STATISTICAL LEARNING AMONG YOUNG AND OLDER ADULTS: SIMILAR YET DIFFERENT?

Jia Hoong Ong[1,2] & Alice H. D. Chan[1]

[1]School of Humanities, Nanyang Technological University, Singapore
[2]School of Psychology and Clinical Language Sciences, University of Reading, United Kingdom
jiahoong.ong@reading.ac.uk, alice@ntu.edu.sg

ABSTRACT

While studies have demonstrated that infants, children, and young adults are capable of statistical learning (SL), it is unclear whether the ability is preserved in older adults and if so, whether they might show a decline. The present study investigates this directly by comparing young and older adults on a standard SL task (word segmentation task). Our results indicate that both age groups did not differ significantly in their overall performance. The two groups, however, differed in their performance on the two distractors used in the task. Furthermore, higher working memory was associated with better SL among older adults, but no such association was seen among young adults. Altogether, this suggests that SL ability is preserved among older adults, but they may be using a different strategy/mechanism compared to young adults, presumably due to age-related decline in the relevant brain structures supporting SL.

Keywords: Statistical learning; older adults; ageing; working memory; individual differences

1. INTRODUCTION

Statistical learning (SL), or the acquisition of knowledge based on regularities and patterns in the input, plays an important role in language acquisition and processing. For instance, SL is implicated in learning phonetic categories [16, 11, 20], word segmentation from continuous speech [26, 27], the acquisition of grammatical categories [18, 25], and accent adaptation [13, 15]. Majority of the SL literature focused on infants, children, and young adults, and the findings indicate that they are capable of SL. Importantly, from infants to adulthood, learners do not differ in their SL performance [27], suggesting that SL is age-invariant (though this appears to be modality-dependent [24, 30]).

Comparatively, little is known about SL among older adults. On the one hand, older adults may show a decline in SL due to age-related decline in memory [32], which is hypothesised to be an important component according to several SL models [1, 33, 34]. On the other hand, some suggest that SL is governed by two distinct mechanisms—an implicit and an explicit process [6]—and so older adults may compensate for their SL performance by relying on mechanism that is spared from age-related decline. A thorough understanding of this issue is necessary as it has implications for second language acquisition and language processing among the elderly.

Studies from the implicit learning (IL) literature may shed light on this issue. IL, or the ability to learn without conscious awareness, is an ability that is closely related to SL. Indeed, some suggest that the two abilities may in fact share underlying processes [22]. Older adults are able to perform IL tasks such as the serial-reaction time task (SRTT) to the same degree as young adults, especially when the task contains simple, deterministic patterns. However, older adults show worse performance than young adults when the input contains complex, probabilistic patterns [12]. In other words, IL appears to be preserved among older adults, but age-related deficit is seen under challenging IL conditions.

The few studies that investigated SL among older adults report that the ability is preserved but some age-related differences are seen [4, 19, 21] For example, when participants were presented with two visual streams and were told to only attend to one of the streams while performing a cover task, older adults learned the statistical regularities of both streams whereas young adults only did so for the attended stream [4]. This was argued to be a direct consequence of an age-related decline in inhibitory control; that is, older adults were presumably less able to inhibit their attention to the unattended stream and thus acquire its regularities. In the auditory domain, using a word segmentation task, older adults show worse performance than young adults under challenging or cognitively demanding conditions [21]. Taken together, this suggests that subtle differences in SL may be seen between young and older adults, particularly when the relevant cognitive processes involved in SL show age-related decline.

However, before such a conclusion is reached, it is important to address certain methodological issues of previous studies. In a study that is directly relevant to the present study [21], participants completed the SL task four times, each under different conditions. Even though different languages were used, and the order of the conditions was counterbalanced, expectations may arise as a result of completing the tasks multiple times. Furthermore, there were only four words in each language, and each word was repeated 108 times during exposure. Thus, the relatively easy learning conditions may have masked other potential age-related differences. In the present study, young and older participants completed one SL task with six words, with each word presented less frequently (32 times) during exposure, in order to determine whether learning still occurs under more challenging conditions, and if so, whether there might be differences between the two age groups.

As a secondary question, we investigate whether working memory (WM) may be involved in SL as equivocal findings have been reported in the literature [21, 31]. This may be a result of the different WM tasks used, which may tap into different aspects of WM. Here, we seek to shed light on this issue by using a WM task that has not been used in the literature that involves both storage (as would be the case with certain WM tasks such as digit span used in previous studies) and manipulation of the stored information.

2. METHOD

2.1. Participants

Twenty young adults (14 females) and 20 older adults (12 females) participated in the experiment. The young adults (Age: M= 23.15, SD= 2.94, Range= 20-29) were undergraduates from a local university whereas the older adults (Age: M= 67.90, SD= 6.18, Range= 60-87) were recruited from the community. All the older participants were cognitively healthy at the time of testing, as indexed by their Mini Mental State Examination-2 (MMSE-2) scores (Range= 27-30). All participants provided their written informed consent prior to participating in the research and they were reimbursed for their participation.

2.2. Tasks

2.2.1. Word segmentation

The task was modelled after previous studies [24, 27]. The language consists of six trisyllabic CVCVCV pseudowords (*mubita, pitinu, nalubu, gupala, bapugi,* and *lituma*) that were synthesised using Mac OS X Speech Service with a female voice. For each syllable, its duration was edited to approximate 280ms, its pitch contour was levelled at 170Hz, and its amplitude was normalised at 70dB.

The task consisted of two parts: exposure and test. In the exposure phase, participants were instructed to listen to a monotone language and to try and pick out the words. The exposure speech stream consisted of the six trisyllabic words concatenated with no pauses and in a pseudorandomised order such that no two consecutive words were the same. Thus, the transitional probability (TP) between syllables within a word is 1.0 whereas TP across words is 0.2. Each word appeared on the exposure stream 32 times in total, with the total exposure period being not more than 3 mins in duration. Participants then completed a two-alternative forced choice test: in every trial, a word and a distractor were presented, separated by a 500ms pause and their presentation order counterbalanced. Participants indicated which of the two sounded like a word from the monotone language and they were encouraged to guess if they were unsure. There were six distractors in total: three part-words and three nonwords. Part-words were formed by concatenating a final syllable from a word and the first two syllables of another word (e.g., *bulitu* from *nalu*<u>*bu*</u> + <u>*litu*</u>*ma*) whereas nonwords consisted of syllables that never followed each other in the exposure stream (e.g., *mutila* from <u>*mu*</u>*bita* + *pi*<u>*ti*</u>*nu* + *gupa*<u>*la*</u>). Each word was paired with a distractor exhaustively, resulting in 36 test trials in total. Accuracy was coded as 1 (correct) and 0 (incorrect).

2.2.2. Working memory

Working memory (WM) was measured using the Auditory Working Memory subtest from Woodcock-Johnson Tests of Cognitive Abilities (3rd Edition). In each trial, participants heard a list of objects and digits. At the end of each list, participants recalled the objects first, in the order that the objects were presented, and then the digits, in the order that the digits were presented. Thus, this task requires participants to store and manipulate the words in the list. The number of words in each list ranged between three and eight. For each list, two full points were awarded for correct recall of objects and numbers. One point was given when they correctly recalled either the objects or the numbers.

2.3. Procedure

Participants completed the tasks as part of a larger study on language learning and cognition. The tasks were completed within a single session and the word segmentation task was always presented first.

3. RESULTS

Following previous studies, we first analysed the data using a series of t-tests. For nonword trials, young adults ($t(19)= 4.82$, $p< .001$) and older adults ($t(19)= 2.45$, $p= .024$) performed above chance, and two groups did not differ significantly from each other (*t*(38)= 1.58, *p*= .122). For part-word trials, young adults did not perform above chance (*t*(19)= 0.07, *p*= .94) whereas older adults did (*t*(19)= 3.27, *p*= .004), but the two groups did not differ significantly from each other (*t*(38)= -1.91, *p*= .063).

Data were then analysed using mixed effects logistic regression models using *lme4* package [10] in R [23]. Pairwise comparisons were done using *lsmeans* package [14]. The dependent variable was Accuracy, a binary outcome. The predictors Distractor (Nonword/Part-word) and Age Group (Young adults/Older adults) were effect-coded whereas WM was centred by Age Group. All predictors as well as all possible interactions were included in the model. The random structure included random intercepts by subject and by item as well as random slopes for Distractor by participant and Age Group by item. The model output is displayed in Table 1.

Table 1: Model output of the mixed effects logistic regression.

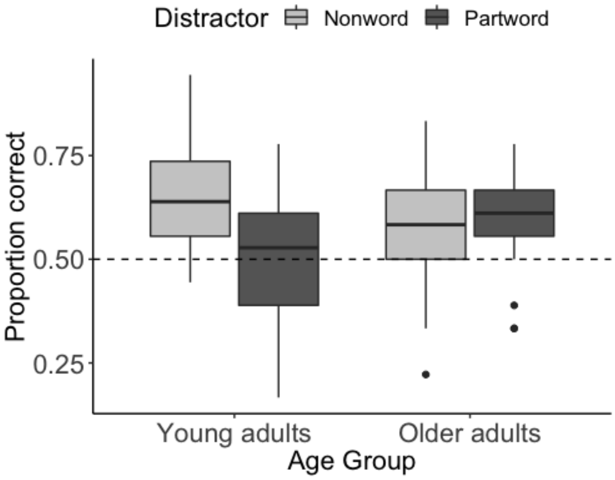
Predictors	Estimated ß (SE)	*z*-value
(Intercept)	0.36 (0.09)	3.98***
Distractor	-0.15 (0.09)	-1.79†
Age Group	0.02 (0.07)	0.27
WM	0.05 (0.02)	3.14**
Distractor × Age Group	0.17 (0.07)	2.57*
Distractor × WM	0.00 (0.02)	0.25
Age Group × WM	0.01 (0.02)	0.75
Distractor × Age Group × WM	-0.02 (0.02)	-1.05

Note: *** *p*< .001, ** *p*< .01, * *p*< .05, † *p*< .1

We found a marginal main effect of Distractor, which was qualified by a Distractor × Age Group interaction (see Figure 1). Among young adults, their performance on nonword test trials were significantly better than that on part-word test trials (ß$_{diff}$= 0.65, SE= 0.21, *p*= .002) whereas no difference between the distractors were seen among older adults (ß$_{diff}$= -0.04, SE= 0.23, *p*= .867). We also found a significant main effect of WM such that the higher their WM, the better their performance on the task. (Note: the same pattern of results was obtained when the outliers seen in Figure 1 were removed.)

We modelled the data separately for each Age Group to verify the results of the interaction. For young adults, only a main effect of Distractor was found (ß= -0.33, SE= 0.11, *p*= .004) such that nonword trials were performed better than part-word trials. No main effect of WM was found (ß= 0.04, SE= 0.03, *p*= .127). For older adults, there was no main effect of Distractor (ß=0.02, SE= 0.11, *p*= .866), confirming the interaction seen in the full model. The only significant predictor was WM (ß= 0.06, SE= 0.02, *p*= .002). This suggests that the main effect of WM in the full model is driven predominantly by the older adults.

Figure 1: Proportion correct on the word segmentation task by Age Group and Distractor Type. The dashed line represents chance performance.



4. DISCUSSION

The present study investigated whether there is an age-related decline in statistical learning (SL) and whether working memory (WM) is associated with SL. Our findings indicate that SL was preserved among older adults and they showed similar overall SL performance to young adults. However, some subtle differences were found between the two age groups. Specifically, (i) older adults did not show the expected difference between nonword distractor trials and part-word distractor trials seen among young adults; and (ii) WM appears to be associated with SL performance among older adults and not young adults (though it should be noted that an Age Group × WM interaction was not significant in the full model and so this should be interpreted with caution). These two findings are discussed in more detail below.

4.1. Differences in the distractor type performance

Consistent with previous studies, young adults were more accurate on nonword distractor trials than part-word distractor trials [27]. This finding is unsurprising given that the differences in TPs between words and nonwords are larger than that between words and part-words. In other words, the nonword trials are objectively ‘easier’ than the partword trials, at least if one were to approach the test trials using TPs. Older adults, however, did not show a difference between the two distractor trials, which suggests that older adults may be using a different strategy/mechanism to complete the task. While unexpected, similar findings were reported in memory research: in these studies, adults were shown images that belong to certain categories (e.g., cats, books, etc.) and then later they were shown a series of images and they had to decide whether they have seen the images before. At test, the images could be old (i.e., an old image of a cat seen during the exposure phase), new but from a related category (i.e., a new image of a cat), or new from an unrelated category (e.g., an image of an umbrella). Whereas young adults showed the expected differentiation in behavioural and neural responses to the new-related and new-unrelated lures, older adults did not [2], suggesting that the two lures were processed similarly by older adults.

Several explanations have been offered to explain the differences in performance between young and older adults in memory research [9]. Two such are: (i) older adults engage in hyper-binding; and/or (ii) older adults tend to use gist-based processing. We argue that these explanations are similarly applicable to our findings. Hyper-binding, or the encoding of irrelevant associations in the input, appears to be more prevalent among older adults, presumably due to the deterioration of their inhibitory control [5]. In our case, it may be that the older adults bound irrelevant information (i.e., syllables) across the entirety of the language such that nonwords no longer have zero TPs, resulting in those distractors to be similar in difficulty as part-words.

Another explanation relates to the tendency of older adults to use gist-based processing. This tendency may result from older adults encoding the input more poorly than young adults, leading to a weaker memory trace of the input [8]. As a result, older adults are said to rely on a sense of ‘familiarity’ in making a memory judgment as opposed to recalling item-specific memory traces. In our case, it may be that the older adults were processing the test items based on a holistic, familiarity judgment rather than on the exact differences in TPs between the syllables, which would require retrieving item-specific traces of the syllables encountered during the exposure phase. This is not to say that older adults do not use TPs – they clearly do, as they are still able to discriminate words from the distractors—but rather, their use of TPs for discrimination judgment is not as fine-grained as that of young adults. This lack of fine-grained discernment was similarly reported in an artificial grammar learning task [29], in which older adults show a smaller difference than young adults in their grammaticality ratings between grammatical and ungrammatical strings. Thus, consistent with our interpretation, older adults in that study were able to utilise statistical information to learn the overall grammar but they were less sensitive to the grammar than young adults.

4.2. Difference in the involvement of working memory

Our results also indicate that young and older adults differed in how WM was associated with their SL performance; specifically, there was no association between WM and SL among young adults, but there was a positive association for older adults. This suggests that older adults may be using a different strategy than young adults when performing the SL task. While the current study does not allow us to differentiate *what* strategy was used by the older adults, we could speculate *why* they may use a different strategy. The segmentation task is said to be supported by the MTL, in particular, the hippocampus [28, 35]. Given that these areas are prone to age-related decline, we speculate that older adults engage their PFC as a compensatory mechanism when performing the task. This view is consistent with several cognitive ageing hypotheses, in which it is argued that that age-related decrease in one area of the brain is compensated by an increase in the (pre)frontal area [3, 7]. Given that the PFC is implicated in complex cognition such as executive functions and cognitive control [17], the WM task in the present study may thus be tapping into the same processes involved in the PFC, which would explain the positive relationship between WM and SL for the older adults. However, the interaction between Age Group × WM in the full model was not significant (which may presumably be due to lack of statistical power) and thus this pattern of results should be interpreted with caution.

5. CONCLUSION

Our preliminary results indicate that there are similarities and differences between young and older adults in their SL ability. While the two groups did not differ in their overall SL performance, older adults did not show the expected difference in performance among different distractor trials that was seen among young adults. Furthermore, when the results were analysed separately, a significant positive association between WM and SL was seen among older adults but not among young adults. We propose that these subtle differences between young and older adults are due to a change in strategies/mechanisms to perform the SL task, which, we argue, result from a shift from using the MTL to the PFC to perform the task. Further studies are needed to corroborate this claim and to determine precisely the strategy used by older adults when engaging in SL.

6. REFERENCES

[1] Arciuli, J. 2017. The multi-component nature of statistical learning. *Philos. Trans. R. Soc. B* 372, 20160058

[2] Bowman, C. R., Dennis, N. A. 2015. Age differences in the neural correlates of novelty processing: The effects of item-relatedness. *Brain Res.* 1612, 2–15

[3] Cabeza, R. 2006. Prefrontal and medial temporal lobe contributions to relational memory in young and older adults. In: Zimmer, H., Mecklinger, A. & Lindenberger, U. (eds) *Handbook of Binding and Memory: Perspectives from Cognitive Neuroscience.* Oxford Scholarship Online, 595–626.

[4] Campbell, K. L., Healey, M. K., Lee, M. M. S., Zimerman, S., Hasher, L. 2012. Age differences in visual statistical learning. *Psychol. Aging* 27, 650–656

[5] Chao, L. L., Knight, R. T. 1997. Prefrontal deficits in attention and inhibitory control with aging. *Cereb. Cortex* 7, 63–69

[6] Daltrozzo, J., Conway, C. M. 2014. Neurocognitive mechanisms of statistical-sequential learning: What do event-related potentials tell us? *Front. Hum. Neurosci.* 8, 1–22

[7] Davis, S. W., Dennis, N. A., Daselaar, S. M., Fleck, M. S., Cabeza, R. 2008. Qué PASA? the posterior-anterior shift in aging. *Cereb. Cortex* 18, 1201–1209

[8] Dennis, N. A., Bowman, C. R., Peterson, K. M. 2014. Age-related differences in the neural correlates mediating false recollection. *Neurobiol. Aging* 35, 395–407

[9] Devitt, A. L., Schacter, D. L. 2016. False memories with age: Neural & cognitive underpinnings. *Neuropsychologia* 91, 346–359

[10] Douglas, B., Maechler, M., Bolker, B., Walker, S. 2015. Fitting linear mixed-effects models using lme4. *J. Stat. Softw.* 67, 1–48

[11] Escudero, P., Williams, D. 2014. Distributional learning has immediate and long-lasting effects. *Cognition* 133, 408–13

[12] Howard, J. H. J., Howard, D. V 2013. Aging mind and brain: Is implicit learning spared in healthy aging? *Front. Psychol.* 4, 817

[13] Kleinschmidt, D. F., Jaeger, T. F. 2015. Robust speech perception: Recognize the familiar, generalize to the similar, and adapt to the novel. *Psychol. Rev.* 122, 148–203

[14] Lenth, R. V 2016. Least-Squares Means: The R Package lsmeans. *J. Stat. Softw.* 69, 1–33

[15] Maye, J., Aslin, R. N., Tanenhaus, M. 2008. The Weckud Wetch of the Wast: Lexical adaptation to a novel accent. *Cogn. Sci.* 32, 543–562

[16] Maye, J., Werker, J. F., Gerken, L. 2002. Infant sensitivity to distributional information can affect phonetic discrimination. *Cognition* 82, B101–B111

[17] Miller, E. K., Cohen, J. D. 2001. An integrative theory of prefrontal cortex function. *Annu. Rev. Neurosci.* 24, 167–202

[18] Mintz, T. H., Wang, F. H., Li, J. 2014. Word categorization from distributional information: Frames confer more than the sum of their (bigram) parts. *Cogn. Psychol.* 75, 1–27

[19] Neger, T., Rietveld, T., Janse, E. 2015. Adult age effects in auditory statistical learning. *Proc. 18th Int. Congr. Phonetic Sci.*

[20] Ong, J. H., Burnham, D., Escudero, P. 2015. Distributional learning of lexical tones: A comparison of attended vs. unattended listening. *PLoS One* 10, e0133446

[21] Palmer, S. D., Hutson, J., Mattys, S. L. 2018. Statistical learning for speech segmentation: Age-related changes and underlying mechanisms. *Psychol. Aging*

[22] Perruchet, P., Pacton, S. 2006. Implicit learning and statistical learning: One phenomenon, two approaches. *Trends Cogn. Sci.* 10, 233–8

[23] R Core Team 2017. R: A language and environment for statistical computing.

[24] Raviv, L., Arnon, I. 2017. The developmental trajectory of children's auditory and visual statistical learning abilities: Modality-based differences in the effect of age. *Dev. Sci.* e12593

[25] Reeder, P. A., Newport, E. L., Aslin, R. N. 2013. From shared contexts to syntactic categories: The role of distributional information in learning linguistic form-classes. *Cogn. Psychol.* 66, 30–54

[26] Saffran, J. R., Aslin, R. N., Newport, E. L. 1996. Statistical learning by 8-month-old infants. *Science.* 274, 1926–1928

[27] Saffran, J. R., Newport, E. L., Aslin, R. N. 1996. Word segmentation: The role of distributional cues. *J. Mem. Lang.* 35, 606–621

[28] Schapiro, A. C., Turk-Browne, N. B. 2015. Statistical learning. In: Toga, A. W. (ed) *Brain Mapping: An Encyclopedic Reference.* Elsevier, 501–506

[29] Schwab, J. F. *et al.* 2016. Aging and the statistical learning of grammatical form classes. *Psychol. Aging* 31, 481–487

[30] Siegelman, N., Bogaerts, L., Elazar, A., Arciuli, J., Frost, R. 2018. Linguistic entrenchment: Prior knowledge impacts statistical learning performance. *Cognition* 177, 198–213

[31] Siegelman, N., Frost, R. 2015. Statistical learning as an individual ability: Theoretical perspectives and empirical evidence. *J. Mem. Lang.* 81, 105–120

[32] Stuart-Hamilton, I. 2012. *The Psychology of Ageing: An Introduction*. Jessia Kingsley Publishers.

[33] Thiessen, E. D. 2017. What's statistical about learning? Insights from modeling statistical learning as a set of memory processes. *Philos. Trans. R. Soc. B* 372, 20160056

[34] Thiessen, E. D., Pavlik, P. I. 2013. iMinerva: A mathematical model of distributional statistical learning. *Cogn. Sci.* 37, 310–43

[35] Turk-Browne, N. B., Scholl, B. J., Chun, M. M., Johnson, M. K. 2010. Neural evidence of statistical learning: Efficient detection of visual regularities without awareness. *J. Cogn. Neurosci.* 21, 1934–1945