT to REGn vs REGr for each block and session. We first focused on day 1 (memory formation stage). A repeated measure ANOVA with factors condition (RANREGn / RANREGr), block (block 1-3 of d1) and between-subjects factor group (OLD/YOUNG) yielded a main effect of condition [F(1, 130) = 82.52, p < .001, η_p^2 =.39], a main effect of block [F(2, 260) = 6.87, p = .002, η_p^2 = .05] and an interaction of condition by block [F(2, 260) = 7.57, p = .001, η_p^2 = .06]. This confirms the general pattern previously observed for this task [61]: whilst RT to REGn patterns remains stable across blocks (no significant difference between blocks), RT to REGr becomes progressively faster with repeated exposure (block 1 vs 2 p = .004; block 1 vs 3 p < .001).

A main effect of group [F(1, 130) = 8.71, p = .004, η_p^2 = .06] and an interaction of condition with group [F(1, 130) = 4.74, p = .031, η_p^2 = .04] confirmed that, across the 3 blocks of d1, the older group were slower overall than the younger group in detecting the REGr patterns (RT to REGr: OLD, mean 1725 ± 182 ms, vs YOUNG, 1622 ± 164, [t(130) = 3.40, p < .001]). An effect of ageing on REGn RT did not quite reach significance when focusing on the first day only (RT to REGn: OLD, mean 1809 ± 173 ms, vs YOUNG, mean 1761 ± 138 ms, [t(130) = 1.75, p = .08]), possibly due to some noise in the YOUNG data in block 2. No three-way interaction was found [F(2, 260) = 0.76, p = .467, η_p^2 = .01].

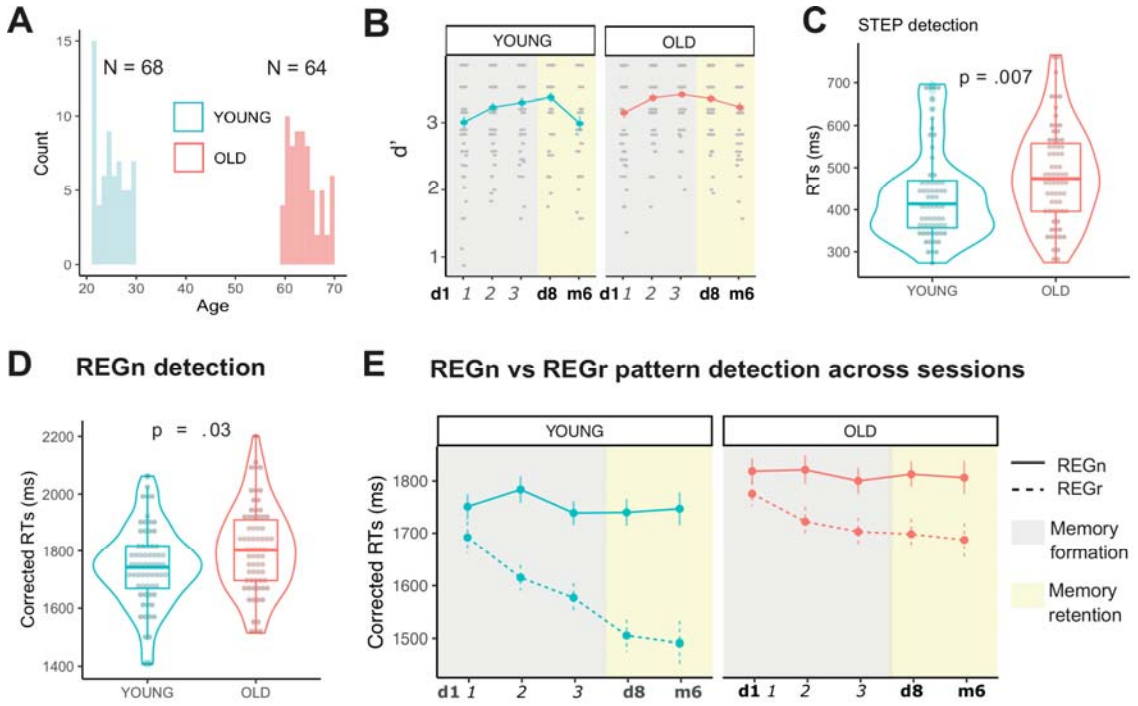


Figure 2. Auditory memory formation in young and older adults (ApMEM task). A) *Participant age distribution.* ***B)*** *d' (sensitivity to emergence of regularity) for the OLD and YOUNG groups. Shaded areas indicate the different stages of memory formation (d1 in grey) and retention (d8, m6 in yellow).* ***C)*** *Median RT to STEP trials (frequency step changes) averaged across the d1, d8 and m6 sessions.* ***(D)*** *Across all sessions (d1, d8 and m6), older listeners exhibited slower REGn RT than young controls.* ***E)*** *RTs to REGr and REGn across the 3 blocks of d1, and the 1 block of d8 and m6. Error bars represent standard error of the mean.*

Computational modelling indicates that reduced performance among older listeners can be explained by reduced echoic buffer duration and faster LTM decay

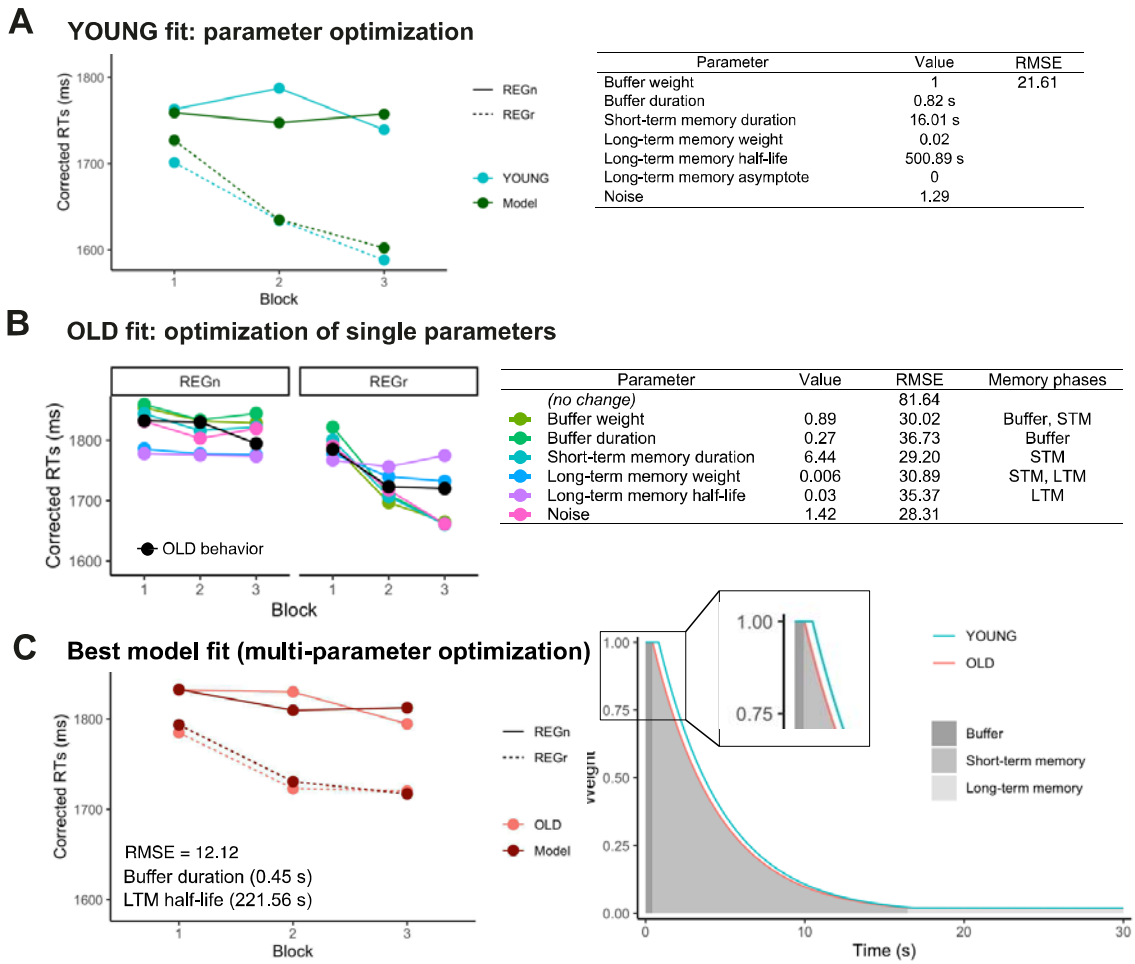
Using a memory-constrained variant of Prediction by Partial Matching [67], a computational model was optimised to fit the observed data on day 1 over blocks 1 to 3 of the ApMEM task for both OLD and YOUNG groups (see methods). We use this model to provide a formal simulation of early memory encoding and long-term memory formation characterising differences between the groups.

Fig. 3A shows the simulated RTs and the parameters optimised to fit the data of the YOUNG group (RMSE = 21.61). The decay kernel that these parameters generate is illustrated in Fig. 3C (right plot). Qualitatively, these parameters show a close correspondence to those obtained when simulating responses of young participants on the same task in [61].

Optimisation to the data of the OLD group first examined whether a single parameter change from the values obtained for YOUNG could explain the differences in RT. The individually optimised parameters and their fit are given in Fig. 3B. It should be noted that, in several cases, the change of a single parameter affects the characteristics of multiple memory phases. No optimisation of only a single parameter managed to adequately fit the observed data for the OLD group, with all models possessing both high RMSEs reflecting an inability to recreate the trajectories of REGn and REGr responses, as displayed in Fig. 3B (left plot). In particular, parameters affecting the buffer, STM,

or overall prediction noise, were unable to reproduce the decrease in learning rate exhibited in block 3 of the observed data for the REGr condition. While LTM weight was able to account for this effect, it could not sufficiently increase simulated RTs for the REGn condition at the same time. Manipulating LTM half-life on its own was unable to produce a fit for either condition.

Next, pairs of parameters were optimised in turn to fit the data for the OLD group (a full list of parameter values and fit of models is given in SUPPLEMENTARY MATERIAL). Multiple parameter sets produced could plausibly fit the observed data, and each contained one parameter controlling properties of the buffer or short-term memory, and one controlling an aspect of long-term memory. The best-fitting combination (RMSE = 12.12) was the optimisation of buffer duration (0.45 s) and LTM half-life (221.56 s). The decay kernel described by these parameters diverges from that of the YOUNG group by having a lower capacity buffer and more rapid long-term decay, as shown in Fig. 3C (right). These differences, and those of the other low-RMSE models fitting the OLD data, indicate that the older group possesses weaker memory formation in both the immediate and long-term mnemonic phases, and that both of these deficits are required to explain the differences observed between the two groups.



Parameter	Value	RMSE
Buffer weight	1	21.61
Buffer duration	0.82 s	
Short-term memory duration	16.01 s	
Long-term memory weight	0.02	
Long-term memory half-life	500.89 s	
Long-term memory asymptote	0	
Noise	1.29	

Parameter	Value	RMSE	Memory phases
(no change)		81.64	
Buffer weight	0.89	30.02	Buffer, STM
Buffer duration	0.27	36.73	Buffer
Short-term memory duration	6.44	29.20	STM
Long-term memory weight	0.006	30.89	STM, LTM
Long-term memory half-life	0.03	35.37	LTM
Noise	1.42	28.31	

Figure 3. Modelling contribution of buffer, STM and LTM to auditory memory formation in young and older adults. A) *Simulated and observed RTs to REGn and REGr conditions for the YOUNG group (left), and the parameter values of the optimised model with RMSE = 21.61(right).* ***B)*** *Simulated RTs and parameter values for models fitting OLD RTs by optimising individual parameter changes from*

those of YOUNG. ***C)*** *Simulated and observed RTs to REGn and REGr conditions for the OLD group using parameter values obtained for YOUNG, modifying buffer duration and LTM half-life yielding RMSE = 12.12 (left), and the memory decay kernels of best fitting models for both groups (right).*

Auditory memory retained for up to 6 months in both older and younger listeners

Fig. 4A displays the RT advantage between REGr and REGn across all experimental sessions. The data revealed that the difference in RTA observed at the end of d1 (and modelled above) persisted when probed at d8 and in m6.

To explicitly test for long-term memory retention, for each subject we compared the RTA at d8 and m6 to that observed in b1 (block 1 of d1). If listeners retained a lasting memory of REGr we expected RTA in d8 and m6 to be different from that in b1. An ANOVA with factors block (b1 / d8) and group yielded main effects of group [$F(1, 130) = 8.69$, $p = .004$, $\eta_p^2 = .06$], block [$F(1, 130) = 26.30$, $p < .001$, $\eta_p^2 = .17$], and no interaction [$F(1, 130) = 3.24$, $p = .074$, $\eta_p^2 = .02$], indicating a greater RTA in d8 than in b1 in both groups [$t(131) = 5.03$, $p < .001$; mean RTA OLD: b1 36.4 ± 189, d8 115 ± .190 ms; YOUNG: b1 71.6 ± 215, d8 234 ± 215 ms], and overall greater in the YOUNG than in the OLD group [$t(130) = 2.94$, $p = .004$].

An ANOVA with factors block (b1 / m6) and group yielded main effects of group [$F(1, 91) = 10.82$, $p = .001$, $\eta_p^2 = .11$], block [$(1, 91) = 16.59$, $p < .001$, $\eta_p^2 = .15$], and no interaction [$F(1, 91) = 1.52$, $p = .221$, $\eta_p^2 = .02$], indicating a greater RTA in m6 than in b1 in both groups [$t(92) = 3.90$, $p < .001$; mean RTA OLD: b1 30.2 ± 198, m6 119 ± .189 ms; YOUNG: b1 89.5 ± 197, d8 255 ± 239 ms], and an overall greater RTA in the YOUNG group [$t(91) = 3.28$, $p = .001$]. This analysis shows that weaker LTM memory is formed in the OLD than YOUNG group, but both groups maintain non-decaying memories of the REGr patterns up to 6 months following initial memory formation.

Lastly, we compared the RTA in d1 block 3 (b3), to those in d8 and m6 (Fig. 4B). A repeated measures ANOVA revealed a main effect of group only [$F(1, 91) = 13.86$, $p < .001$, $\eta_p^2 = .12$], consistent with the overall larger RTA among the young listeners. There was no main effect of block, nor an interaction (block: [$F(2, 182) = .97$, $p = .381$, $\eta_p^2 = .01$]; interaction: [$F(2, 182) = .93$, $p = .397$, $\eta_p^2 = .01$]), confirming a plateauing of the RTA after d1 in both groups – consistent with an enduring memory trace. We also found that RTA in b3 positively correlated with RTA in d8 (spearman's rho = 0.198, p = 0.022), and that RTA in d8 correlated with RTA in m6 (spearman's rho = 0.230, p = 0.026). This indicates a good reliability of individual effects even in online settings.

To summarise, an effect of age emerged across different time scales. In addition to weaker early mnemonic stages (Fig. 2D), the older group exhibited weaker long-term memory, as reflected by smaller RTA in the last block of d1, d8 and in m6. However, there was no evidence of a decline in memory e.g. between d1 and d8 or d8 and m6.

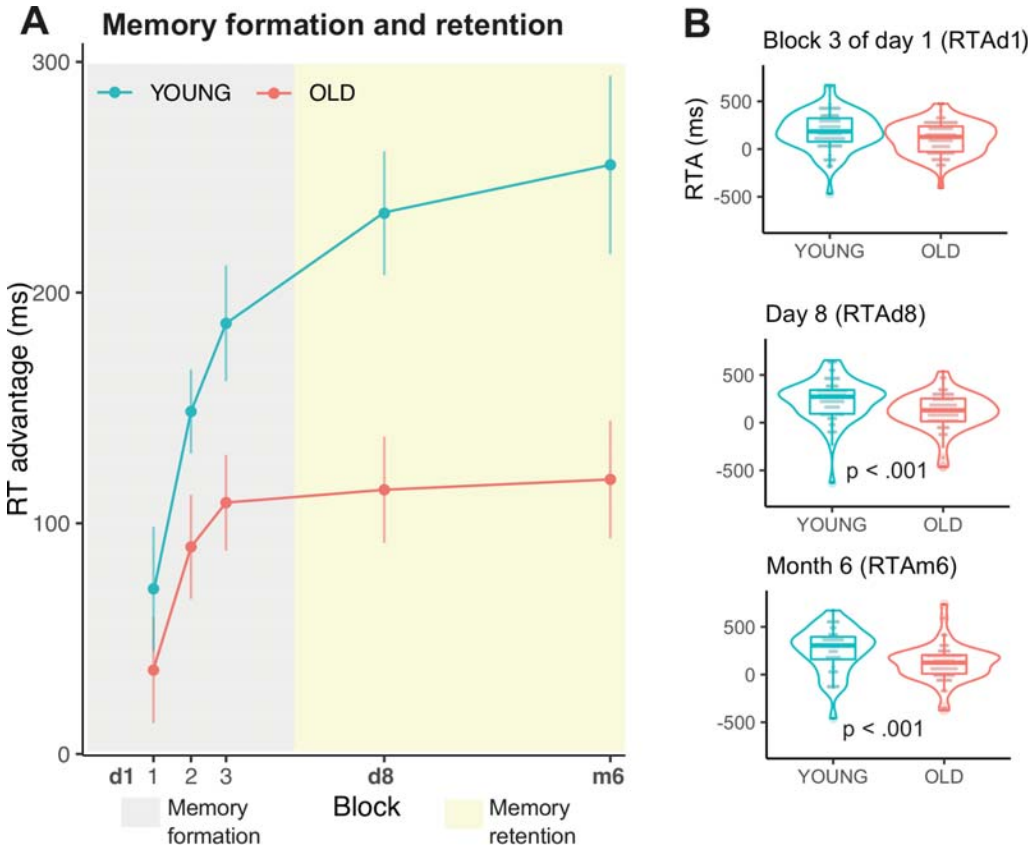


Figure 4. Auditory memory formation and retention in the young and older groups (ApMEM task). (A) Long-term *memory dynamics across d1, d8 and m6 in young and older listeners, quantified as RT advantage (RTA) of REGr over REGn novel patterns.* ***(B)*** *Violin plots showing the individual data for the indexes extracted from the ApMEM task to quantify memory formation (the RTA computed over the last block of the first exposure on d1) and memory retention (the RTA computed on d8 and m6). Statistically significant differences between group means (Mann-Whitney U) are indicated.*

No link between explicit and implicit auditory memory formation on day 1. Explicit memory for the REGr patterns was assessed at the end of d8 with a surprise familiarity task (Fig. 5A). Each REGr was presented once only amongst a large set of foils (REGn) and participants judged if the pattern was familiar. MCC (Mathew correlation coefficient) was used to measure the quality of subjects' binary classification. Both groups exhibited above chance performance (OLD: V = 1734, p < .001; YOUNG: V = 1765.5, p < .001), but performance was poorer in older listeners (W = 1589, p = .040, CI [-1.37e-01 -5.16e-05], mean MCC OLD: .173 ±.217, YOUNG: .253 ± .183) (Fig. 5A). As in our previous findings [61], explicit memory scores did not correlate with the RTA observed in the last block of d1 (spearman's Rho = -.025; p = .777) nor with that in d8 (spearman's Rho = .056; p = .170). These analyses confirm the implicit nature of the RTA measures obtained with ApMEM before running the explicit familiarity task.

Age-related decline in visual-sequence memory and processing speed, but no link with ApMEM RTA. At the group level, older participants showed slower median RTs in the CRT (W = 4000, p < .001, CI [68.80 100.09], mean OLD: 421 ± 97.2; YOUNG: 319 ± 40.5), and greater variability (SD of trials: W = 3227, p < .001, CI [11.63 27.06]) (Fig. 5B-C), reflecting a well-known effect of age-related

processing speed impairment [73]. The median RTs in CRT correlated with the RTs in our control STEP condition (Rho = .312, $p < .001$) confirming that STEP RTs are a good measure for correcting differences in baseline speed.

The older group exhibited worse performance in the Corsi blocks task (OLD vs YOUNG: $t(130) = -3.56$, $p < .001$, mean OLD: 4.54 ± .59; YOUNG: 4.80 ± .537), confirming a the expected age-related decline in visual sequence memory (Fig. 5D) [28,74,75].

In line with previous findings on the SART [76], there was no age-related decline in in the sustained attention accuracy (% 'no-go' fail: W = 1967, p = .343, CI [-11.10 3.10], mean OLD: 33.5 ± 20.3; YOUNG: 38.1 ± 24.5) (Fig. 5E). As expected, RTs were slower in the OLD vs the YOUNG group in the 'go-trials' (W = 11073, $p < .001$, CI: [19.62 52.54], mean OLD: 378 ± 66.5; YOUNG: 348 ± 85.7). The speed-accuracy trade-off was similar between groups: accuracy in 'no go' trials was predicted by RT ($\chi^2(1) = 25.29$, $p < .001$), but not by group or the interaction between RTs and group ($p > 0.1$).

We conducted linear regression analyses to understand to what extent group-specific variance in ApMEM is predicted by performance on these cognitive tasks. We also included weekly hours of physical activity and years of musical training as possible predictors of ApMEM performance. Physical activity has been listed amongst the factors reducing the risk of cognitive and memory decline [77,78]. Evidence from a meta-analysis has linked musical practice in healthy ageing with cognitive benefits both in domain-specific functions (auditory perception) and more general ones [79]. For each outcome measure of ApMEM (Fig. 4B) and group, we performed a linear regression analysis with the predictors: CRT standard deviation, SART RTs, Corsi mean sequence length, and the above-mentioned demographic scores. None of the models were significant in the older group (all p-values > .11). A similar analysis in the younger cohort also yielded non-significant models (all p-values > 0.14). Overall this pattern of results indicates that the variability in ApMEM is not driven by general processing speed, visuo-spatial sequential memory or sustained attention, and might thus reflect age-related deficits specific to auditory memory.

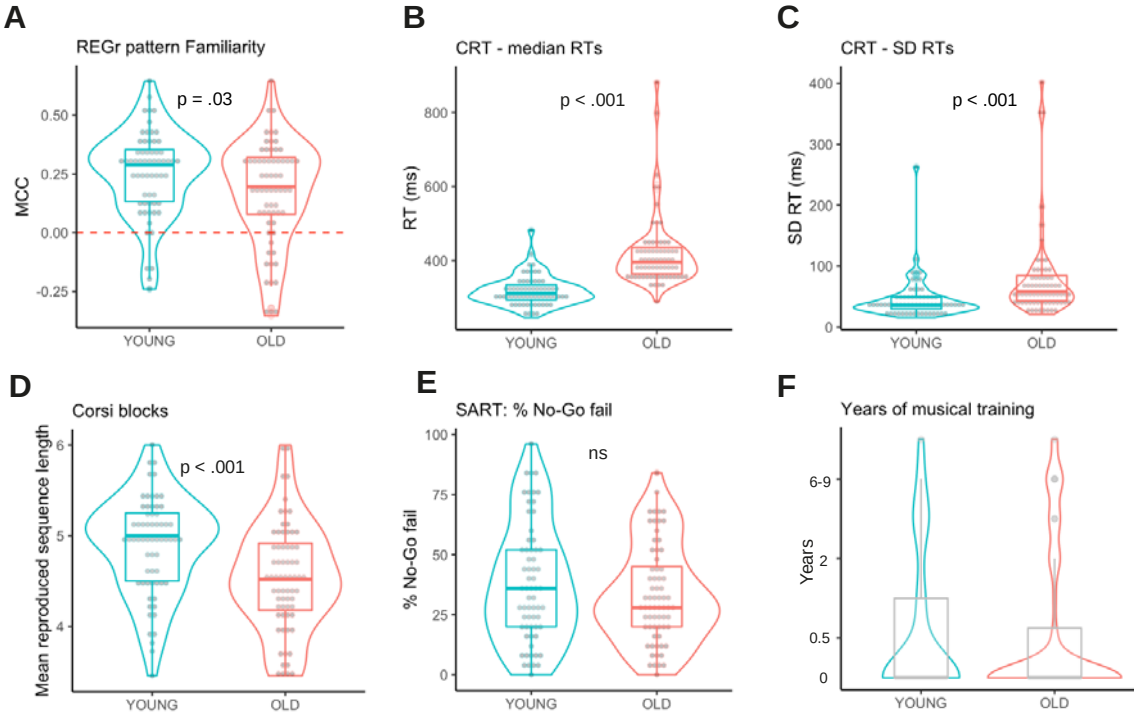


Figure 5. Distribution of performance, across the various tasks, in the young and older groups. *Statistically significant differences between group means (Mann-Whitney U) are indicated. Worse performance in OLD than in YOUNG group was observed in* ***(A)*** *Explicit familiarity scores on the reoccurring patterns obtained at the end of day 8* ***(B)*** *CRT both in terms of median and* ***(C)*** *variance of RTs, and* ***(D)*** *Corsi-blocks. Accuracy on the SART task* ***(E)*** *did not differ between groups. The two groups did not differ in terms of* ***(F)*** *musical training (W = 2076, p = 0.580).*

Discussion

Despite its role in supporting fundamental aspects of auditory perception, how implicit sensory memory is affected by ageing remains poorly understood. Existing work has predominantly focused on verbal material and tasks that require cued reporting which are susceptible to factors such as rehearsal strategies or interference during long retention periods. For example, it is not known whether the "accelerated forgetting" effect recently characterised in older listeners [53,54] is specific to explicit memory, or also extends to more basic auditory mnemonic representations.

Here, we introduce a paradigm that distils and quantifies age-related deficits in core implicit memory mechanisms that support fundamental aspects of auditory scene analysis. We compared younger (aged between 20 and 30 years old) and older adults (aged between 60 and 70 years old) with an RT-based implicit memory test. Participants are required to detect emerging regular patterns from random rapid sequences of tones. Patterns are novel in most trials, but unbeknownst to the participants, a few distinct patterns reoccur identically throughout the experimental sessions. The progressively growing RT advantage of reoccurring vs novel patterns demonstrates that mnemonic traces for the specific reoccurring patterns become more salient in memory through reoccurrence. Notably, the stimuli are arbitrary, and too fast to allow conscious tracking of the sequence events; thus, they minimise active tracking processes and mnemonic interference with real-world sounds. This test allowed us to obtain 'pure' measures of memory at different stages:

from early mnemonic stages to long-term memory formation and retention (for up to 6 months). We found that compared to young adults, older participants were slower in detecting novel patterns and exhibited a smaller RT advantage in detecting reoccurring patterns, indicating deficits in echoic / short memory and long-term memory formation. A computational model of auditory sequence memory fit to the data on the first day of exposure also suggests age-related limitations in both early and long-term mnemonic components. In contrast to demonstrations of accelerated forgetting of verbal material with ageing, here older adults maintained stable memory traces for the reoccurring patterns – an unaltered RT advantage – up to 6 months after the first exposure.

Processing speed does not explain the reduced auditory memory effect in the older cohort

It is important to note that the between-group difference in the derived RT-based measures of memory cannot be explained by general reduced processing speed in the aged cohort [69,80]. This is supported by three arguments: First, RTs were corrected by RT to a simple stimulus change (STEP condition) interspersed in the main ApMEM task. This ensured that performance in pattern detection was controlled for inter-individual biological (e.g., subject's general state of vigilance, or the time taken to perceive an auditory change, to generate a response) or equipment-based differences (e.g. keyboard latency) introducing non-memory specific variability. Second, the worse performance of older than younger adults in the control choice-RT task (CRT) showed the expected age effect on processing speed, and it correlated with the ApMEM STEP control condition, confirming the latter as a valid measure for correcting general differences in baseline speed. Finally, the regression analyses, including control tasks as predictors of ApMEM performance, showed that CRT did not contribute to variability in any of ApMEM-related memory measures. This, together with the absence of correlations between the ApMEM memory measures and tasks associated with sustained attention (SART) and visual-sequence memory suggest that between-group differences in ApMEM reflects age-related deficits specific to auditory memory.

Deficits in both early and long-term mnemonic stages contribute to the age-related performance decline

We showed that older compared with younger participants exhibited overall slower RT in response to novel patterns and formed a smaller RTA for the reoccurring patterns. Whilst a-priori, it could be possible for a single underlying factor (e.g. associated with weaker short-term memory) to explain the deficits in older people, modelling demonstrated that optimisation of parameters associated with both early and later mnemonic stages was necessary to accurately model older listeners' performance.

We modelled the effect of ageing on memory by optimising the model memory decay kernel pertaining to: (1) a high-fidelity echoic memory buffer; (2) a STM phase; and (3) an exponentially decaying LTM phase. Each of these is associated with parameters describing their duration, relative weight, and rate of decay. Amongst the four best-fitting models, each contained one parameter controlling properties of the buffer or short-term memory, and one controlling an aspect of long-term memory. The best fitting model suggested that shorter echoic buffer duration (YOUNG: .82 s; OLD: .45 s) as well as more rapid long-term decay (YOUNG: 500.89 s; OLD: 221.56 s) contributed to the age-related performance decline on day 1.

The first 'pre-perceptual' stage of temporarily holding auditory information allows listeners to bind incoming events with the just heard ones in order to perceive a coherent representation of sequential sounds (e.g., a sequence of single tones as a motive). Limitations of early memory stages

constrain listening in various ways, including poor speech-related pattern recognition [23]. The computational modelling results support the interpretation that the overall slower detection of REGn patterns in the older adults is a consequence of limited buffer / short-term memory components, with temporal capacity of the echoic memory buffer being perhaps the most limiting factor. This is in line with previous hypotheses [81] and consistent with neurophysiological evidence showing a somewhat weaker ability to represent information in echoic memory in older than young adults: when peripheral hearing sensitivity is controlled for, elderly people show diminished amplitude and longer latencies of mismatch negativity responses to tones that deviate from regularities underlying an unfolding sequence [26,39].

Parameters affecting the early stages of auditory memory alone could reproduce the age-effect on responses to novel patterns, but were unable to reproduce the diminished RTA for reoccurring patterns. A combination of reduced memory buffer and more rapid long-term decay (LTM half-life) best accounted for both aspects of performance differences between groups. Overall, this suggests that age-related performance decline could be underpinned by reduced functionality of the early auditory pathway affecting early core auditory cortical regions involved in echoic memory [82,83], and fronto-temporal and hippocampal networks implicated in encoding and maintenance of tone-patterns [30,31,84] and associated with early signs of cognitive decline [85–87] and dementia [88,89].

The decay kernel optimised to older adults' data provides an initial model of how limitations at multiple stages of memory may explain different cognitive performance between populations. The parameters for this model were optimised to fit blocks 1 to 3 of the ApMEM task performed on d1. As the constant long-term decay of the model predicts that memory should eventually reduce to zero, d8 and m6 are beyond the scope of this modelling in its current form. Modelling such time spans, while still being able to recreate effects within the first three blocks, would require a non-trivial addition that could account for memory consolidation over the intervening time periods.

Could the observed age-related memory effects arise from poorer hearing sensitivity in older listeners? We consider this unlikely for several reasons: firstly, the sensitivity to the presence of regularities was high among older adults and did not differ between older and young listeners. Secondly, during the instructions stage, participants were given the opportunity to adjust the sound volume to as high a level as needed. This ensured that all sounds were sufficiently audible. Finally, only participants who passed the headphone/binaural hearing test (see methods) were included.

No evidence of long-term forgetting with ageing: memory traces to arbitrary tone-patterns are retained for up to 6 months from initial exposure.

The older compared with the young group formed weaker memory during the first day of exposure as quantified by a smaller RTA. However, just as observed in the younger group, the RTA in older listeners persisted for 8 days and 6 months after the initial exposure. This very long-lasting effect observed in both groups is noteworthy considering that the RTA was not driven by the explicit familiarity judgments, and participants did not retain explicit awareness of the session 6 months later.

There is growing interest in tests taxing memory circuit functionality at delayed recall because memory problems at this stage could indicate incipient dementia [57,90]. That auditory patterns were not forgotten at delays of 8 days nor 6 months in the older cohort is in contrast to the body of work on accelerated long-term forgetting (ALF) for verbal material in ageing [53–55,60,91]. ALF of verbal material has been reported also in pre-symptomatic autosomal dominant Alzheimer's disease [57], patients with temporal lobe epilepsy [92], and it may represent a failure of memory consolidation

processes [56] due to altered integrity of hippocampal-neocortical (temporal) connections [93]. One explanation of the discrepancy between verbal tasks and ApMEM may reside in the very low probability, compared to verbal material, that subjects were exposed to ApMEM-like sequences outside of the experimental sessions. This might have minimised phenomena such as forgetting due to interference with real world stimuli [94], which perhaps affects older more than younger adults [95]. An alternative explanation might reside in the different nature of the memorization process involved in verbal memory vs ApMEM tasks. Whilst, the former requires subjects to actively memorise and recall, ApMEM relies on implicit memorization through repetition. Implicit memory may thus be rooted in robust biological substrates [96,97] less vulnerable to availability of processing resources, attention or interference, and so more preserved by ageing.

One important open question is why such long-term auditory implicit memory is overall resilient to time decay even later in life? The remarkable examples of preserved auditory memory for music in severe cases of dementia [98–100] suggest that implicit auditory memory has a privileged status in the brain. In young listeners, implicit auditory memory based on repeated exposure has been demonstrated for many sound types, ranging from white noise [101–103], click trains [104], discrete sequences of tones [61,84,105,106], tone clouds [107], and naturalistic textures [103,108]. All these studies capitalise on Hebb-type learning tasks [109], whereby regardless of subject awareness, recognition of reoccurring patterns improves compared to novel ones simply due to reinforcement through repetition. Repetition is perhaps the simplest cue inducing learning because it indicates the presence of patterns potentially relevant for behaviour [110]. Patterns have often a communicative function and are indeed implicitly learned through repetition in human [111–113] and non-human animals [114–118]. The long-term memory of tone-sequences observed here even in older adults might thus reflect this primordial predisposition of the brain to remember patterns even when they sparsely reoccurr.

In conclusion, ageing is associated with poorer auditory echoic / short-term memory and long-term memory formation than young listeners, but not with forgetting. We speculate that ageing might affect frontal-auditory and hippocampal circuits underlying memory formation, but once formed auditory memories of rapid tone-patterns remain accessible for months after the initial exposure even in older listeners. This result might be explained by absence of interference with memory traces of arbitrary stimuli, unlikely to be encountered in daily life [90,119], and suggests preserved long-term implicit auditory memory in ageing. Future studies combining human neuroimaging, animal models and synaptic simulations should shed light on the underlying circuits and neuronal mechanisms [97,120,121].

Methods

Power analysis.

We initially ran an online pilot experiment of ApMEM (N = 20, age between 20-30 years old). The RTA effects size across 3 blocks was hp2 = .22. We expected the difference between groups to be potentially small (hp2 = .02). A prospective power calculation (beta = 0.8; alpha = 0.05) for an ANOVA within-between interaction yielded a required total sample size of N = 41 per group. We set our online target sample size to N = 90 per group to account for drop outs (expected ~30%) due to headphone check exclusion and the unsupervised and longitudinal nature of the experiment. Experimental procedures were approved by the research ethics committee of University College London and informed consent was obtained from each participant.

Participants.

Two participant groups were recruited via the Prolific platform (https://www.prolific.co/). A group of younger participants (age range 20-30 years old; N=93) and a group of older participants (age range 60-70 years old; N=98). A subset of the participants in the older group (N=50) had participated in a previous study [28]. Inclusion criteria included being a native speaker of British English, general good health, no known hearing problems or cognitive impairment (all based on self-report). Participants using low quality audio equipment, or those suffering from binaural hearing loss, were screened out using the test introduced in Milne et al (2020). 29 participants in the older group, and 24 of the younger participants failed the screen and their data were therefore not analysed. Additional exclusion criteria were: (a) poor performance in the main apMEM test (mean d' < 1.5 across blocks in day 1 or day 8; N = 1 in the older group, N = 4 in the younger group excluded.) (b) poor performance on the attentional checks (mean RT to STEP changes larger than 2 STD away from the group mean; N=1 in the younger group excluded). A final N = 132 was analysed: N = 68 (28 female) in the older group, and N = 64 (33 female) in the younger group. Six months later we ran an additional (surprise) session. From the original pool, N = 104 participants (N = 63 older, N = 41 younger group) signed up, and, after exclusion as mentioned above, N = data from 93 subjects were analysed.

General procedure.

This study was implemented in the Gorilla Experiment Builder platform (www.gorilla.sc) [122] and delivered across three sessions: day 1 (d1), day 8 (d8) and month 6 (m6) (Fig. 1B). Participants were initially recruited only for d1 and d8. They were later invited to participate in the m6 session. Participants were recruited via the Prolific platform and remunerated based on an hourly wage of £ 8. Participants who performed below 70% of accuracy in the practice of the main ApMEM task (see below) were prevented from continuing and received a partial compensation for the time spent on the experiment.

On **day 1 (60 minutes)**, participants first completed a headphone / binaural hearing check (Milne et al 2020; strict test version). The test is based on a binaural pitch signal that is only audible over headphones (i.e. where L and R audio channels are delivered separately to each ear). Passing the test requires reasonable quality audio equipment (headphones with separate R and L channels) and preserved binaural hearing [123]. People who failed this very first stage were excluded from the analysis. Next, participants performed the Auditory pattern Memory task (ApMEM; 3 blocks). The main task was preceded by a short practice with a simplified version of the stimuli (see below). People who did not reach 70% accuracy in this practice stage were stopped from continuing the experiment and received a partial compensation. ApMEM was followed by a series of cognitive tests - Sustained Attention to Response Task (SART), the Corsi blocks task, and the Choice Reaction Time (CRT) task - presented in random order across participants (Fig. 6A). More details about each task are provided below. At the end of the session, participants completed a short questionnaire about their listening environment and equipment, their physical activity habits (numbers of hours per week), level of education (ranked as No formal qualification, Secondary education, High school diploma/A-levels, Technical/community college, Undergraduate degree, Graduate degree, Doctorate degree) and years of musical training (ranked as 0, 0.5, 1, 2, 3-5, 6-9, 10 or more).

On **day 8 (15 minutes)**, participants completed the headphone check followed by a single ApMEM block. The session ended with a surprise familiarity test for REGr (see details of ApMEM test above). **6 Months** later participants who completed d1 and d8 were re-invited for another surprise session **(15 minutes)**. This included the headphone check and 1 block of ApMEM.

Tasks.

Headphone / binaural hearing check: This test was used to exclude from the analysis participants with poor sound equipment. We used the strict version of the test [124,125].

ApMEM task. The ApMEM task was used to measure multiple stages of auditory memory [61]. Stimuli (Fig. 1A) were sequences of 50-ms tone-pips of different frequencies generated at a sampling rate of 22.05 kHz and gated on and off with 5-ms raised cosine ramps. Twenty frequencies (logarithmically-spaced values between 222 and 2,000 Hz; 12% steps; loudness normalised based on iso226) were arranged in sequences with a total duration varying between 5.5 and 6 s. The specific order in which these frequencies were successively distributed defined different conditions that were otherwise identical in their spectral and timing profiles. RAN ('random') sequences consisted of tone-pips arranged in random order. This was implemented by sampling uniformly from the pool with the constraint that adjacent tones were not of the same frequency. Each frequency was equiprobable across the sequence duration. The RANREG (random-to-regular) sequences contained a transition between a random (RAN), and a regularly repeating pattern: Sequences with initially randomly ordered tones changed into regularly repeating cycles of 20 frequencies (an overall cycle duration of 1 s; new on each trial). The change occurred between 2.5 and 3 s after sequence onset such that each RANREG sequence contained 3 REG cycles. RAN and RANREGn (RANREG novel) conditions were generated anew for each trial and occurred equiprobably. Additionally, and unbeknownst to participants, 3 different REG patterns reoccurred identically several times within the d1, d8 and m6 sessions (RANREGr condition, reoccurring). The RAN portion of RANREGr trials was always novel. Each of the 3 regular patterns (REGr) reoccurred 3 times per block (every ~ 2 minutes; i.e. 9 presentations overall in d1, and 3 in d8 and 3 in m6). Reoccurrences were distributed within each block such that they occurred at the beginning (first third), middle and end of each block. Two control conditions were also included: sequences of tones of a fixed frequency (CONT), and sequences with a step change in frequency partway through the trial (STEP). The STEP trials served as a lower bound measure of individuals' reaction time to simple acoustic changes. They were also used as attention checks – no, or very slow (see below) responses to STEP trials indicated insufficient task engagement.

Each session of the main task was preceded by a volume adjustment stage. Participants heard a few sounds from the main task and were instructed to adjust the volume to a comfortable listening level. In the main task, participants were instructed to monitor for transitions (50% of trials) from random to regular patterns (RANREG) and frequency changes in STEP stimuli, and press a keyboard button as soon as possible upon pattern detection. On day 1, to acquaint participants with the task, two practice runs were administered. The first practice contained 24 sequences consisting of simplified versions of the stimuli (10 RAN, 10 RANREGn, 2 STEP, 2 CONT), in that sequences were presented at a slower tempo (10 Hz) and contained regularities of 10 tones. The second practice consisted of 21 sequences (9 RAN, 9 RANREGn, 2 STEP, 1 CONT) presented at a faster tempo (20 Hz) and containing regularities of 20 tones, as in the main task. The main task consisted of 3 blocks on d1, 1 on d8 and 1 on m6 sessions. Each block lasted about 6 minutes and contained 43 stimuli (18 RAN, 9 RANREGn, 9 RANREGr, 5 STEP, 2 CONT), with ISI of 1 s. Feedback on accuracy and speed was provided at the end of each trial as in our previous work [61]: a red cross for incorrect responses, and a tick after correct responses. The colour of the tick was green if responses were 'fast' (< 2200 ms from REG onset or <500 ms from the step frequency chance), and orange otherwise. This served to

encourage participants to respond as quickly as possible. The inter-block intervals were set to have a maximum duration of 3 minutes so as to keep the overall duration of the exposure equal across participants. Altogether, on day 1 instructions and practice took approximately 20 min and the main task lasted 18 minutes. On day 8 and month 6 the ApMEM task took 8 min, 2 of which consisted of 20 trials practice.

d' (computed across RANREGn and RANREGr conditions) served as a general measure of sensitivity to regularity. Responses that occurred after the onset of the regular pattern were considered hits, whilst responses to random trials were marked as false alarms. Participants whose d' was smaller than 1.5 were excluded from the analysis as this indicated poor pattern sensitivity. The core analysis focused on the response times (RTs) to the onset of regular patterns as in [61]. RT was defined as the time difference between the onset of the regular pattern or the frequency step change and the participant's button press. For each participant, RTs beyond 2 SD from the mean were discarded. Individuals identified as outliers in the RTs to the STEP condition were excluded from the analysis as this indicated low task engagement. The median STEP RTs computed per session were used as a measure of the latency of the response to a simple acoustic change, and subtracted from the RTs to RANREGn and RANREGr to yield a lower-bound estimate of the computation time required for pattern detection. Lastly, for each subject we computed indexes of RT advantage (RTA) of REGr over REGn to quantify memory at different time points. To do so, we first corrected the RTs to REGr trials by the median RTs to REGn in each block. Then, to calculate the RTA by block, we computed the median RTA across the 3 intra-block presentations and the 3 different REGr patterns.

ApMEM familiarity surprise task. Explicit memory for REGr was examined with a surprise task at the end of day 8. The 3 REGr patterns presented in the ApMEM (only one instance per REGr) were intermixed with 18 REGn patterns, as in [61]. Participants were instructed to indicate which patterns sounded 'familiar'. The task took approximately 2 minutes to complete. Classification was evaluated using the Matthew Correlation Coefficient (MCC) score which ranges between 1 (perfect classification) to -1 (total misclassification) [126,127]. Before starting the task, participants were played a few sounds similar to those in the upcoming task, and asked to adjust the volume to a comfortable listening level.

Choice reaction time task (CRT). The CRT task is an established measure of individual variability in processing speed and known to be linked with age-related decline in higher-level cognitive functions [69]. Subjects were required to respond as soon as possible with the index or middle finger to a cue appearing with equal probability on the left or right box displayed on the screen. The task comprised 20 trials and took approximately 1 minute to complete. The task has two outcome measures: The central tendency (the median RT), and intraindividual variability (the raw standard deviation of the RTs), known to show marked increase with age [73,128].

Corsi blocks (visual-sequence memory) task. Nine identical black squares were presented on the screen. On each trial, following a fixation duration (500 ms) a number of blocks flashed (briefly changed colour from black to yellow; flash duration 500ms; inter-flash-interval 250 ms) in a sequence. Participants had to reproduce the order of the sequence by mouse clicking on the correct blocks. The initial sequence length was 2 blocks. Correct responses resulted in a length increase and incorrect responses in a length decrease. Overall participants completed 20 trials. The task took approximately 5 minutes to complete. As an outcome measure, we computed the mean sequence

duration. This score is considered to reflect the ability to remember the temporal order of spatial sequences and it is known to deteriorate with ageing [28,74,75].

Sustained Attention to Response test (SART). The ApMEM task is attentionally demanding and memory formation may be affected by the listener's capacity to sustain focused attention. The SART task was used to measure individual vigilance and propensity to inattention [72]. Participants were asked to respond by pressing a button to serially presented frequent 'go' visual stimuli (digits from 0 to 9, except 3) but maintain a readiness to withhold a response to rare and unpredictable no-go trials (the digit 3). The task took approximately 8 minutes to complete. The key outcome measure was the % 'no-go' fail – quantifying listeners' ability to successfully stay "on task".

Statistical analyses. Performance was statistically tested with linear analyses of variance (ANOVA) implemented in the R environment using the 'ezANOVA' function [129]. P-values were Greenhouse-Geisser adjusted when sphericity assumptions were violated. Post hoc t-tests were used to test for differences in performance between conditions across blocks and groups. A Bonferroni correction was applied by multiplying p values by the number of comparisons. Resulting values below the significance level of .05 are indicated as n.s. – non-significant. Non-parametric tests were used where normality of the outcome distribution and homogeneity of variances were violated. To isolate the contributions of different tasks to ApMEM performance, we used hierarchical linear regressions.

PPM-decay modelling. Observed data from the ApMEM task were computationally modelled using a memory constrained Prediction by Partial Matching (PPM) model. PPM is a variable-order Markov modelling technique that estimates likelihoods for the occurrence of symbolic sequential events, given the number of occurrences of n-grams of varying size within a training sequence, smoothing between models of different orders [130,131].

Conventional models using PPM possess a perfect memory for all events in their training data, regardless of proximity to the modelled event. In order to model the effects of human memory on learning, Harrison et al. (2020) implemented a PPM model with the ability to down-weight occurrences in the model over time, based on a customisable decay kernel. As used here, the kernel contained three phases: (1) a high-fidelity echoic memory buffer, defined by a weight and a duration; (2) a short-term memory (STM) phase that decays exponentially from the weight of the buffer to the starting weight of the next phase over a given duration; and (3) an exponentially decaying long-term memory (LTM) phase, defined by a starting weight and a half-life; (examples of decay kernels and their phases can be seen in Fig. 3C). Additionally, varying levels of noise were added to event probabilities, replicating similar imperfections in human memory.

All stimuli in blocks 1 to 3 of day 1 were modelled, as presented for each stimulus set, maintaining the tone, stimulus and block timings of the task. Models were trained dynamically, estimating a probability for each tone, given the sequence preceding it and all preceding stimuli, which was converted into information content (negative log-base-2 probability). Models were limited to a maximum n-gram length of 5 symbols (an order bound of 4). As in [61], changes in information content were identified for REGn and REGr stimuli using the nonparametric change-point detection algorithm of [132], a sequential application of the Mann-Whitney test, while controlling for a Type I error rate of 1 in 10000.

Model parameters were optimised so as to find the decay configuration that best reproduced the observed data of the younger group using Rowan's Subplex algorithm, as

implemented in the NLopt package [133,134]. Initial parameter values were adapted from the manually fitted parameters of [61]. To account for the increased variability of change points due to modelling prediction noise, for every optimisation iteration modelling was repeated 30 times, refreshing model memory between each. Repeated change points were then averaged for individual stimuli. Optimisation sought to minimise the root-mean-square error (RMSE) between observed RTs and modelled change points, when averaged for each block, for each of the REGn and REGr conditions.

To characterise differences between the older and younger groups, first, observed data for the older group were modelled by optimising a single parameter while holding all others to the values obtained for the younger group. The fit of these models is shown in Fig3. B. As no single change of only an individual parameter adequately reproduced the observed data, pairs of parameters were then optimised with remaining parameters the same as those for the younger group. The parameters of the best-fitting of these models, based on RMSE, were selected as those characterising the older group.

Funding

This work was supported by a BBSRC grant (BB/P003745/1) to M. C., the NIHR UCLH BRC Deafness and Hearing Problems Theme, and an ARUK award and Marie Skłodowska-Curie Individual fellowship to R. B.

Competing interests

The authors declare that no competing interests exist.

Data availability

The datasets for this study will be make publicly available upon peer-reviewed publication.

References

1. Raz, N. & Lindenberger, U. Only Time will Tell: Cross-sectional Studies Offer no Solution to the Age-Brain-Cognition Triangle—Comment on Salthouse (2011). *Psychol Bull* **137**, 790–795 (2011).
2. Cansino, S. Episodic memory decay along the adult lifespan: A review of behavioral and neurophysiological evidence. *Int. J. Psychophysiol.* **71**, 64–69 (2009).
3. Füllgrabe, C. On the Possible Overestimation of Cognitive Decline: The Impact of Age-Related Hearing Loss on Cognitive-Test Performance. *Front. Neurosci.* **14**, (2020).
4. Nyberg, L., Lövdén, M., Riklund, K., Lindenberger, U. & Bäckman, L. Memory aging and brain maintenance. *Trends Cogn. Sci.* **16**, 292–305 (2012).
5. Salthouse, T. A. Neuroanatomical substrates of age-related cognitive decline. *Psychol. Bull.* **137**, 753–784 (2011).
6. Koen, J. D. & Yonelinas, A. P. The Effects of Healthy Aging, Amnestic Mild Cognitive Impairment, and Alzheimer's Disease on Recollection and Familiarity: A Meta-Analytic Review. *Neuropsychol. Rev.* **24**, 332–354 (2014).
7. Koen, J. D. & Yonelinas, A. P. Recollection, not familiarity, decreases in healthy ageing: Converging evidence from four estimation methods. *Memory* **24**, 75–88 (2016).
8. Schneider, B. A. & Pichora-Fuller, M. K. Implications of perceptual deterioration for cognitive

aging research. in *The Handbook of Aging and Cognition* 155–219 (2000).
9. Wayne, R. V. & Johnsrude, I. S. A review of causal mechanisms underlying the link between age-related hearing loss and cognitive decline. *Ageing Res. Rev.* **23**, 154–166 (2015).
10. Rimmele, J., Sussman, E., Keitel, C., Jacobsen, T. & Schröger, E. Electrophysiological evidence for age effects on sensory memory processing of tonal patterns. *Psychol Aging* **27**, 384–398 (2012).
11. Sluming, V. *et al.* Voxel-Based Morphometry Reveals Increased Gray Matter Density in Broca's Area in Male Symphony Orchestra Musicians. *Neuroimage* **17**, 1613–1622 (2002).
12. Griffiths, T. D. *et al.* How Can Hearing Loss Cause Dementia? *Neuron* 1–12 (2020). doi:10.1016/j.neuron.2020.08.003
13. Johnson, J. C. S. *et al.* Hearing and dementia: from ears to brain. *Brain* **144**, 391–401 (2021).
14. Humes, L. E., Busey, T. A., Craig, J. & Kewley-Port, D. Are age-related changes in cognitive function driven by age-related changes in sensory processing? *Attention, Perception, Psychophys.* **75**, 508–524 (2013).
15. Cheng, C. H., Hsu, W. Y. & Lin, Y. Y. Effects of physiological aging on mismatch negativity: A meta-analysis. *Int. J. Psychophysiol.* **90**, 165–171 (2013).
16. Rieckmann, A. & Bäckman, L. Implicit learning in aging: Extant patterns and new directions. *Neuropsychol. Rev.* **19**, 490–503 (2009).
17. Janacsek, K., Fiser, J. & Nemeth, D. The best time to acquire new skills: Age-related differences in implicit sequence learning across the human lifespan. *Dev. Sci.* **15**, 496–505 (2012).
18. Nees, M. A. Have we forgotten auditory sensory memory? Retention intervals in studies of nonverbal auditory working memory. *Front. Psychol.* **7**, 1–6 (2016).
19. Atkinson, R. C. & Shiffrin, R. M. Human Memory: A Proposed System and its Control Processes. *Psychol. Learn. Motiv. - Adv. Res. Theory* **2**, 89–195 (1968).
20. Winkler, I. & Cowan, N. From sensory to long-term memory: Evidence from auditory memory reactivation studies. *Exp. Psychol.* **52**, 3–20 (2005).
21. Cowan, N. On short and long auditory stores. *Psychol. Bull.* **96**, 341–370 (1984).
22. Heilbron, M. & Chait, M. Great Expectations: Is there Evidence for Predictive Coding in Auditory Cortex? *Neuroscience* **389**, 54–73 (2018).
23. Massaro, D. W. & Cohen, M. M. Preperceptual Auditory Storage in Speech Recognition. in 226–245 (Springer, Berlin, Heidelberg, 1975). doi:10.1007/978-3-642-81000-8_14
24. Baldeweg, T. Repetition effects to sounds: Evidence for predictive coding in the auditory system [1]. *Trends Cogn. Sci.* **10**, 93–94 (2006).
25. Rimmele, J. M., Sussman, E. & Poeppel, D. The role of temporal structure in the investigation of sensory memory, auditory scene analysis, and speech perception: A healthy-aging perspective. *Int. J. Psychophysiol.* **95**, 175–183 (2015).
26. Pekkonen, E. *et al.* Aging Effects on Auditory Processing: An Event-Related Potential Study. *Exp. Aging Res.* **22**, 171–184 (1996).
27. Jääskeläinen, I. P., Varonen, R., Näätänen, R. & Pekkonen, E. Decay of cortical preattentive sound discrimination in middle-age. *Neuroreport* **10**, 123–126 (1999).
28. Bianco & Chait, M. No link between Speech-in-noise perception and Auditory short-term memory – evidence from a large cohort of older and younger listeners. *PsyArXiv Preprints* (2022). doi:10.31234/osf.io/mjg4q
29. Fogerty, D., Humes, L. E. & Busey, T. A. Age-related declines in early sensory memory: Identification of rapid auditory and visual stimulus sequences. *Front. Aging Neurosci.* **8**, 1–16 (2016).
30. Barascud, N., Pearce, M. T., Griffiths, T., Friston, K. & Chait, M. Brain responses in humans reveal ideal-observer-like sensitivity to complex acoustic patterns. *Proc. Natl. Acad. Sci.* **113**, E616-25 (2016).
31. Kumar, S. *et al.* Representations of specific acoustic patterns in the auditory cortex and

hippocampus. *Proc. R. Soc. B Biol. Sci.* **281**, 20141000 (2014).
32. Brown, S., Martinez, M. J. & Parsons, L. M. Passive music listening spontaneously engages limbic and paralimbic systems. *Neuroreport* **15**, 2033–2037 (2004).
33. Schapiro, A. C., Gregory, E., Landau, B., McCloskey, M. & Turk-Browne, N. B. The necessity of the medial temporal lobe for statistical learning. *J. Cogn. Neurosci.* **26**, 1736–1747 (2014).
34. Burunat, I., Alluri, V., Toiviainen, P., Numminen, J. & Brattico, E. Dynamics of brain activity underlying working memory for music in a naturalistic condition. *Cortex* **57**, 254–269 (2014).
35. Watanabe, T., Yagishita, S. & Kikyo, H. Memory of music: Roles of right hippocampus and left inferior frontal gyrus. *Neuroimage* **39**, 483–491 (2008).
36. Schmithorst, V. J. Separate cortical networks involved in music perception: Preliminary functional MRI evidence for modularity of music processing. *Neuroimage* **25**, 444–451 (2005).
37. Fleischman, D. A. Repetition priming in aging and Alzheimer's disease: An integrative review and future directions. *Cortex* **43**, 889–897 (2007).
38. Swords, G. M., Nguyen, L. T., Mudar, R. A. & Llano, D. A. Auditory system dysfunction in Alzheimer disease and its prodromal states: A review. *Ageing Res. Rev.* **44**, 49–59 (2018).
39. Cooper, R. J., Todd, J., McGill, K. & Michie, P. T. Auditory sensory memory and the aging brain: A mismatch negativity study. *Neurobiol. Aging* **27**, 752–762 (2006).
40. Pekkonen, Hirvonen, J., Jääskeläinen, I. P., Kaakkola, S. & Huttunen, J. Auditory sensory memory and the cholinergic system: Implications for Alzheimer's disease. *Neuroimage* **14**, 376–382 (2001).
41. Kiang, M., Braff, D. L., Sprock, J. & Light, G. A. The relationship between preattentive sensory processing deficits and age in schizophrenia patients. *Clin. Neurophysiol.* **120**, 1949–1957 (2009).
42. Näätänen, R., Paavilainen, P., Rinne, T. & Alho, K. The mismatch negativity (MMN) in basic research of central auditory processing: A review. *Clin. Neurophysiol.* **118**, 2544–2590 (2007).
43. Squires, K. C., Wickens, C., Squires, N. K. & Donchin, E. The effect of stimulus sequence on the waveform of the cortical event-related potential. *Science (80-.).* **193**, 1142–1146 (1976).
44. Atienza, M., Cantero, J. L. & Dominguez-Marin, E. Mismatch negativity (MMN): An objective measure of sensory memory and long-lasting memories during sleep. *Int. J. Psychophysiol.* **46**, 215–225 (2002).
45. Haeaschel, C., Vernon, D. J., Dwivedi, P., Gruzelier, J. H. & Baldeweg, T. Event-related brain potential correlates of human auditory sensory memory-trace formation. *J. Neurosci.* **25**, 10494–10501 (2005).
46. Näätänen, R. *et al.* The mismatch negativity (MMN) - A unique window to disturbed central auditory processing in ageing and different clinical conditions. *Clin. Neurophysiol.* **123**, 424–458 (2012).
47. Lukács, Á. & Kemény, F. Development of different forms of skill learning throughout the lifespan. *Cogn. Sci.* **39**, 383–404 (2015).
48. Petkov, C. I. & ten Cate, C. Structured Sequence Learning: Animal Abilities, Cognitive Operations, and Language Evolution. *Top. Cogn. Sci.* **12**, 828–842 (2020).
49. Christiansen, M. H. Implicit Statistical Learning: A Tale of Two Literatures. *Top. Cogn. Sci.* **11**, 468–481 (2019).
50. Schevenels, K., Altvater-Mackensen, N., Zink, I., De Smedt, B. & Vandermosten, M. Aging effects and feasibility of statistical learning tasks across modalities. *Aging, Neuropsychol. Cogn.* **00**, 1–30 (2021).
51. Herff, S. A., Zhen, S., Yu, R. & Agres, K. R. Age-dependent statistical learning trajectories reveal differences in information weighting. *Psychol. Aging* **35**, 1090–1104 (2020).
52. Murphy, D. R., Craik, F. I. M., Li, K. Z. H. & Schneider, B. A. Comparing the effects of aging and background noise on short-term memory performance. *Psychol. Aging* **15**, 323–334 (2000).
53. Wearn, A. R. *et al.* Accelerated long-term forgetting in healthy older adults predicts cognitive decline over 1 year. *Alzheimer's Res. Ther.* **12**, 1–9 (2020).

54. Elliott, G., Isaac, C. L. & Muhlert, N. Measuring forgetting: A critical review of accelerated long-term forgetting studies. *Cortex* **54**, 16–32 (2014).
55. Mary, A., Schreiner, S. & Peigneux, P. Accelerated long-term forgetting in aging and intra-sleep awakenings. *Front. Psychol.* **4**, 1–11 (2013).
56. Hoefeijzers, S., Dewar, M., Della Sala, S., Zeman, A. & Butler, C. Accelerated long-term forgetting in transient epileptic amnesia: An acquisition or consolidation deficit? *Neuropsychologia* **51**, 1549–1555 (2013).
57. Weston, P. S. J. *et al.* Accelerated long-term forgetting in presymptomatic autosomal dominant Alzheimer's disease: a cross-sectional study. *Lancet Neurol.* **17**, 123 (2018).
58. Palmer, S. D., Hutson, J. & Mattys, S. L. Statistical learning for speech segmentation: Age-related changes and underlying mechanisms. *Psychol. Aging* **33**, 1035–1044 (2018).
59. Schneider, B. A., Daneman, M. & Pichora-Fuller, M. K. Listening in aging adults: From discourse comprehension to psychoacoustics. *Can. J. Exp. Psychol.* **56**, 139–152 (2002).
60. Manes, F., Serrano, C., Calcagno, M. L., Cardozo, J. & Hodges, J. Accelerated forgetting in subjects with memory complaints A new form of Mild Cognitive Impairment? *J Neurol* **255**, 1067–1070 (2008).
61. Bianco, R. *et al.* Long-term implicit memory for sequential auditory patterns in humans. *Elife* 9:e56073 (2020).
62. Rosen, S. Temporal information in speech: acoustic, auditory and linguistic aspects. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences* **336**, 367–373 (1992).
63. Di Liberto, G. M. *et al.* Cortical encoding of melodic expectations in human temporal cortex. *Elife* 9:e51784 (2020).
64. Bianco, R., Ptasczynski, L. E. & Omigie, D. Pupil responses to pitch deviants reflect predictability of melodic sequences. *Brain Cogn.* **138**, 103621 (2020).
65. Pearce, M. T., Ruiz, M. H., Kapasi, S., Wiggins, G. a & Bhattacharya, J. Unsupervised statistical learning underpins computational, behavioural, and neural manifestations of musical expectation. *Neuroimage* **50**, 302–13 (2010).
66. Kern, P., Heilbron, M., Lange, F. P. de & Spaak, E. Cortical activity during naturalistic music listening reflects short-range predictions based on long-term experience. *bioRxiv* 2022.06.08.495241 (2022).
67. Harrison, P. M. C., Bianco, R., Chait, M. & Pearce, M. T. *PPM-Decay: A computational model of auditory prediction with memory decay*. *PLoS Computational Biology* **16**, (2020).
68. Quiroga-Martinez, D. R. *et al.* Decomposing neural responses to melodic surprise in musicians and non-musicians: Evidence for a hierarchy of predictions in the auditory system. *Neuroimage* **215**, (2020).
69. Salthouse, T. A. The processing-speed theory of adult age differences in cognition. *Psychol. Rev.* **103**, 403–428 (1996).
70. Kessels, R. P. C., Van Zandvoort, M. J. E., Postma, A., Kappelle, L. J. & De Haan, E. H. F. The Corsi Block-Tapping Task: Standardization and normative data. *Appl. Neuropsychol.* **7**, 252–258 (2000).
71. Corsi, P. M. Human memory and the medial temporal region of the brain. *Diss. Abstr. Int. Sect. B. Sci. Eng.* **34**, 891 (1972).
72. Manly, T., Davison, B., Heutink, J., Galloway, M. & Robertson, I. H. Not enough time or not enough attention? Speed , error and self-maintained control in the Sustained Attention to Response Test (SART). *Clin. Neuropsychol. Assess.* **3**, 167–177 (2000).
73. Hultsch, D. F., MacDonald, S. W. S. & Dixon, R. A. Variability in reaction time performance of younger and older adults. *Journals Gerontol. - Ser. B Psychol. Sci. Soc. Sci.* **57**, 101–115 (2002).
74. Beigneux, K., Plaie, T. & Isingrini, M. Aging effect on visual and spatial components of working memory. *Int. J. Aging Hum. Dev.* **65**, 301–314 (2007).
75. Fournet, N. *et al.* Evaluating short-term and working memory in older adults: French

normative data. *Aging Ment. Health* **16**, 922–930 (2012).
76. de Kerangal, M., Vickers, D. & Chait, M. The effect of healthy aging on change detection and sensitivity to predictable structure in crowded acoustic scenes. *Hear. Res.* **399**, 108074 (2021).
77. Nyberg, L. & Pudas, S. Successful Memory Aging. *Annu. Rev. Psychol.* **70**, 219–243 (2019).
78. Sofi, F. *et al.* Physical activity and risk of cognitive decline: A meta-analysis of prospective studies. *J. Intern. Med.* **269**, 107–117 (2011).
79. Román-Caballero, R., Arnedo, M., Triviño, M. & Lupiáñez, J. Musical practice as an enhancer of cognitive function in healthy aging - A systematic review and meta-analysis. *PLoS One* **13**, 1–23 (2018).
80. Zajac, I. T. & Nettelbeck, T. Auditory speed tasks as potential candidates for the study of cognitive ageing. *Aging, Neuropsychol. Cogn.* **25**, 167–185 (2018).
81. Herrmann, B., Maess, B. & Johnsrude, I. S. A neural signature of regularity in sound is reduced in older adults. *Neurobiol. Aging* **109**, 1–10 (2022).
82. Teichert, T. & Gurnsey, K. Formation and decay of auditory short-term memory in the macaque monkey. *J. Neurophysiol.* **121**, 2401–2415 (2019).
83. Ng, C. W. & Recanzone, G. H. Age-Related Changes in Temporal Processing of Rapidly-Presented Sound Sequences in the Macaque Auditory Cortex. *Cereb. Cortex* **28**, 3775–3796 (2018).
84. Bonetti, L. *et al.* Brain recognition of previously learned versus novel temporal sequences: a differential simultaneous. 1–14 (2022).
85. Du, A. T. *et al.* Magnetic resonance imaging of the entorhinal cortex and hippocampus in mild cognitive impairment and Alzheimer's disease. *J. Neurol. Neurosurg. Psychiatry* **71**, 441–447 (2001).
86. Frisoni, G. B. *et al.* Hippocampal and entorhinal cortex atrophy in frontotemporal dementia and Alzheimer's disease. *Neurology* **52**, 91–100 (1999).
87. Allen, J. S., Bruss, J., Brown, C. K. & Damasio, H. Normal neuroanatomical variation due to age: The major lobes and a parcellation of the temporal region. *Neurobiol. Aging* **26**, 1245–1260 (2005).
88. Benhamou, E. & Warren, J. D. Disorders of music processing in dementia. *Music Aging Brain* 107–149 (2020). doi:10.1016/b978-0-12-817422-7.00004-3
89. Hsieh, S., Hornberger, M., Piguet, O. & Hodges, J. R. Neural basis of music knowledge: Evidence from the dementias. *Brain* **134**, 2523–2534 (2011).
90. Ryan, T. J. & Frankland, P. W. Forgetting as a form of adaptive engram cell plasticity. *Nat. Rev. Neurosci.* 1–14 (2022). doi:10.1038/s41583-021-00548-3
91. Davis, H. P. *et al.* Acquisition, recall, and forgetting of verbal information in long-term memory by young, middle-aged, and elderly individuals. *Cortex* **39**, 1063–1091 (2003).
92. Blake, R. V., Wroe, S. J., Breen, E. K. & McCarthy, R. A. Accelerated forgetting in patients with epilepsy. Evidence for an impairment in memory consolidation. *Brain* **123**, 472–483 (2000).
93. Alvarez, P. & Squire, L. R. Memory consolidation and the medial temporal lobe: A simple network model. *Proc. Natl. Acad. Sci. U. S. A.* **91**, 7041–7045 (1994).
94. Davis, R. L. & Zhong, Y. The Biology of Forgetting—A Perspective. *Neuron* **95**, 490–503 (2017).
95. Wais, P. E. & Gazzaley, A. Distractibility during retrieval of long-term memory: Domain-general interference, neural networks and increased susceptibility in normal aging. *Front. Psychol.* **5**, 1–12 (2014).
96. Bailey, C. H., Kandel, E. R. & Si, K. The persistence of long-term memory: A molecular approach to self-sustaining changes in learning-induced synaptic growth. *Neuron* **44**, 49–57 (2004).
97. Ohno, T. *et al.* Short-term plasticity and long-term potentiation mimicked in single inorganic synapses. *Nat. Mater.* **10**, 591–595 (2011).
98. Jacobsen, J. H. *et al.* Why musical memory can be preserved in advanced Alzheimer's disease.

Brain **138**, 2438–2450 (2015).
99. Baird, A. & Samson, S. Memory for music in Alzheimer's disease: Unforgettable? *Neuropsychol. Rev.* **19**, 85–101 (2009).
100. Benhamou, E. *et al.* Decoding expectation and surprise in dementia: the paradigm of music. *Brain Commun.* **3**, (2021).
101. Agus, T. R., Thorpe, S. J. & Pressnitzer, D. Rapid Formation of Robust Auditory Memories: Insights from Noise. *Neuron* **66**, 610–618 (2010).
102. Dauer, T., Henry, M. J. & Herrmann, B. Auditory perceptual learning depends on temporal regularity and certainty. *J. Exp. Psychol. Hum. Percept. Perform.* **48**, 755–770 (2022).
103. Ringer, H., Schröger, E. & Grimm, S. Perceptual Learning and Recognition of Random Acoustic Patterns. *Audit. Percept. Cogn.* **00**, 1–23 (2022).
104. Kang, H., Agus, T. R. & Pressnitzer, D. Auditory memory for random time patterns. *J. Acoust. Soc. Am.* **142**, 2219–2232 (2017).
105. Herrmann, B., Araz, K. & Johnsrude, I. S. Sustained neural activity correlates with rapid perceptual learning of auditory patterns. *Neuroimage* **238**, 118238 (2021).
106. Leek, M. R. & Watson, C. S. Auditory perceptual learning of tonal patterns. *Percept. Psychophys.* **43**, 389–394 (1988).
107. Agus, T. R. & Pressnitzer, D. Repetition detection and rapid auditory learning for stochastic tone clouds. *J. Acoust. Soc. Am.* **150**, 1735–1749 (2021).
108. Woods, K. J. P. & Mcdermott, J. H. Schema learning for the cocktail party problem. *Proc. Natl. Acad. Sci.* **115**, 1–10 (2018).
109. Hebb, D. O. Brain mechanisms and learning. in *Distinctive features of learning in the higher animal* (ed. Delafresnaye, J. F.) 37–46 (New York: Oxford University Press., 1961).
110. McDermott, J. H., Wrobleski, D. & Oxenham, A. J. Recovering sound sources from embedded repetition. *Proc. Natl. Acad. Sci.* **108**, 1188–1193 (2011).
111. Saffran, J. R., Aslin, R. N. & Newport, E. L. Statistical Learning by 8-Month-Old Infants. *Science (80-.).* **274**, 1926–1928 (1996).
112. Aslin, R. N. Statistical learning: a powerful mechanism that operates by mere exposure. *Wiley Interdiscip. Rev. Cogn. Sci.* **8**, 1–7 (2017).
113. Smalle, E. H. M., Page, M. P. A., Duyck, W., Edwards, M. & Szmalec, A. Children retain implicitly learned phonological sequences better than adults: a longitudinal study. *Dev. Sci.* **21**, (2018).
114. Hauser, M. D., Newport, E. L. & Aslin, R. N. Segmentation of the speech stream in a non-human primate: Statistical learning in cotton-top tamarins. *Cognition* **78**, B53–B64 (2001).
115. Wilson, B. *et al.* Auditory artificial grammar learning in Macaque and Marmoset monkeys. *J. Neurosci.* **33**, 18825–18835 (2013).
116. Cazala, A., Giret, N., Edeline, J.-M. & Del Negro, C. Neuronal Encoding in a High-Level Auditory Area: From Sequential Order of Elements to Grammatical Structure. *J. Neurosci.* **39**, 6150–6161 (2019).
117. Lu, K. & Vicario, D. S. Statistical learning of recurring sound patterns encodes auditory objects in songbird forebrain. *Proc. Natl. Acad. Sci.* **111**, 14553–14558 (2014).
118. Soyman, E. & Vicario, D. S. *Rapid and long-lasting improvements in neural discrimination of acoustic signals with passive familiarization*. *PLoS ONE* **14**, (2019).
119. Hardt, O., Nader, K. & Nadel, L. Decay happens: The role of active forgetting in memory. *Trends Cogn. Sci.* **17**, 111–120 (2013).
120. Gershman, S. J. The molecular memory code and synaptic plasticity: a synthesis. 1–51 (2022).
121. Poeppel, D. & Idsardi, W. We don't know how the brain stores anything, let alone words. *Trends Cogn. Sci.* **26**, 1054–1055 (2022).
122. Anwyl-Irvine, A. L., Dalmaijer, E., Hodges, N. & Evershed, J. Online Timing Accuracy and Precision: A comparison of platforms, browsers, and participant's devices. 1–22 (2020). doi:10.31234/osf.io/jfeca

123. Sanchez Lopez, R., Bianchi, F., Fereczkowski, M., Santurette, S. & Dau, T. Data-Driven Approach for Auditory Profiling and Characterization of Individual Hearing Loss. *Trends Hear.* **22**, 1–12 (2018).
124. Milne, A. *et al.* An online headphone screening test based on dichotic pitch. *Behav. Res. Methods* (2020). doi:10.1101/2020.07.21.214395
125. Bianco, Mills, G., de Kerangal, M., Rosen, S. & Chait, M. Reward Enhances Online Participants' Engagement With a Demanding Auditory Task. *Trends Hear.* **25**, (2021).
126. Boughorbel, S., Jarray, F. & El-Anbari, M. Optimal classifier for imbalanced data using Matthews Correlation Coefficient metric. *PLoS One* **12**, 1–17 (2017).
127. Powers, D. M. W. Evaluation: From Precision, Recall and F-Factor to ROC, Informedness, Markedness & Correlation. *Tech. Rep. SIE-07-001* (2007).
128. Der, G. & Deary, I. J. Age and Sex Differences in Reaction Time in Adulthood: Results From the United Kingdom Health and Lifestyle Survey. *Psychol. Aging* **21**, 62–73 (2006).
129. Michael Lawrence, M. A. Package 'ez': Easy Analysis and Visualization of Factorial Experiments. 4–5 (2016). doi:10.1515/zava.93.1.1
130. Bunton, S. On-line stochastic processes in data compression. (University of Washington, 1996).
131. Cleary, J. & Witten, I. Data compression using adaptive coding and partial string matching. *EEE Trans. Commun.* **32**, 396–402 (1984).
132. Ross, G. J., Tasoulis, D. K. & Adams, N. M. Nonparametric monitoring of data streams for changes in location and scale. *Technometrics* **53**, 379–389 (2011).
133. Rowan, T. Functional stability analysis of numerical algorithms. (1990).
134. Johnson, S. G. The NLopt nonlinear-optimization package. *http://github.com/stevengj/nlopt* (2020).