

日本語の動詞流暢性課題における年齢の影響

The Effects of Age on the Japanese Verb Fluency Task

古澤 里菜 FURUSAWA, Rina

● 国際基督教大学
International Christian University

李 勝勲 LEE, Seunghun J.

● 国際基督教大学, ベンダ大学
International Christian University, University of Venda

遠藤 歩華 ENDOH, Honoka

● 国際基督教大学
International Christian University

吉田 智行 YOSHIDA, Tomoyuki

● 国際基督教大学
International Christian University

有井 巴 ARII, Tomoe

● 横浜市立大学
Yokohama City University

 **Keywords** 言語流暢性課題, 日本語, 年齢, クラスター, スイッチ
Verb Fluency task, Japanese, aging, cluster, switch

ABSTRACT

言語流暢性課題は、語彙検索の能力を測るために使われるスクリーニングテストである。本研究は、日本語における若年層と老年層の動詞における言語流暢性課題の基準を示すことを目的とする。61人の日本語母語話者の対する実験結果は、若年層よりも老年層のほうが少ない動詞を発することが示された。クラスタ分析では、クラスタとスイッチの数は両者ともに老年層のほうが少ないが、クラスタの大きさには差が観察されなかった。

Verb fluency task is a screening test that shows sensitivity to lexical retrieval abilities. This study aims to establish a baseline for verb fluency in Japanese by comparing data from 61 younger and elderly Japanese speakers. The results show that the elderly group produced fewer verbs than the younger group, but compared to Lee et al. (2013), many more correct responses were produced by both groups. An analysis using the Lancaster Sensorimotor Norms found that the number of clusters and switches was smaller for the elderly participants, although the cluster size did not show any difference.

1. Introduction

A Fluency task is a test that requires participants to produce as many items as possible in a specified time, usually under one minute. It is widely used for clinical purposes since it is sensitive to lexical retrieval abilities (Ito & Hatta, 2006; Sauzéon et al., 2004). Fluency tasks are conducted in various categories, such as naming animals, flowers, or verbs. In particular, the Verb Fluency task focuses on how participants retrieve verbs from their lexicon. Most studies on verb fluency have targeted languages that do not allow null pronouns, such as English (Alegret et al., 2018; Piatt et al., 1999; Woods et al., 2005), and there has been little work on languages that allow null pronouns, such as Japanese and Korean (Iwasaki et al., 2008; Yoo et al., 2023). Languages such as Japanese are verb-oriented, as verbs in isolation can be used as a single sentence. For example, the single-word phrase *tabetayo* “(I) already ate (a meal)” excludes both the overt subject and the overt object of the verb *taberu* “to eat” (Sakamoto, 2016; Takahashi, 2008). The productivity of verb-only sentences in colloquial speech suggests that differences in the results between speakers of non-pro-drop languages and pro-drop languages may emerge in the lexical retrieval test focusing on verbs.

The focus on the clinical population in previous studies means that there has been limited research conducted on healthy elderly speakers. Most known studies compared young individuals with clinical populations (Adams et al., 1989; Bose et al., 2017; Kim et al., 2011), introducing two fixed variables: age and clinical status. Young and healthy individuals were compared with elderly clinical individuals¹. This trend suggests that it is not straightforward to immediately determine which factor is responsible for the observed differences in lexical retrieval abilities. An earlier study that explored verb fluency tasks in Japanese was Lee et

al. (2013), which examined both young and elderly groups and reported that the elderly group produced fewer verb items while producing a higher number of errors. Reports on production abilities have shown age-related effects, such as difficulties for elderly populations to recall well-known words (Burke et al. 1991; James & Burke 2000; Shafto et al. 2007). Kemper (2012) analyzed the underlying mechanism for such age-related effects and concluded that the age difference in the temporal buffer reflects that of working memory.

The goal of this paper is to contribute additional empirical data on the differences and similarities between healthy young and elderly participants, comparing with results reported in Lee et al. (2013)². Our results found a smaller number of items and a larger number of errors for elderly participants than younger participants, which replicates the result of Lee et al. (2013). This study adds three new aspects based on the results of the verb fluency task. First, we report the verbs produced by the participants, which were not reported in previous studies. We expand our analysis by focusing on the verbs produced by the participants. Second, the results of the analyses of switch count and cluster count across the participant groups are reported. Troyer (2000)³ proposes these analyses because the clustering analysis demonstrates how participants retrieve lexical items that belong to a subcategory, and the switching analysis shows how participants switch subcategories in their retrieval process when a said subcategory has been exhausted (p. 370). Troyer (2000) considers the clustering of lexical items an automatic process during the lexical retrieval, whereas the switching of subcategories is a process that requires effort by participants because it involves “cognitive flexibility” when participants shift between subcategories. When a specific subcategory of verbs within a lexicon is activated, participants will produce verbs that are

semantically similar to those previously produced.

In this study, verbs are classified based on Lancaster sensorimotor norms (Lynott et al., 2020). These norms were established by measuring the sensory strength of 39,707 verbs. The number of participants who provided their judgment was 3,500 people, a huge number considering the usual perception studies. To this date, the Lancaster sensorimotor norms are the largest scale of norm measures for classifying the sensory strengths of verbs. If Troyer's conjecture is on the right track, we hypothesize that the elderly group will produce fewer switches: if age affects the ability to perform lexical retrieval, they would not be able to perform as many "effortful processes" as the younger group. In contrast, the process of creating clusters is automatic, which suggests that the number of clusters would show no difference between the elderly and the young participants. Contrary to our expectations, however, this prediction is not borne out in this study.

Finally, we perform an analysis of the divided time segments within the one-minute task. The verb fluency task is conducted for 60 seconds. Since the number of clusters is expected not to be different between the two age groups, we analyzed production patterns during the first 30 seconds and the last 30 seconds. We hypothesize that the distribution of verb items produced within the 60 seconds is not necessarily uniform: the first half typically would have a larger number of items than the last half.

The structure of this paper is as follows. In section 2, we describe the data collection methods of the verb fluency task, which was conducted as part of a larger study. Section 3 reports various results obtained from the data by comparing the younger group and the elderly group, which is followed by a discussion section. This paper aims to augment existing studies by adding how aging affects the results of the verb fluency task, which

has been understudied⁴.

2. Methods

2.1 Participants

We recruited 62 native speakers of Japanese by balancing the number of younger participants and the elderly participants (31 in each group). In the younger group (age average = 21, $SD = 1.31$), 23 participants were female, and all of them graduated from high school or university (average year of education = 14.7, $SD = 1.25^5$). We excluded data from one elderly participant because he self-reported experience with a clinical issue. The age of the remaining thirty elderly participants (19 of them female) ranges from 67 to 90 years old (60s: $n = 2$, 70s: $n = 15$, 80s: $n = 12$, 90s: $n = 1$). Their educational background ranges from junior college to graduate school (average years of education = 13.8, $SD = 3.16$).

2.2 Data Collection Procedure

The verb fluency task was part of a larger set of experiments, lasting approximately 30-90 minutes. All parts of the experiment were recorded via Zoom or a smartphone for transcription of the produced data. The experiments were conducted over a period of one year. Due to the COVID-19 situation, the experiment in 2023 summer was conducted online with 12 elderly and 10 young speakers. Data from the rest of the participants (18 elderly and 21 younger speakers) were collected in an in-person experiment.

In the verb fluency task, participants were asked to produce as many verbs as possible in 60 seconds. The preparatory instruction included a 15 seconds-trial test in which participants were asked to produce as many fruit names as possible. In the online session, some participants did not understand the meaning of the word "verb (*doushi* in Japanese)." As such, the remaining sessions

offered explanations about what a verb is and provided participants with three examples of verbs: *haku* “sweep,” *nomu* “drink” and *ageru* “give.” After completing all the tasks in the experiment, participants were compensated for their time.

2.3 Post-experiment Processing

Recordings were manually transcribed by the first three authors. Some data was transcribed with the help of the online-based Clipto app that offers speech-to-text transcription (<https://www.clipto.com/>). The data were then translated into English and categorized into six perceptual modalities and five action effectors following the Lancaster sensorimotor norms (Lynott et al., 2020). The perception modalities classify verbs into types of sensory motors as follows: Visual, Auditory, Haptic, Gustatory, Interoceptive, and Olfactory. The action effectors classify verbs by body parts that perform a verbal action, such as Mouth, Foot_leg, Head, Hand_arm, and Torso. Lancaster’s norms of 39,707 verbs were accessible from the website (<https://www.lancaster.ac.uk/psychology/l norms/>), with each verb being assigned the dominant perception modality based on the highest mean rating. Using the same method, verbs were also assigned with dominant action effectors (Lynnott & Connell, 2009). We automatically categorized the Japanese verbs produced by the participants using the VLOOKUP function on the Google spreadsheet, which prevented errors from being included in the norm coding.

The first author performed further data processing: (a) to facilitate the cluster analysis, all pairs that were uttered by the same speaker in a sequence were classified into the same cluster, and (b) when the subcategory of a sequence of two verbs was different, the following verb was marked to identify the location of the cluster switch. The number of items, clusters, cluster switch, and cluster size (how many items exist in one cluster)

were calculated for each participant for both perceptual and action clusters, following the method introduced in Troyer (2000). The calculation itself was automated using Google spreadsheet functions. The process is illustrated using a list produced by participant 003 in Table 1. The first ten verbs produced by this speaker are listed in the second column. The action norms corresponding to the verbs and the clusters are shown in the third and fourth columns. The data in Table 1 has two clusters under the norm “Head”: the first cluster has *read*, *listen*, *become open*, and the second cluster has *lost*, *listen*. Switches occur when there is a change in the action norm. In Table 1, six switches are observed: from Mouth (Nr. 1) to Foot_leg (Nr. 2), then to Head (Nr. 3), then to Foot_leg (Nr. 6), then to Head (Nr. 7), then to Hand_arm (Nr. 8), and to Head (Nr. 9).

Participants also produced items that were treated as errors. The exclusion criteria included items that do not belong to verbs; for example, the item *kakekko* “running competition” is a noun. When a participant produced the same verb more than once, only the first occurrence of the repeated verbs was counted. We also excluded verbs that were identical to the three examples (*haku* “sweep,” *nomu* “drink” and *ageru* “give”). The 61 participants produced 1271 items with 139 errors (10.9 %).

Table 1
First Ten Verbs Produced by Participant 003

Order	Verbs	Action norms	Cluster	Switch
1	eat	Mouth		
2	run	Foot_leg		Switch 1
3	read	Head	Cluster 1	Switch 2
4	listen	Head	Cluster 1	
5	become open	Head	Cluster 1	
6	go	Foot_leg		Switch 3
7	go home	Head		Switch 4
8	shut	Hand_arm		Switch 5
9	close	Head	Cluster 2	Switch 6
10	listen	Head	Cluster 2	
... (more items)				

3. Results of the Verb Fluency Task by Japanese Speakers

3.1 Descriptive Results

This section reports the type of verbs that are being produced by the participants. The verbs produced by the younger group are listed in Table 2. More than 30% of the participants produced verbs in Table 2a, and fewer participants produced verbs in Table 2b. Most of these verbs belong to fairly high-frequency verbs.

Verbs that were retrieved by the elderly participants are listed in Table 3. Overall, the

elderly participants produced fewer words, but it is also the case that fewer verbs are produced commonly by the elderly participants. Compared to 16 verb types produced by ten or more younger speakers, the elderly group shows only 10 of such verbs, which demonstrates that more diverse types of verbs were produced by the elderly participants.

Table 4 compares the top 10 verbs produced by the younger group and the elderly group. Participants in both generations have identical three words in the top 3: *hashiru* “run,” *aruku* “walk,” and *kiku* “listen.”

Table 2

Verbs (Translated from Japanese) Used by More Than One Speaker in the Young Group (n = 31)

a. Produced by ten or more participants (16 verbs)

run (27)	walk (23)	listen (20)	look (19)	eat (18)	write (18)
fly (16)	talk (15)	sit (14)	uncover (13)	throw (13)	read (12)
shut (10)	knock (10)	kick (10)	move (10)		

b. Produced by less than 10 participants

(9) cut, touch	(8) punch, join, go to bed, get up, stand, strike, hit	(7) put on
(6) close, transfer, pull, swim, make, have	(5) go, laugh, go up, sing, scream, say, speak, cry, get out	(4) wear, get off, take, do skip, enter, get out, angry/happen, think, grasp
(3) bloom, bounce, inhale, taste, make it break, grab, stab, erase, lower, buy, draw, wash,		
(2) do jump, drop something, look for, scold, kill, drink, connect, discard, pull/stretch, peck, dance, yell, do cook, stick, polish, bake, fall, put out, squat, grieve, fight, slip, put, chew, feel, smell, hand over, get		

Table 3

Verbs (Translated from Japanese) Used by More Than One Speaker in the Elderly Group (n = 30)

a. Produced by ten or more participants (10 verbs)

run (19)	walk (15)	listen (14)	fly (14)	move (12)	swim (11)
eat (11)	talk (10)	go to bed (10)	look (10)		

b. Produced by less than 10 participants

(8) stand	(7) sit, read, go up, slip	(6) write, get on, cry, get up
(5) knock, sleep, laugh, cut	(4) angry/happen, kick, go, put it away, get off, put on, play, wipe, put out, sing, chew	
(3) take, throw, stop, bounce, wash, learn, think, touch, go down, close		
(2) receive, do driving, change, smell, shut, uncover, discard, exist, rest, do(yaru), polish, wear, arrive, squat, dive, stretch, be pleased, bend, make it stretch, lick, scoop, put between, do studying, pat, touch on		

Other verbs in Table 4 that are produced by a number of participants are *tobu* “fly,” *taberu* “eat,” *miru* “see,” and *hanasu* “talk.” About half of the participants in the younger group produced *kaku* “write,” *suwaru* “sit,” and *akeru* “uncover,” which were produced by fewer participants in the elderly group. More than 30% of the elderly participants produced *ugoku* “move,” *oyogu* “swim,” and *neru* “go to bed,” all of which were produced only by a limited number of participants.

The result of a *t*-test demonstrates that the average number of verbs produced by participants in each group is significantly different ($p < .001$) between the two groups, with or without the error items, see Table 5. On average, the younger group produced more verbs than the elderly group. The younger group produced fewer items (45 tokens, 6.17%; on average 1.45) that are counted as errors than the elderly group (94 tokens, 17.34%; on average, 3.13). The error rate was lower in young speakers than in elderly speakers (cf. Lee et al., 2013).

Elderly participants in this study had an age distribution between 60s and 90s, which is wider than the age range (65 - 79) of participants in Lee et

al. (2013). The early elderly group (those who are under 80, $n = 17$) and the late elderly group (over 80, $n = 13$) did not show significant differences in the number of all items produced by each participant ($p = .16$). When the errors were excluded from the count, the young elderly group produced 16.59 verbs ($SD = 6.59$), whereas the older-elderly group produced 12.77 verbs ($SD = 3.92$). A marginal difference ($p = .06$) between the two groups was found, suggesting that the late elderly group produced more errors than the early elderly group. Other measures reported in later sections did not show any significant differences between these two groups. As such, we report the remaining results without dividing the elderly group.

Perceptual modalities and action effectors based on the Lancaster Sensorimotor Norms (Lynott et al., 2020) were extracted using the web-based interface⁶. Iida & Akita (2023) report the perceptual strength norms of 510 Japanese words that include ideophones and other lexical items, but in the absence of large-scale sensorimotor norms for Japanese verbs, the translation from Japanese to English was adopted in this study. Results of J. Lee & Shin (2023) suggest that the use of English

Table 4
Top 10 Verbs Based on the Production by Elderly and Younger Participants ($n = 61$). Highlighted Cells Indicate Verbs Produced by Both the Younger and the Elderly Groups. Error Tokens are Excluded from this Table

Rank	Elderly	Younger	Rank	Elderly	Younger
1	run (19)	run (27)	6	swim (11)	write (18)
2	walk (15)	walk (23)	7	eat (11)	fly (16)
3	listen (14)	listen (20)	8	talk (10)	talk (15)
4	fly (14)	look (19)	9	go to bed (10)	sit (14)
5	move (12)	eat (18)	10	look (10)	uncover, throw (13)

Table 5
The Number of Items Without Errors

	Average number of verbs without errors	Average number of verbs produced
younger group	22.06 ($SD = 4.59$)	23.52 ($SD = 5.21$)
elderly group	14.93 ($SD = 5.83$)	18.07 ($SD = 6.02$)
t-test	$p < .00001$	$p < .001$

translations provides comparable results, suggesting that the use of English translation is a possible alternative prior to creating sensorimotor norms of close to 40,000 verbs.

The results of verbs classified by perceptual modalities are shown in Figure 1. For the analysis, each verb is counted once based on its dominant perceptual modality. The younger group and the elderly group show comparable patterns in retrieving verbs that share the same perceptual modality. Most verbs produced by the participants share the visual modality ($n = 499$ in the younger group, $n = 364$ in the elderly group). Distant second and third modalities are auditory and haptic. Almost no verb that is dominantly olfactory is produced.

While both younger and elderly groups show a strong bias toward producing verbs that have visual modality, the dominant action strength does not always show parallel patterns between the two groups. Participants in the younger group tend to produce verbs that involve action using Hands or

Arms ($n = 270$), which is followed by actions using Head ($n = 211$) as well as Foot and Leg ($n = 142$). The dominant action strength of the verbs produced by the elderly participants displays different patterns. Verbs that create an action using Head ($n = 194$) are produced more than other actions, such as Hand_arm ($n = 145$) and Foot_leg ($n = 129$)⁷. Both groups also produced a number of verbs that are dominantly performed using Mouth ($n = 99$ for the younger group; $n = 69$ for the elderly group). Participants in the current study barely produced verbs that dominantly use Torso as shown in Figure 2.

The information about perceptual modalities and action strength corresponding to each verb provides us with tools to analyze the results of the verb fluency task. Perceptual modalities show an overwhelming majority of verbs produced with Visual, which would result in the creation of a lower number of clusters and fewer switches; if most tokens belong to the Visual modality, most responses will be analyzed as having clusters with

Figure 1
Number of Tokens (Including Errors) Classified by Perceptual Modalities Using Lancaster Sensory Norms

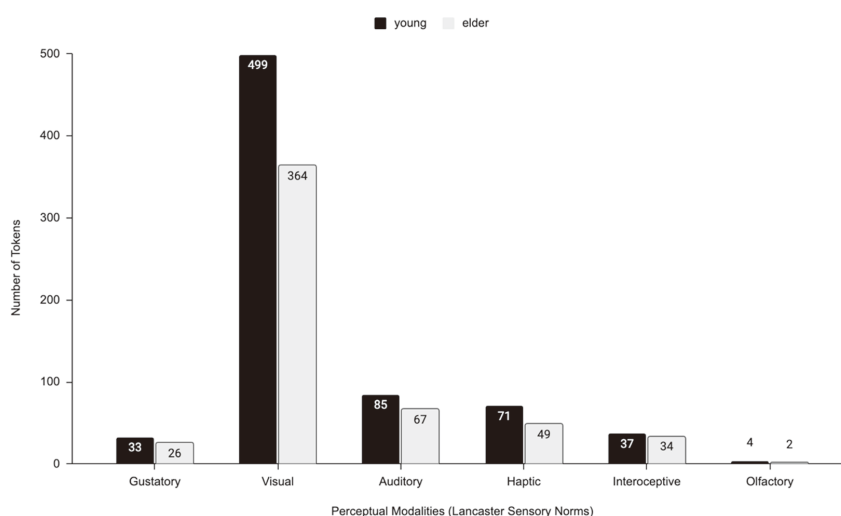
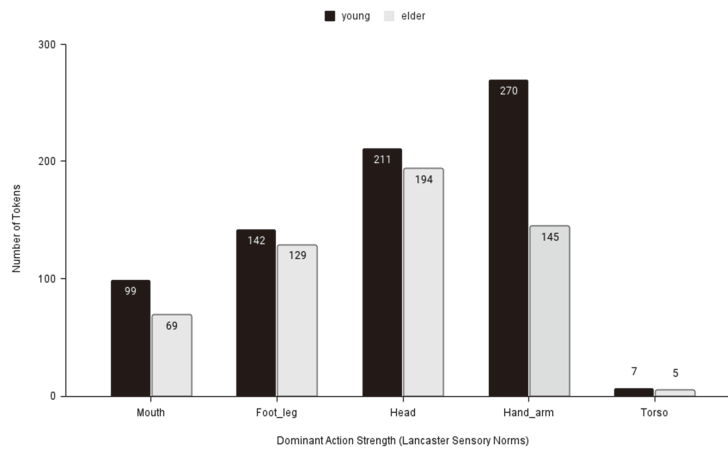


Figure 2

Number of Tokens (Including Errors) Classified by Dominant Action Strength Using Lancaster Sensory Norms



the Visual modality. As such, the cluster and switch analyses in section 3.2 will utilize the results of the dominant action strength in calculating clusters and switches in our data set, which will provide a step into understanding the patterns of lexical retrieval by age group.

3.2 Cluster Size, Number and Switch

Recall that the verb fluency task is a lexical retrieval task within 60 seconds, during which participants are asked to produce as many verbs as possible. The set of verbs produced by a participant may look random, but analyzing the number of clusters, the number of switches, and the cluster size based on action strength offers a window into understanding how participants access their lexicon to retrieve target words. Additionally, we report the switch rate and the cluster rate, which refers to the number of switches or clusters divided by the total number of items produced by a participant. Troyer (2000), however, cautions that the number of switches is one of the factors that determine the total number of items uttered, suggesting that switch rate “would be tantamount to correcting cause for its effect” (p372). Thus, the primary

comparison between the two age groups is based on the raw counts of clusters and switches.

In Table 6, the action strength results are shown. Both the number of clusters and the number of switches are higher in the younger group than in the elderly group. A two-tailed *t*-test shows that the difference between the two groups is significant. As reported in Table 5, participants in the younger group produced more verbs (avg. 22.06) than the participants in the elderly group (avg. 14.93). The difference in the number of items produced is independent of the number of clusters, suggesting that the younger group actively uses the clustering strategy by semantic fields (i.e., action strength) when retrieving verbs from their lexicon. The elderly group shows less use of this clustering. The younger group produced more clusters, but the size of the clusters between the two age groups is not significantly different, which aligns with the difference in the number of switches between the groups. If the size of the clusters is constant, an increase in the number of clusters means that there would be an increase in the number of switches. The cluster rate and switch rate in the two groups do not differ from one another.

Table 6

Clusters and Switches Classified by Action Strength. N of Cluster Means the Number of Clusters, and the Same Goes for N of Switch. The Asterisk Indicates a Significant Difference at $\alpha = .05$

	N of Cluster	N of Switch	Cluster size	Cluster rate	Switch rate
elderly	4.40 (<i>SD</i> = 1.67)	10.27 (<i>SD</i> = 3.96)	1.62 (<i>SD</i> = 0.77)	0.24 (<i>SD</i> = 0.05)	0.56 (<i>SD</i> = 0.10)
young	6.03 (<i>SD</i> = 1.78)	12.74 (<i>SD</i> = 3.48)	1.82 (<i>SD</i> = 0.92)	0.26 (<i>SD</i> = 0.07)	0.54 (<i>SD</i> = 0.10)
p-value	< .001*	.012*	.342	.289	.448

Table 7

Perceptual Modalities. N of Cluster Means the Number of Clusters, and the Same Goes for N of Switch. The Asterisk Indicates a Significant Difference at $\alpha = .05$

	N of Cluster	N of Switch	Cluster size	Cluster rate	Switch rate
elderly	3.57 (<i>SD</i> = 1.45)	7.67 (<i>SD</i> = 3.12)	2.87 (<i>SD</i> = 1.58)	0.20 (<i>SD</i> = 0.06)	0.43 (<i>SD</i> = 0.13)
young	4.94 (<i>SD</i> = 1.50)	10.23 (<i>SD</i> = 3.63)	2.88 (<i>SD</i> = 1.76)	0.21 (<i>SD</i> = 0.06)	0.43 (<i>SD</i> = 0.13)
p-value	< .001*	.005*	.974	.448	.971

In Table 7, the results are shown based on perceptual modalities. As most participants, regardless of the age group, predominantly produced verbs that are classified as Visual, the number of clusters by perceptual modalities is fewer, but the average cluster size is larger. *T*-tests by age show that the difference in the number of clusters and the number of switches is significant in this classification.

3.3 A Shorter Time Window

The verb fluency task is administered for 60 seconds. We noticed that most participants produced more verbs in the earlier part of the one minute. The recording of the sessions enabled us to create a list of verbs that are produced in the first 30 seconds and another list of verbs that are produced in the last 30 seconds (see Adams et al., 1989; Bose et al., 2017 for analysis of 15-second intervals; Kim et al., 2011 for 30-second intervals). The average number of verbs produced by participants in each group is shown in Figure 3.

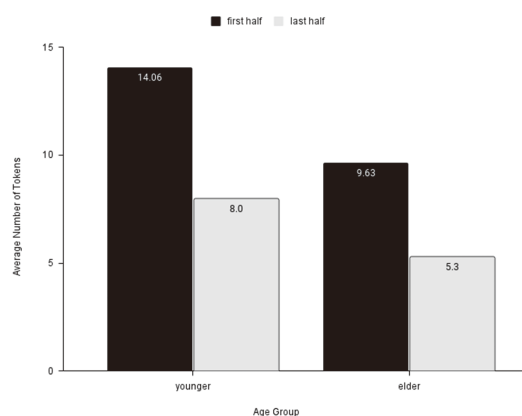
In both groups, more items were produced during the first half than the last half. On average, the younger group produced 14.06 verbs during the

first 30 seconds and 8 verbs during the remaining time. The elderly group showed the same tendency but a fewer number of verbs. The ratio difference between the groups is similar. The participants in the younger group produced 42% fewer verbs in the second half of the one minute. Likewise, the elderly participants recalled 45% fewer verbs in the second half of the minute.

Focusing on the data produced during the first 30

Figure 3

Count of Items Per Person Without Error Tokens in the First Half and the Second Half; Each Half is 30 Seconds Long



seconds, we explored whether the shorter time frame shows sensitivity to the age effect by comparing the two age groups. The younger group produced 14.06 verbs ($SD = 3.45$) whereas the participants in the elderly group produced 9.63 verbs ($SD = 4.03$). The results of a t -test show the participants in these two groups produced significantly different numbers of verbs ($p < .0001$).

The results of the cluster analysis and the switch analysis based on action strength in Table 8 demonstrate that the two age groups show significant differences in the number of clusters produced, the number of switches that are present, and the size of clusters. The elderly group produced fewer numbers of clusters, fewer number of switches, and smaller size of clusters.

In perceptual modularity categories, as shown in Table 9, the number of clusters and switches are significantly different between the elderly group and the younger group, but the cluster size was not different. Since the perceptual modality of most verbs was Visual, this result may not be informative enough.

When results of the elderly group produced in the first 30 seconds are analyzed separately for the early elderly group (under 80) and the late elderly group (over 80), the results showed some sensitivity

to the age group. The number of correct answers was marginally different (4.12 in the early elderly vs. 3.59 in the late elderly, $p = .09$). The verbs categorized by the action strength showed more switches (6.88 by early elderly versus 4.62 in late elderly) and larger switch rate (0.57 by early elderly versus 0.45 by late elderly) ($p < .05$).

4. Discussion

4.1 Number of Verbs Produced by Participants

Reports on the results of the verb fluency task are uncommon, and not many reference points are available except for Lee et al. (2013). Similar to that study, our data reported in section 3.1 shows that the difference in the number of items between the younger group and the older group is significant ($p < .001$). A closer look, however, reveals that the participants in Lee et al. (2013) produced fewer verbs than our participants. The younger group in Lee produced on average 14.5 verbs ($SD = 3.8$), but the younger participants in this study produced on average 22.06 verbs ($SD = 4.59$). The difference in the elderly group was also large. The participants in Lee's study produced about 7.7 verbs ($SD = 4.0$), whereas participants in this study produced on average 14.93 verbs ($SD = 5.83$); almost twice as

Table 8
Action Strength Category in the First Half. The Asterisk Indicates a Significant Difference at $\alpha = .05$

	N of Cluster	N of Switch	Cluster size	Cluster rate	Switch rate
elderly	2.77 ($SD = 1.28$)	5.90 ($SD = 2.90$)	2.57 ($SD = 0.71$)	0.25 ($SD = 0.10$)	0.52 ($SD = 0.13$)
young	3.45 ($SD = 1.36$)	8.00 ($SD = 1.87$)	3.13 ($SD = 1.18$)	0.24 ($SD = 0.09$)	0.55 ($SD = 0.12$)
p-value	.047*	.004*	.028*	.648	.405

Table 9
Perceptual Modalities in the First Half. The Asterisk Indicates a Significant Difference at $\alpha = .05$

	N of Cluster	N of Switch	Cluster size	Cluster rate	Switch rate
elderly	1.70 ($SD = 0.88$)	4.10 ($SD = 2.16$)	5.44 ($SD = 3.76$)	0.16 ($SD = 0.09$)	0.38 ($SD = 0.15$)
young	2.71 ($SD = 1.19$)	6.13 ($SD = 3.24$)	4.67 ($SD = 2.62$)	0.19 ($SD = 0.06$)	0.41 ($SD = 0.15$)
p-value	< .001*	.006*	.374	.168	.497

many verbs were produced.

Individual differences between the two studies would be the main reason for the difference in the number of produced verbs when comparing the two studies. We examined the years of education reported in the two studies. The average year of education is 14.7 ($SD = 1.25$) for the younger group and 13.8 ($SD = 3.16$) for the elderly group. The participants in Lee (2013) had on average 14.2 ($SD = 1.2$) years of education in the younger group and 11.3 ($SD = 1.5$) years of education in the older group. Although the elderly group in Lee et al. (2013) experienced on average 2 years less education than the elderly participants in our study, the younger group in the two studies had comparable years of education, suggesting that the years of education are not the reason for the different number of items in the two studies.

4.2 Rethinking Clusters, Switches, and Cluster Size

During the verb fluency task, participants show performance by retrieving lexical items. In the retrieving process, the optimal method (Troyer, 2000) would be producing words within a subcategory (i.e., clustering). Upon the exhaustion of the subcategory, participants would generate words from a new subcategory; this is called switching. Based on results from phonemic fluency and semantic fluency tasks, Troyer (2000) further states that generating clusters of a subcategory is a relatively automatic process, whereas switching is a relatively effortful process due to the cognitive pressure that involves the retrieval of words from one subcategory to another subcategory. Troyer's (2000) results are based on noun fluency tasks (phonemic, semantic, and animal), and Lee et al. (2013) report results of five different tasks (animal, vegetable, company, celebrity, verb).

This study is based on the verb fluency task where the younger group outnumbers the elderly

group both in the number of clusters and the number of switches. In Troyer (2000, pp. 372-373), age is reported to have a different effect size. As a predictor of switching the effect size was small on phonemic fluency, but large on semantic fluency and animal fluency. Concerning clustering, age had a minimal effect size on all three types of tasks.

Our data does not align with the results reported in Troyer (2000, p. 370), which suggested that clustering is a relatively automatic process, and switching is a relatively effortful process. We observe age effects, as both the number of clusters and the number of switches decrease in the elderly group, suggesting that dividing the process of clustering and switching based on data produced within a short period of time (one minute) is not straightforward.

An interesting observation is the difference in the number of correctly produced verbs between the participants in the early elderly group and in the late elderly group. Whether this difference indicates a gradual age effect in retrieving verbs within a short period of time is beyond the scope of this paper⁸.

4.3 The Age Effect

The elderly group produced fewer verbs as well as fewer clusters and fewer switches compared to the younger group. This tendency suggests that the verb fluency task is sensitive to the age difference. In the normalized values, such as the cluster size, the cluster rate and the switch rate, the two groups demonstrate comparable patterns, indicating that the ratio is not what triggers the difference between the groups. The age effect is also observed when only the first 30 seconds are considered. In addition to fewer verbs, clusters and switches, the cluster size is also smaller when the verbs are classified by its action strength (but not by its perceptual modalities).

5. Conclusion

This study reports the results of a verbal fluency task on the non-clinical Japanese speakers in two age groups. The main goal of this study has been to provide baseline results that control only for age, unlike previous studies that compare healthy younger participants with clinically old participants. The age effect was observed when the two groups were compared in terms of the number of verbs, the number of switches, and the number of clusters, while the average cluster size was mainly constant. Results from the first 30 seconds further suggest that identifying the difference does not hinge upon the entire 60 seconds of the time, usually used in the verb fluency task.

It might be the case that the verb fluency task coupled with other fluency tasks could be used to screen participants in terms of language usage, especially when researchers do not have direct access to other standardized screening tests. We await future studies that compare results of the verb fluency task with other cognitive screening tasks.

Acknowledgment

This research is supported by the JSPS Bilateral Grant with NRF of Korea in 2023 (#JPJSBP120238812). We thank for comments by Prof. Jee Eun Sung and the members of the Neurogenic Communication & Brain Lab at the Ewha Womans University.

Notes

1 A reviewer pointed out that there are some studies that compared healthy participants of different age groups which have been omitted (e.g., Zeng et al., 2019). Zeng et al. (2019) examined the effects of bilingual experience on the processing abilities of children, young and elderly adults. There may be an interesting connection between the scope of this study and Zeng et al (2019), but we thought that the focus of the two studies is not directly comparable.

- 2 A reviewer asked about the exact mechanism of how aging affects working memory or lexical retrieval ability. Although the mechanism is interesting to investigate, this study aims to build empirical data on the aging effect, and our experiment was not designed to explore the mechanism. As such, how working memory works differently in different ages is not discussed in this paper.
- 3 According to Troyer (2000), this observation is not robust. Whether clustering is easier than switching appears to vary depending on the type of clinical issues participants have. Despite this inconsistency, measuring clustering and switching still provides valuable insights into the real-time cognitive processes during fluency tests.
- 4 A reviewer and editor pointed out that the authors could go more in-depth about the choice of sensorimotor norms for this study, the significance of verb types in the verbal fluency task (e.g. perception modalities vs action effectors) and any relevant issues such as categorization of verb types, and the automatic nature of clusters vs. non-automatic nature of switching. It is indeed interesting to look at the different sensorimotor characteristics of the verbs produced. However, our study aims to add to the empirical data on the verbal fluency of verbs, which has yet to be fully reported and analyzed. Sensorimotor norms are widely used for analyzing fluency tests on other languages (e.g. Lee, J., & Shin, J.-A., 2023), and we followed their methods in this study so that our study can be comparable to such studies. Since the focus of this paper is reporting Japanese data on verb fluency, we did not discuss much the choice of sensorimotor norms for this study, the significance of verb types in the verbal fluency task (e.g. perception modalities vs action effectors), and any relevant issues such as categorization of verb types.
- 5 One participant's response contains an error on the birth date, so the age and education year information of that participant is not included here.
- 6 https://embodiedcognitionlab.shinyapps.io/sensorimotor_norms/ (accessed on August 29, 2024).
- 7 A reviewer suggested the idea that the difference in the verb produced by the two age groups is related to how physically active they are. For example, older participants produced more head verbs than hand verbs, unlike the younger participants, showing the opposite pattern. The

reviewer wondered if this might be because older participants tend to be less physically active but perhaps have more experience with head-related verbs. We agree that one can speculate in this way as to why we observed such opposite tendencies between the two groups, but the experiment in this study did not specifically address this issue. We took this idea for further research direction.

- 8 A reviewer wondered whether we could find patterns in the order of the switches. We took samples from 10% of the participants (3 younger and 3 elder). Our impression is that there are no discernible patterns in the order of switches in both perception and action. Also, we looked at the interaction between the size of clusters and the types of clusters, and there is an interspeaker variation as to what kind of actions and perception verbs are preferred in clusters.

References

- Adams, M. L., Reich, A. R., & Flowers, C. R. (1989). Verbal Fluency Characteristics of Normal and Aphasic Speakers. *Journal of Speech, Language, and Hearing Research*, 32(4), 871-879. <https://doi.org/10.1044/jshr.3204.871>
- Alegret, M., Peretó, M., Pérez, A., Valero, S., Espinosa, A., Ortega, G., Hernández, I., Mauleón, A., Rosende-Roca, M., Vargas, L., Rodríguez-Gómez, O., Abdelnour, C., Berthier, M. L., Bak, T. H., Ruiz, A., Tárraga, L., & Boada, M. (2018). The Role of Verb Fluency in the Detection of Early Cognitive Impairment in Alzheimer's Disease. *Journal of Alzheimer's Disease*, 62(2), 611-619. <https://doi.org/10.3233/JAD-170826>
- Bose, A., Wood, R., & Kiran, S. (2017). Semantic fluency in aphasia: Clustering and switching in the course of 1 minute: Semantic fluency in aphasia. *International Journal of Language & Communication Disorders*, 52(3), 334-345. <https://doi.org/10.1111/1460-6984.12276>
- Burke, D. M., MacKay, D. G., Worthley, J. S., & Wade, E. (1991). On the tip of the tongue: What causes word finding failures in young and older adults?. *Journal of memory and language*, 30(5), 542-579.
- Iida, H., & Akita, K. (2023). Perceptual Strength Norms for 510 Japanese Words, Including Ideophones: A Comparative Study with English. *Proceedings of the Annual Meeting of the Cognitive Science Society*, 45, 2201-2207.
- Ito, E., & Hatta, T. (2006). Reliability and validity of verbal fluency tasks. *The Japanese Journal of Neuropsychology*, 22(2), 146-152.
- Iwasaki, N., Vinson, D. P., Vigliocco, G., Watanabe, M., & Arciuli, J. (2008). Naming action in Japanese: Effects of semantic similarity and grammatical class. *Language and Cognitive Processes*, 23(6), 889-930. <https://doi.org/10.1080/01690960801916196>
- James, L. E., & Burke, D. M. (2000). Phonological priming effects on word retrieval and tip-of-the-tongue experiences in young and older adults. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 26(6), 1378.
- Kim, H., Kim, J., Kim, D. Y., & Heo, J. (2011). Differentiating between Aphasic and Nonaphasic Stroke Patients Using Semantic Verbal Fluency Measures with Administration Time of 30 Seconds. *European Neurology*, 65(2), 113-117. <https://doi.org/10.1159/000324036>
- Kemper, Susan (2012): The interaction of linguistic constraints, working memory, and aging on language production and comprehension. In Naveh-Benjamin Moshe & NobuoOhta (Hrsg.), *Memory and Aging*, 31-47. New York, NY: Psychology Press
- Lee, D., Sawada, Y., Nakamura, H., Tokuchi, R., & Fujimoto, N. (2013). Effects of Word Class and Aging on Verbal Fluency. *Higher Brain Function Research*, 33(4), 421-427. <https://doi.org/10.2496/hbfr.33.421>
- Lee, J., & Shin, J.-A. (2023). The cross-linguistic comparison of perceptual strength norms for Korean, English and L2 English. *Frontiers in Psychology*, 14, 1188909. <https://doi.org/10.3389/fpsyg.2023.1188909>
- Lynott, D., & Connell, L (2009). Modality exclusivity norms for 423 object properties. *Behavior Research Methods*, 41, 558-564. <https://doi.org/10.3758/BRM.41.2.558>
- Lynott, D., Connell, L., Brysbaert, M., Brand, J., & Carney, J. (2020). The Lancaster Sensorimotor Norms: Multidimensional measures of perceptual and action strength for 40,000 English words. *Behavior Research Methods*, 52(3), 1271-1291. <https://doi.org/10.3758/s13428-019-01316-z>
- Piatt, A. L., Fields, J. A., Paolo, A. M., & Tröster, A. I. (1999). Action (verb naming) fluency as an executive function measure: Convergent and divergent evidence of validity. *Neuropsychologia*, 37(13), 1499-1503. [https://doi.org/10.1016/S0028-3932\(99\)00066-4](https://doi.org/10.1016/S0028-3932(99)00066-4)
- Sakamoto, Y. (2016). Phases and argument ellipsis in Japanese. *Journal of East Asian Linguistics*, 25(3), 243-274. <https://doi.org/10.1007/s10831-016-9145-6>
- Sauzéon, H., Lestage, P., Raboutet, C., N'Kaoua, B., & Claverie, B. (2004). Verbal fluency output in

- children aged 7-16 as a function of the production criterion: Qualitative analysis of clustering, switching processes, and semantic network exploitation. *Brain and Language*, 89(1), 192-202. [https://doi.org/10.1016/S0093-934X\(03\)00367-5](https://doi.org/10.1016/S0093-934X(03)00367-5)
- Shafto, M. A., Burke, D. M., Stamatakis, E. A., Tam, P. P., & Tyler, L. K. (2007). On the tip-of-the-tongue: neural correlates of increased word-finding failures in normal aging. *Journal of cognitive neuroscience*, 19(12), 2060-2070.
- Takahashi, D. (2008). *Noun Phrase Ellipsis*. Oxford University Press. <https://doi.org/10.1093/oxfordhb/9780195307344.013.0015>
- Troyer, A. K. (2000). Normative Data for Clustering and Switching on Verbal Fluency Tasks. *Journal of Clinical and Experimental Neuropsychology*, 22(3), 370-378. [https://doi.org/10.1076/1380-3395\(200006\)22:3;1-V;FT370](https://doi.org/10.1076/1380-3395(200006)22:3;1-V;FT370)
- Woods, S. P., Scott, J. C., Sires, D. A., Grant, I., Heaton, R. K., Tröster, A. I., & Group, T. H. N. R. C. (Hnrc). (2005). Action (verb) fluency: Test-retest reliability, normative standards, and construct validity. *Journal of the International Neuropsychological Society*, 11(4), 408-415. <https://doi.org/10.1017/S1355617705050460>
- Yoo, Y. R., Paek, E. J., & Sung, J. E. (2023). Aging-related Differences in Emotional Valence from the Verb Fluency Task Depending on the Time Period. *Communication Sciences & Disorders*, 28(4), 791-804. <https://doi.org/10.12963/csd.231002>