# Effects of Experience of Listening to Short Sentences Containing ANEWs on Memory: An Analysis Based on Pupillary Responses during Listening and Visual Behavior during Impression Evaluation

Katsuko T. Nakahira
*Nagaoka University of Technology*
Nagaoka, Niigata, Japan
Email: katsuko@vos.nagaokaut.ac.jp

Sho Hasegawa
*Nagaoka University of Technology*
Nagaoka, Niigata, Japan
Email: s211070@stn.nagaokaut.ac.jp

Shunsuke Moriya
*Nagaoka University of Technology*
Nagaoka, Niigata, Japan
Email: s213345@stn.nagaokaut.ac.jp

Muneo Kitajima
*Nagaoka University of Technology*
Nagaoka, Niigata, Japan
Email: mkitajima@kjs.nagaokaut.ac.jp

***Abstract*— The present study focuses on the memory of verbal information by short sentences provided auditorily and examines the relationship between the attributes of emotion-inducing words, i.e., Affective Norms for English Words (ANEW), in the verbal information and the memory of the information contained in the short sentences. We have developed a cognitive model of auditory information comprehension, in which auditory information is sequentially input from the outside, stored in working memory, cognitively understood, and the memory network is updated to reflect the results of these processes. According to this model, the content retained in working memory during listening should influence subsequent recall, as externally provided auditory information is transient. Furthermore, previous studies have shown that the presence of ANEWs also affects memory performance. In the current study, 32 short sentences (each read aloud within 15-20 seconds) were created by embedding two ANEWs—one evoking a positive emotion and the other a negative emotion—within four different patterns. The interval between the presentation of the first and second ANEW varied: either 1 second or 7 seconds. Seventeen subjects listened to these short sentences, recalled their content, and we recorded the number of items recalled. The eye-tracker was used to measure pupillary responses during listening and visual behavior during the selection of evaluation items related to the impression of the listening experience. We analyzed the relationship between total pupil diameter change, number of fixations, impression evaluation time required, impression evaluation results, interval between ANEWs, positive/negative ANEW combination patterns, and recall performance. The following suggestions for the mechanism leading to high recall performance were obtained: 1) when the way ANEW pairs appear in a short sentence increases the load on their cognitive processing; 2) when the cognitive processing load is not high but the impression of short sentences are strong. Finally, we discussed ways to apply the findings from this study to the design of auditory information that facilitates memory.**

***Keywords*— *Affective Norms for English Words; Cognitive Process; Two Minds; Pupillary Response; Eye Movement; Recall;***

## I. Introduction

In our daily lives, we receive physical and chemical stimuli from the outside world through our sensory organs and process them cognitively by combining them with various information stored in the brain. To understand these processes, information obtained from perceptual processes is very useful. In particular, suggestions on the cognitive processes provided by visual information, i.e., eye movement and pupil measurements, are important and have been reviewed by many researchers.

For example, Mahanama et al. [1] summarize the methods for measuring eye gaze and pupil response. Vilotijević [2], for example, suggests that the intensity of pupillary response is influenced not only by the physical characteristics of light but also by cognitive factors such as gaze.

Studies on cognitive load, eye movement, and pupillary response, for example, include the following. Skaramagkas et al. [3] focus on visual attention, emotional arousal and cognitive workload, presenting a review of metrics related to gaze and pupil tracking (gaze, fixations, saccades, blinks, pupil size variation, etc.) that are used to detect emotional and cognitive processes. Other studies include one that examined whether eye movements respond to the viewer's cognitive load, and another that examined pupil dilation as a measure of lexical retrieval.

Research on pupillary response and emotion (valence-arousal pair) is also ongoing. In the past, pupil size variation during and after auditory emotional stimulation was examined by [4]. They suggested that pupil dilation is greater for emotionally negative/positive stimuli than for neutral stimuli. Henderson et al. [5] investigated pupil response during emotional imagery, and found that pupil dilation during emotional (pleasant or unpleasant) imagery was significantly greater than pupil dilation during neutral imagery. They suggested that pupil diameter was increased during emotional (pleasant or unpleasant) imagery compared to neutral imagery. Oliva et al. [6] investigated the relationship between pupil response and the process of emotion recognition, and showed that during emotion recognition, the time course of pupil response is driven by the decision-making process. Tarnowski et al. [7] investigated emotion recognition using eye tracking. They calculated 18 features related to eye movements (fixations and saccades) and pupil diameter, and found up to 80% classification accuracy for high arousal and low valence, low arousal and middle valence, high arousal and high valence, and low arousal and low valence, and high arousal and high valence.

Thus, eye movement and pupil response have the potential to detect the occurrence of emotion in response to emotional stimuli input to humans. Based on the above, we have been continuously conducting basic research on designing multi-modal content that can be easily recalled by real-time detection of such emotional responses [8][9][10][11][12]. So far, we have suggested the following for content with visual-auditory multimodal information: 1) It is possible to optimize the load of perception information processing by adjusting the interval between perception information. 2) The degree of emotion generated by Affective Norms for English Words (ANEW) may suppress the characteristics of the total change in pupillary response. 3) We constructed a cognitive model of the process of sequential input of externally provided auditory information, its retention in working memory, its recognition, and its updating in the memory network.

However, previous experiments have not yet investigated the relationship between the impression evaluation of the presented stimulus, the total change in pupillary response, and the recall rate after listening to the presented stimulus. For this reason, this paper investigates the relationship among eye movement data, total change in pupillary response, and recall rate at the time of selecting rating items regarding the impression of listening experience, based on pupillary response data during short sentence listening and experiences with multiple positive and negative two-level valences. We investigate the relationship between the total change in pupillary response and recall rate.

In Section II, we describe the impression evaluation process and possible measurements based on the cognitive model presented in [12]. In Section III, we describe the experiment and its analysis results. In Section IV, we discuss the relationship between pupillary response to ANEW and recall.

## II. Valence Evaluation Process and Possible Biometric Mesurments

In this section, based on the cognitive model of listening to auditory information proposed by Nakahira et al. [12], possible biometric measurements are selected for impression evaluation after listening to auditory information.

Figure 1 shows the main flow. The response model for human emotion shown in Figure 1 is the Construction-Integration Model (CI-model) proposed by Kintsch [13][14][15]. The CI-model is a theory of discourse comprehension consisting of a construction step and an integration step, in which symbolic features are integrated using the connectionist technique. The cognitive model was constructed as follows, incorporating this theory. The presented auditory information is matched with long-term memory for each specific packet of sound waves, which are basically groups of phonemes that correspond to specific concepts. The matching results are returned together with the valence and other information associated with the concept. The returned matching results stay in working memory as a group for a certain period of time. The more the concepts overlap in working memory, the stronger the connection between the concepts becomes. A series of concepts are returned to long-term memory as related, but in this case, various node relationships, such as between concepts and between concepts and valences, are strengthened.

The possibility of measuring human data by applying Figure 1 is examined based on the overall picture of the pupil response shown in Nakahira et al. [12]. Since the presented stimuli assumed in this paper are sentences provided by auditory stimuli, the nature and arrangement of ANEWs in the narration are considered to influence the listener. Therefore, the control parameters are valence and arousal, which indicate the nature of ANEWs, and the interval when there are multiple ANEWs. In this paper, we deal specifically with valence among the properties of ANEWs. In this paper, two ANEWs are prepared. If positive valence is denoted as $V_{++}$ and negative valence as $V_{--}$, then a pair of ANEWs can be denoted as $(V_{++},\ V_{--})$. The interval between ANEWs is denoted by $dT$.

The effect of these ANEWs appears in the pupil diameter of the listener, and the timing of the effect is considered to occur during and after listening to the text, respectively. In this study, we measure the total change in pupillary response, $r_{all}$, as proposed in Nakahira et al. [12]. The $r_{all}$ measured each time an ANEW appears is considered a response to local emotion. In this study, since we deal with two ANEWs, we denote them as $r^1_{all}$ and $r^2_{all}$, respectively. In addition, at the end of the sentence, the impression, or emotion, of the whole sentence heard should be generated, and the resulting total change in pupillary response is expressed as a $r^e_{all}$.

Next, the following parameters are considered measurable during impression evaluation.

- **Impression evaluation speed**:
  The speed of impression evaluation is affected by the clarity of the impression of the listening sentence. The time taken by subject $p$ to evaluate the impression of the presented stimulus $i$ is measured and is denoted as $t^i_p$.
- **Visual behavior during impression evaluation**:
  The choices when the subject responds to the impression evaluations are given as visual information. We consider the viewpoint movement during the viewing of options as an element related to the decision-making process of selecting options. The number of stopping points $n^i_{p,f}$ of $p$ for $i$ can be regarded as the number of times the subject searches for alternatives. The $j$th stop time $t^{i,j}_{f,p}$ during choice browsing indicates the approximate degree to which the user has thought about the choices.

In addition to the above measurements, the following analysis is possible.

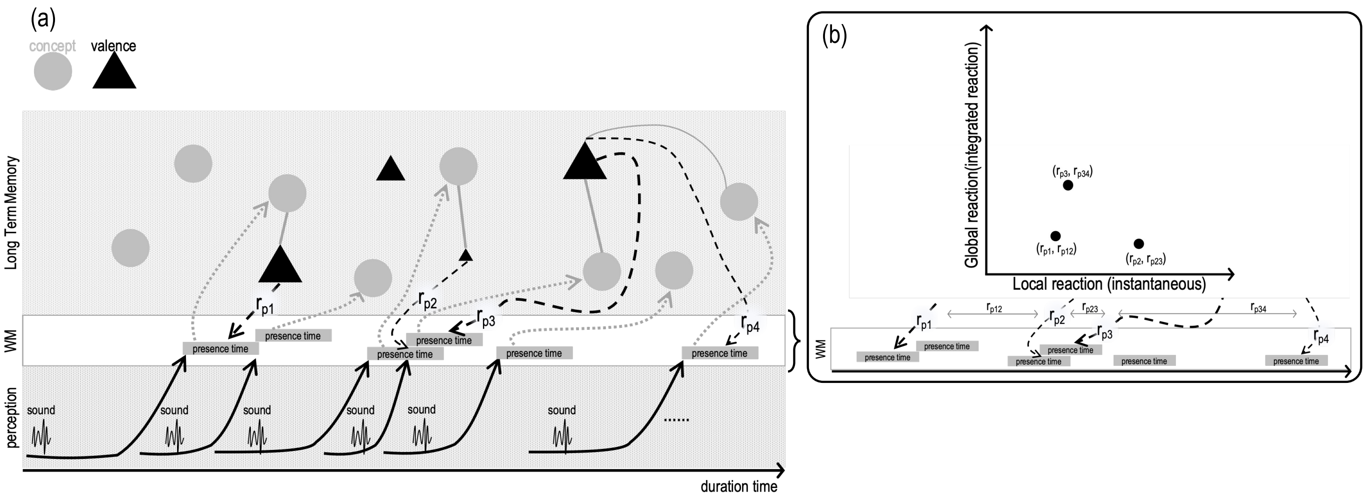


Figure 1. Cognitive model of this paper based on CI model. (a) Input - Cognitive process. (b) Working memory processing - output process. Reposting of the figure in [12].

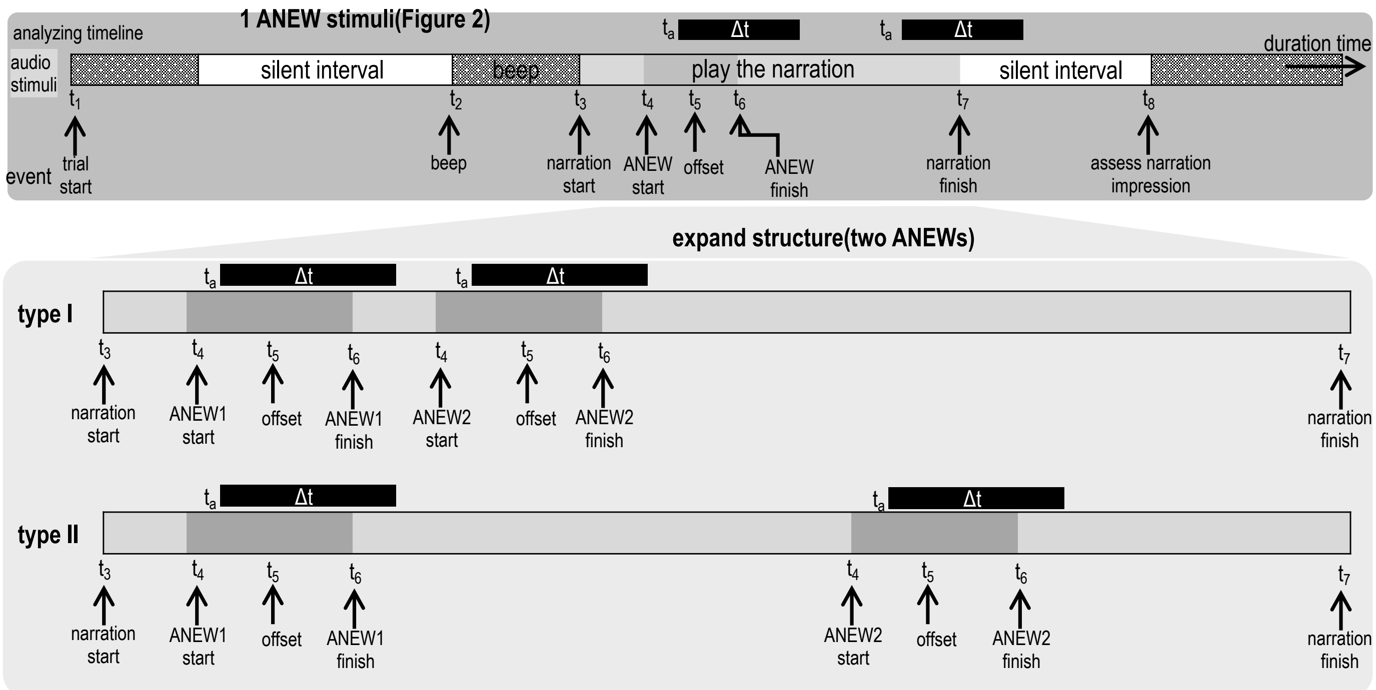


Figure 2. Timeline of presented stimuli including multiple ANEWs. Reposting of the figure in [12].

### A. *Total Fixation Time during Valence Evaluation*

In the impression evaluation, it is difficult to assess the degree of concentration on the choices, i.e., the degree of thinking, from $t_p^i$ alone. Therefore, to confirm the degree to which the subject concentrated on checking the options during $t_p^i$, the total pause time during the impression evaluation, $t_{f,p}^i$, is determined. This is given by the following equation:

$$t_{f,p}^i = \sum_j^{n_f^i} t_{f,p}^{i,j}$$

### B. $T_{c,p}^i : t_p^i - t_{f,p}^i$ *ratio*

$t_p^i$ and $t_{f,p}^i$ are measured in actual time required. However, when comparing only the actual time, it is difficult to distinguish between cases where the impression evaluation takes a relatively long time and those where it takes time only under certain conditions. To solve this problem, the ratio of $t_p^i$ to $t_{f,p}^i$, $T_{c,p}^i$, is calculated as one of the indices and used as one of the characteristic quantities of impression evaluation. $T_{c,p}^i$ is calculated using the following equation:

TABLE I. PARAMETERS PREPARED IN THIS PAPER.

| valence pair | $dT$ | |
|---|---|---|
| | 1 sec | 7 sec |
| $(V_{++},\ V_{++})$ | 4 | 4 |
| $(V_{--},\ V_{--})$ | 4 | 4 |
| $(V_{++},\ V_{--})$ | 4 | 4 |
| $(V_{--},\ V_{++})$ | 4 | 4 |

$$T^i_{c,p} = \frac{t^i_{f,p}}{t^i_p}$$

## III. EXPERIMENT AND ANALYSIS

### A. Preliminary Experiment

In order to evaluate the validity of the content reviewed in Section II, the same experiment was conducted on 17 participants using the experimental technique in multiple ANEWs indicated in Nakahira et al. [12]. The tasks of the participants were as follows. The participants were presented with 32 sentences with various valence pairs and $dT$s, eight sentences at a time, at random. The participants listen to the presented narration, and after each sentence, they evaluate their impression of the narration in seven levels from positive to negative and in eight levels from "no impression" to "impression unknown". The results of the impression evaluation are recorded by selecting an option displayed on the terminal and inputting the result with the mouse operation. After that, a recall test is given for each of the eight sentences, and the participant verbally states the contents of the recall. The experiment was conducted in a quiet room under constant illumination to prevent disturbance by other stimuli.

The parameters used in the experiment were as follows: first, the combinations of the ANEW and $dT$ values are shown in the table, and the number of presented stimuli is 32 sentences. The values of valence for $V_{++}$ ranged from 7.70 to 9.00. The values of valence for $V_{--}$ were selected from 1.00 to 1.99. The values of valence for $V_{++}$ were from 7.70 to 9.00, and the values of valence for V– were from 1.00 to 1.99. The number of syllables per presented stimulus was around 65. To examine the relationship between memory and the number of syllables, a recall test was conducted on the content of each sentence after every eight sentences were listened to, and the degree of recall was checked.

The following is an analysis of the parameters set in Section 2 based on the data obtained in the experiment.

### B. Basic Statistics for Metrics

First, the overall picture of the data obtained in this study is shown in table II. In the table, $\sigma$ indicates the standard deviation, $Q_{1/4}$ indicates the first quartile, $Q_{2/4}$ indicates the median, and $Q_{3/4}$ indicates the third quartile. Table II shows the statistics of the available data without any particular classification. The following can be read from the table.

The difference between $t^i_p$ and $t^i_{f,p}$ in the time taken to evaluate impressions is close to one second. From this fact, it is expected that some trends in the time-related quantities can be seen by classifying the response results. For $r_{all}$, it is difficult to find a large difference between $r^1_{all}$ and $r^2_{all}$. On the other hand, a certain difference is observed between $r^e_{all}$ and $r^{1/2}_{all}$. This is expected to provide some insight into the amount of $r^{1/2}_{all}$ starting from $r^e_{all}$. We made relation analysis based on metrics $t^i_p,\ t^i_{f,p},\ T^i_{c,p},\ n^i_{p,f}$ and belows:

TABLE II. BASIC STATISTICS FOR THE ENTIRE DATA OBTAINED IN THE PRELIMINARY EXPERIMENT.

| parameter | average | $\sigma$ | $Q_{1/4}$ | $Q_{2/4}$ | $Q_{3/4}$ |
|---|---|---|---|---|---|
| $t^i_p$[ms] | 3958.69 | 2188.11 | 2529.50 | 3411.50 | 4666.50 |
| $t^i_{f,p}$[ms] | 2933.94 | 2590.75 | 1138.39 | 2220.33 | 4140.00 |
| $T^i_{c,p}$ | 0.67 | 0.32 | 0.4 | 0.7 | 1 |
| $n^i_{p,f}$ | 7.54 | 5.61 | 4 | 7 | 10 |
| $r^1_{all}$[mm] | 2.68 | 2.53 | 0.83 | 1.92 | 3.66 |
| $r^2_{all}$[mm] | 2.75 | 2.47 | 0.89 | 2.02 | 3.82 |
| $r^e_{all}$[mm] | 2.43 | 2.48 | 0.73 | 1.53 | 3.33 |

- valence pair and $dT$
- $r_{all}$
- results of recall test
- results of $S_p$

### C. Relation between $t^i_p,\ t^i_{f,p},\ T^i_{c,p}, n^i_{p,f}$ and valence pair or $dT$

Figure 3 (a) $\sim$ (d) shows boxplots for $T^i_{c,p}$ and $n^i_{p,f}$, the valence pair, and $dT$. (a) in the figure shows boxplots for the valence pair and $T^i_{c,p}$, (b) in the figure shows boxplots for the valence pair and $n^i_{p,f}$, (c) in the figure shows boxplots for $dT$ and $T^i_{c,p}$, and (d) in the figure shows boxplots for $dT$ and $n^i_{p,f}$. Here, the analysis for the value pair is considered as follows: Normally, the valence pair should be considered only in terms of the value of the valence for ANEW. However, in the experiment in this paper, there is a possibility that the valence of one pair strengthens or weakens the valence of the other pair as a result of their interaction. For this reason, we classified the valence pairs by including $dT$.

Based on the figure, we performed a one-way ANOVA for the valence pairs $t^i_p,\ t^i_{f,p}, T^i_{c,p}$, and $n^i_{p,f}$. The results showed a significant difference in $T^i_{c,p}$ $(F(7, 536) = 3.018, p = 0.004)$. Multiple Comparison Procedure using Tukey's method revealed significant differences in the following combinations:

- $(V_{++},\ V_{++},\ dT = 1) - (V_{--},\ V_{--},\ dT = 1)$,
- $(V_{++},\ V_{++},\ dT = 1) - (V_{--},\ V_{--},\ dT = 7)$,
- $(V_{++},\ V_{++},\ dT = 1) - (V_{--},\ V_{++},\ dT = 7)$

### D. Relation between $t^i_p,\ t^i_{f,p},\ T^i_{c,p}, n^i_{p,f}$ and $r_{all}$

Next, a one-way ANOVA was performed for each of $r^{1/2/e}_{all}$ with the valence pair and $dT$. However, no significant results were obtained in either case. Therefore, a test of no correlation was performed to confirm the existence of correlation between $r^{1/2/e}_{all}$ and $t^i_p, t^i_{f,p}, T^i_{c,p}$, and $n^i_{p,f}$. As a result, the following correlations were confirmed:

- $(V_{++},\ V_{++},\ dT = 1)$
  correlation coefficient $\rho = -0.24$ $(p = 0.05)$ between $r^1_{all} - r^e_{all}$ and $T^i_{c,p}$

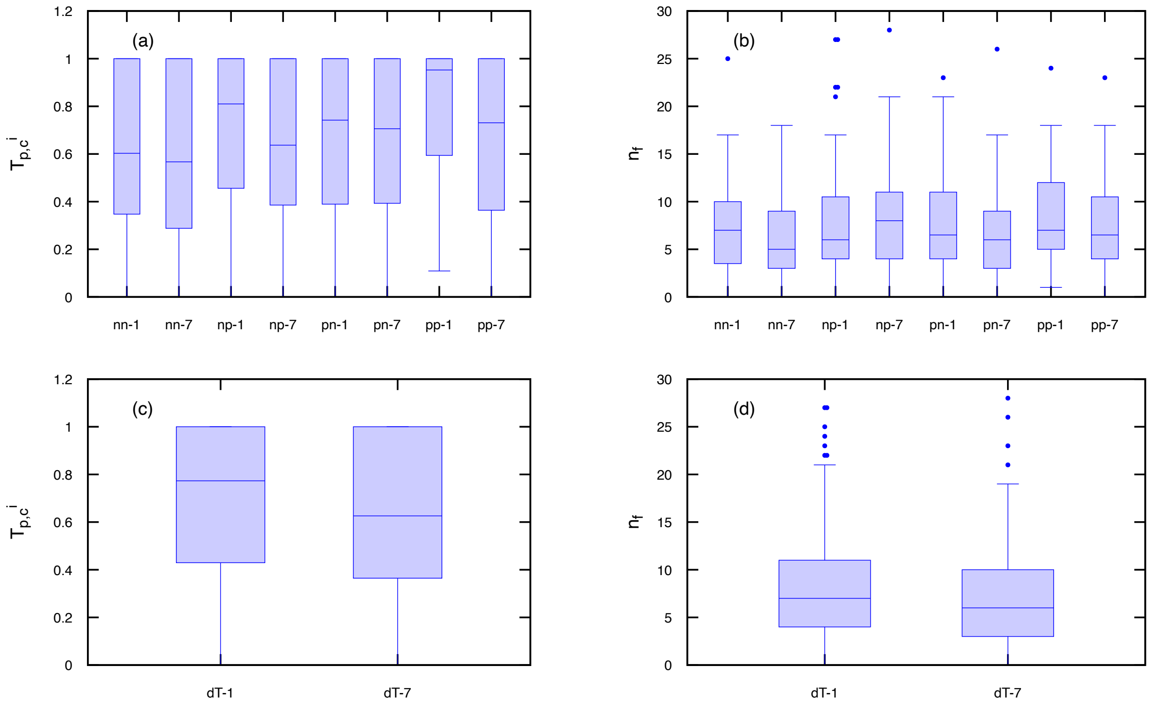


Figure 3. Boxplots for $T^i_{c,p}$ and $n_f$, with $T^i_{c,p}$ on the left and nf on the right. Positive valence is represented by **p**, negative valence by **n**, and the interval between the combination and ANEW is represented by a number.

TABLE III. NUMBER OF CORRECT ANSWERS FOR RECALL WITH ANEW1, ANEW2, AND $dT$ COMBINATIONS.

| | | | |
|---|---|---|---|
| $dT = 1$ | | | |
| | | valence type of ANEW1 | |
| | | $V_{++}$ | $V_{--}$ |
| valence type of ANEW2 | $V_{++}$ | **34** | 23 |
| | $V_{--}$ | 27 | 21 |
| $dT = 7$ | | | |
| valence type of ANEW2 | $V_{++}$ | *17* | 25 |
| | $V_{--}$ | 21 | **28** |

- $(V_{++},\ V_{++},\ dT = 7)$
  correlation coefficient $\rho = -0.237$ $(p = 0.05)$ between $r^1_{all} - r^2_{all}$ and $T^i_{c,p}$

However, all of them are weak correlation coefficients.

### E. Relation between $t^i_p$, $t^i_{f,p}$, $T^i_{c,p}, n^i_{p,f}$ and result of the recall test

All the results have been analyzed without considering the participants' reactions. Next, we conduct an analysis that takes the participant's responses into account. First, we analyze the relationship between the results of the recall test and the various quantities. The results of the recall test were recorded by removing the recency effect and the primacy effect. In this paper, a recall test was conducted for each of the eight presented stimuli. Therefore, the results of the first and eighth sentences were excluded from the number of recalled sentences, and the sentences that were able to reproduce the meaning of the sentence were judged to be "recalled". The results are shown in Table III.

A one-way ANOVA was performed on the measured quantities according to this classification, but no significant differences were found. The results of the test of no correlation on the values of the quantities with respect to the value pairs showed the correlations shown in Table IV.

### F. Relation between $t^i_p$, $t^i_{f,p}$, $T^i_{c,p}$, $n^i_{p,f}$, $r^{1/2/e}_{all}$ and Result of Impression Evaluation

Finally, we analyze the relationship between the classification of the subjects' impression ratings and the various measurements. In Nakahira et. al. [12], the results of the participants' impression ratings of $S_p$ were classified according to the distance from the ANEW valence contained in the presented stimulus as follows:

$$
\begin{array}{rl}
\text{VES}: & |V_1 - S_p| < 1,\ |V_2 - S_p| < 1 \\
\text{V1ES}: & |V_1 - S_p| < 1,\ |V_2 - S_p| \geq 2 \\
\text{V1LS}: & 1 \leq |V_1 - S_p| < 2,\ |V_2 - S_p| \geq 2 \\
\text{V2ES}: & |V_1 - S_p| \geq 2,\ |V_2 - S_p| < 1 \\
\text{V2LS}: & 1 \leq |V_2 - S_p| < 2,\ |V_1 - S_p| \geq 2 \\
\text{N/A}: & |V_1 - S_p| \geq 2,\ |V_2 - S_p| \geq 2
\end{array}
$$

This paper also classifies $S_p$ accordingly. An ANOVA was conducted on the various measures with $S_p$ as the explanatory variable, but no significant differences were obtained. Strictly speaking, a one-way ANOVA showed significant differences when $S_p$ was classified as V1ES/V1LS, but the number of samples in these classifications was less than 20, which was not precise enough to perform ANOVA. For this reason, this

TABLE IV. RESULTS OF THE TEST OF NO CORRELATION BETWEEN MEASUREMENTS FOR EACH CORRECT/INCORRECT ANSWER IN THE RECALL.

| memory availability | valence pair | measurements | | | $\rho$ | $p-value$ |
|---|---|---|---|---|---|---|
| entire set | $(V_{++},\ V_{--}, dT=1)$ | $r^2_{all}$ | and | $n_f$ | 0.313 | 0.009 |
| | $(V_{++},\ V_{++},\ dT=1)$ | $r^1_{all}-r^e_{all}$ | and | $T^i_{c,p}$ | -0.240 | 0.049 |
| | $(V_{++},\ V_{++},\ dT=7)$ | $r^1_{all}$ | and | $n_f$ | 0.303 | 0.012 |
| | | $r^2_{all}$ | and | | 0.258 | 0.034 |
| recalled | $(V_{++},\ V_{--}, dT=1)$ | $r^2_{all}$ | and | $t^i_{f,p}$ | 0.477 | 0.026 |
| | | | and | $n_f$ | 0.534 | 0.011 |
| | $(V_{--},\ V_{++},\ dT=7)$ | $r^1_{all}-r^e_{all}$ | and | $t^i_{f,p}$ | 0.564 | 0.008 |
| | | | and | $T^i_{c,p}$ | 0.543 | 0.011 |
| not recalled | $(V_{--},\ V_{++},\ dT=1)$ | $r^2_{all}$ | and | $n_f$ | 0.302 | 0.040 |
| | $(V_{++},\ V_{++},\ dT=1)$ | $r^1_{all}-r^e_{all}$ | and | $T^i_{c,p}$ | -0.351 | 0.024 |
| | $(V_{++},\ V_{++},\ dT=7)$ | $r^1_{all}$ | and | $n_f$ | 0.323 | 0.017 |
| | | $r^2_{all}$ | and | | 0.347 | 0.288 |
| | | $r^1_{all}-r^e_{all}$ | and | | 0.275 | 0.044 |
| | | $r^2_{all}-r^e_{all}$ | and | | 0.310 | 0.023 |

TABLE V. RESULTS OF TEST OF NO CORRELATION BETWEEN MEASUREMENTS FOR EACH TYPE OF $S_p$.

| assess type of $S_p$ | measurments | | | $\rho$ | $p-value$ |
|---|---|---|---|---|---|
| VES | $r^1_{all}$ | and | $n_f$ | 0.21 | 0.011 |
| V2ES | $r^1_{all}-r^e_{all}$ | and | $t^i_p$ | 0.258 | 0.005 |
| | | and | $t^i_{f,p}$ | 0.298 | 0.001 |
| | | and | $n_f$ | 0.228 | 0.014 |
| | $r^1_{all}$ | and | $t^i_{f,p}$ | 0.236 | 0.010 |
| | | and | $n_f$ | 0.222 | 0.016 |
| V2LS | $r^1_{all}-r^2_{all}$ | and | $n_f$ | -0.219 | 0.038 |
| | $r^1_{all}-r^e_{all}$ | and | $t^i_p$ | -0.298 | 0.004 |
| | | and | $t^i_{f,p}$ | -0.321 | 0.002 |
| | | and | $n_f$ | -0.208 | 0.050 |

paper focuses only on the results of the ANOVA for VES, V2ES, and V2LS.

The results of the test of no correlation on the $S_p$ based on the various quantities, show the correlations presented in the table. As a reference, for V1ES and V1LS, a correlation greater than $\rho = 0.5$ was observed.

## IV. DISCUSSION:

Based on the $r^{1/2/e}_{all}$ and impression ratings obtained from the experimental results when presented stimuli with $t^i_p,\ t^i_{f,p}, T^i_{c,p},\ n^i_{p,f}$ and ANEW are given, we discuss the characteristics of the presented stimuli that are likely to be recalled.

### A. Consideration of Observed Metrics

Table VI shows the medians of the presenting stimulus conditions in this paper, $dT$, the valence pair, and the presence or absence of recall and the type of $S_p$ obtained from the experiment. The table does not include value pairs for which the results of the test of no correlation were not significant except for $(V_{--},\ V_{--}, dT=7)$. First, focusing on $T^i_{c,p}$, the values are lower for $dT=7,$ VES, $(V_{--},\ V_{++},\ dT=7)$ and higher for V1ES, $(V_{++},\ V_{++},\ dT=1)$, $(V_{--},\ V_{++},\ dT=1)$. Next, $t^i_p$ value is higher for V1ES, V1LS, and V2ES. The $t^i_{f,p}$ values are higher for V1ES, V1LS, and $(V_{++},\ V_{++},\ dT=1)$, and lower for $(V_{--},\ V_{++},\ dT=7)$. The total change in pupillary response is difficult to understand, but in general, the value is higher for V2LS and lower for the $r^{1/e}_{all}$ of $(V_{--},\ V_{++},\ dT=7)$.

Based on the above consideration, we discuss the strong correlation with the total change in pupillary response for each value pair according to Table IV. The reason for this is that the degree of emotional induction to the presented stimulus is related to the degree of interference of multiple ANEWs, although emotional is induced by ANEWs, as assumed in Section II. When multiple ANEWs are presented to listeners during the interval of a $dT$, the emotions induced by the ANEWs may interfere with each other. If the $dT$ is short enough to cause interference in emotion induction, then the emotions induced by the multiple ANEWs may interfere with each other, and the final emotion that remains in the listener's mind may induce more emotions than the original ANEWs induced, or it may repel the original ANEWs and not induce any emotions at all. The phenomenon that $dT$ does not induce emotion is likely to occur. If the $dT$ is long enough that it cannot cause interference in the emotion induction, then the emotion induced in the listener by multiple ANEWs will be the same as that induced by a single ANEW, or will induce only very weak interference.

In this study, the interference of induced emotion is considered to appear in the total change in pupillary response as shown in Figure 1. The quantities, $t^i_p,\ t^i_{f,p},\ n^i_{p,f}$, which occur in the participants' rating of emotion induction, rate the degree of emotion induced by the presented stimulus received by the listener as a valence, and the result is selected. Additionally, there may be some relationship between the selected $S_p$ and $t^i_p,\ t^i_{f,p},\ n^i_{p,f}$. From this point of view, Tables V and IV show a clear correlation with the total change in pupillary response to the recalled presented stimulus with the valence pair of $(V_{--},\ V_{++},\ dT=7)$, $(V_{++},\ V_{--},\ dT=1)$. Moderate correlations were found for $(V_{--},\ V_{++},\ dT=7)$ and $(V_{++},\ V_{--},\ dT=1)$ to $r^1_{all}-r^e_{all}$ and $r^2_{all}$, respectively, with $t^i_{f,p}$ and $T^i_{c,p}$.

A comparison of this with the results from Table VI is as follows. $(V_{++},\ V_{--},\ dT=1)$ does not show a trend of highly distorted $t^i_p,\ t^i_{f,p}$, and $n^i_{p,f}$ from the overall data. On the other hand, for $(V_{--},\ V_{++},\ dT=7)$, $T^i_{c,p}$ is rather low,



Copyright (c) IARIA, 2025. ISBN: 978-1-68558-260-9 

TABLE VI. SUMMARY OF MEASUREMENT QUANTITIES.

| | | num. of data | num. of recall | ratio | $T^i_{c,p}$ | $t^i_p$ | $t^i_{f,p}$ | $r^1_{all}$ | $r^2_{all}$ | $r^e_{all}$ | $n^i_{p,f}$ |
|---|---|---|---|---|---|---|---|---|---|---|---|
| all | | 544 | 161 | 0.30 | 0.7 | 3411.5 | 2220.33 | 1.92 | 2.02 | 1.53 | 7 |
| $dT$ | 1 | 272 | 85 | 0.31 | 0.77 | 3416 | 2335 | 1.83 | 2 | 1.47 | 7 |
| | 7 | 272 | 76 | 0.28 | *0.63* | 3393 | *1923.68* | 2 | 2.12 | 1.57 | 6 |
| recall | Yes | 161 | | – | 0.71 | 3436 | 2101.3 | 1.7 | 1.71 | 1.46 | 6 |
| | No | 380 | | – | 0.69 | 3401 | 2236.29 | 1.99 | 2.2 | 1.64 | 7 |
| type of evaluation | VES | 149 | 52 | 0.35 | *0.62* | 3433 | 2232 | 1.85 | 2.02 | 1.5 | 6 |
| | V1ES | 10 | 1 | 0.10 | **0.96** | **4880.5** | **3944.27** | *0.94* | 1.94 | *1.12* | 9.5 |
| | V1LS | 19 | 3 | 0.16 | 0.72 | **4240** | **3027.25** | 1.72 | 3.27 | 1.7 | 7 |
| | V2ES | 117 | 35 | 0.30 | 0.79 | 3702 | 2645.18 | 1.92 | 1.79 | 1.41 | 7 |
| | V2LS | 90 | 25 | 0.28 | 0.68 | 3304.5 | 2004.99 | **2.29** | 2.03 | **1.94** | 6 |
| valence pair | $(V_{++},\ V_{++},\ dT=1)$ | 68 | 26 | 0.38 | **0.95** | 3646.5 | **3123.6** | 1.85 | 2.19 | *1.39* | 7 |
| | $(V_{++},\ V_{++},\ dT=7)$ | 68 | 14 | 0.21 | 0.73 | 3616.5 | 2072.66 | 2.09 | 2.31 | 1.71 | 6.5 |
| | $(V_{++},\ V_{--},\ dT=1)$ | 68 | 22 | 0.32 | 0.74 | 3149.5 | 2157.43 | 1.83 | 1.93 | 1.51 | 6.5 |
| | $(V_{--},\ V_{++},\ dT=1)$ | 68 | 19 | 0.28 | **0.81** | 3436 | 2335.55 | 1.89 | 1.55 | 1.43 | 6 |
| | $(V_{--},\ V_{++},\ dT=7)$ | 68 | 21 | 0.31 | *0.64* | 3617.5 | *1890.56* | **2.26** | 1.92 | 1.74 | 8 |
| | $(V_{--},\ V_{--},\ dT=7)$ | 68 | 24 | 0.35 | *0.57* | 3045.5 | *1735.94* | 1.96 | 1.98 | 1.47 | *5* |

$t^i_{f,p}$ is very low, and $r^1_{all}$ is relatively higher. In addition, $r^1_{all}$ shows a slightly larger change than the others. However, all of these valence pairs are comparable to each other in terms of the recall situation for the pair, as shown in Table III.

These results suggest that the pupil's response to a particular value pair and the response to recall may be somehow related to the pupil's response to a particular value pair.

### B. Relation between Recall and Observed Metric at $dT = 1$

Table III shows the relationship between the number of recalls and the observed metrics. According to Table III, $(V_{++},\ V_{++})$ is the most frequently recalled, followed by $(V_{++},\ V_{--})$.

First, we focus on $(V_{++},\ V_{++})$. There is no correlation between the various quantities in Table IV. From Table VI, $r^e_{all}$ shows lower values and $T^i_{c,p}$ shows higher values. The $r^e_{all}$ indicates that the pupil diameter change was almost completely contained, suggesting that significant emotional induction remained. For $T^i_{c,p}$, the participants were gazing almost anywhere on the screen until $S_p$ was determined. This suggests that $(V_{++},\ V_{++})$ rated $S_p$ without significant emotional induction and their interference, but $T^i_{c,p}$ became longer as a result of hesitation over which to select when choosing $S_p$. This may be the reason why $T^i_{c,p}$ is longer than $S_p$. This may have led to cognitive load, and may have promoted recall due to the load.

Next, we focus on $(V_{++},\ V_{--})$. According to Table IV, the total change in pupillary response at the time of recall is correlated with $r^2_{all}$ and $t^i_{f,p},\ n^i_{p,f}$. Considering the difference from the overall data, it shows a correlation between $r^2_{all}$ and $t^i_{f,p}$, and checking the $S_p$ values, 13 were classified as V2ES and five were set as V2LS, indicating that the answers were generally based on the second ANEW impression. According to Table VI, there are no areas that stand out compared to the overall trend. This suggests that the $(V_{++},\ V_{--})$ responses were based on the impressions received from the presented stimuli. In other words, the degree of cognitive load was low, suggesting that the impression was deep and thus facilitated recall.

### C. Relation between recall and observed metric at $dT = 7$

We refer to the relationship between the number of recalls and the observed quantities shown in Table III. According to the table, $(V_{--},\ V_{--})$ is the most frequently recalled, followed by $(V_{--},\ V_{++})$.

First, we focus on $(V_{--},\ V_{--})$. Table IV shows no correlations or strong trends among the metrics. Table VI shows that $T^i_{c,p},\ t^i_{f,p},\ n^i_{p,f}$ both have low values. This indicates that the participants had almost no hesitation in determining $S_p$, and they checked the screen in a short amount of time. In other words, the degree of cognitive load is considered to be low, and may be due to the fact that the impression was deep, which facilitated recall.

Next, we focus on $(V_{--},\ V_{++})$. Table IV shows that the pupil diameter change at the time of recall is correlated with $r^1_{all} - r^e_{all}$ and $t^i_{f,p},\ T^i_{c,p}$. Since both have a common meaning, we can essentially assume that there is a correlation between $T^i_{f,p}$ and $r^1_{all} - r^e_{all}$. The $S_p$ values for the recalled presented stimuli were eight for V2ES and seven for V2LS, indicating that the responses were generally based on the second ANEW impression. According to Table VI, $T^i_{c,p}, t^i_{f,p}$ are very low, while $r^1_{all}$ is high. This suggests that a relatively high emotional induction occurred at the beginning of $(V_{--},\ V_{++})$, but the second ANEW impression was recalled independently of it, with that the difference from the emotional induction being high. This may be due to the cognitive load that facilitated the recall of the second ANEW impression.

### D. Two Trends triggering recall

Based on the above, it can be concluded that there are several types of recall triggered by the presented stimulus with emotion induction.

- The degree of cognitive load due to emotion induction is low, but the stimulus is very impressive, which promotes recall.
- Promotion of recall due to a high degree of cognitive load induced by emotion induction

The former can be regarded as System 1, i.e., intuitive recall due to bias, and the latter as System 2, i.e., recall due to thought, as indicated by Kahneman [16].

In the case of a presented stimulus containing multiple ANEWs, it is necessary to consider emotion induction due to their interference with each other. It is not sufficient to prepare ANEW pairs at any time, nor is it sufficient to continuously prepare ANEW pairs of the same or different types. According to Table III, even for the same $(V_{++},\ V_{++})$ valence pair, $dT = 1$ and $dT = 7$ produce significantly different results for recall. Notably, the recall rate is almost halved for $dT = 7$ compared to $dT = 1$, suggesting a strong influence of interference generated during ANEW's emotion induction.

In this way, one possibility for a set of ANESs that promotes recall may be the valence pair and the placement of ANEWs within the range where they can interfere with each other. However, while many research institutes have made certain measurements of ANEW valence and can predict $S_p$ evaluations based on these measurements, the interference between valences is not the same from person to person due to differences in individual responses to the valences and individual differences in the degree of interference between valences. The final response will differ from person to person. In the future, based on our findings, we believe that it will be possible to design content that is tailored to each individual after measuring the interference characteristics of the individual's emotion induction to the value pair using LGR-Map and other methods.

## V. Conclusion and Future Works

In this paper, we analyzed the relationship between the valence of ANEWs in linguistic information and recall, using the total change in pupil diameter. We sought to understand the degree of emotional interference by the combination of valence pair and interval time between valence pairs for two ANEWs in short sentences, employing metrics such as the time required for impression evaluation, the total time spent during impression evaluation, the ratio of both, and the number of stopping points. Additionally, we also examined the effect of the combination of valence pair and interval time on the recall rate. The results showed that responses with a high recall rate were 1) The response with a high recall rate had a low cognitive load due to emotion induction, but was very impressive, which promoted recall. 2) The response with a high recall rate was associated with a high degree of cognitive load by emotion induction. These findings suggest the presence of two types of memories, namely System 1 and System 2, as described by Kahneman. Specifically, they correspond to bias-induced memory and thought-induced memory.

These characteristics may be intrinsic to each person, or they may be general properties that do not depend on the person. In the future, the development of a metric to discriminate between the two states of recall will be valuable for designing content that is easier to recall according to the individual.

## Acknowledgment

This work was supported by JSPS KAKENHI Grant Number 19K12246 / 19K12232 / 20H04290 / 22K12284 / 23K11334 , and National University Management Reform Promotion Project.

## References

[1] B. Mahanama et al., "Eye movement and pupil measures: A review," Frontiers in Computer Science, vol. 3, 2022. [Online]. Available: https://par.nsf.gov/biblio/10313299

[2] A. Vilotijević and S. Mathôt, "Functional benefits of cognitively driven pupil-size changes," WIREs Cognitive Science, vol. 15, no. 3, 2024, p. e1672. [Online]. Available: https://wires.onlinelibrary.wiley.com/doi/abs/10.1002/wcs.1672

[3] V. Skaramagkas et al., "Review of eye tracking metrics involved in emotional and cognitive processes," IEEE Reviews in Biomedical Engineering, vol. PP, 03 2021, pp. 1–1.

[4] T. Partala and V. Surakka, "Pupil size variation as an indication of affective processing," International Journal of Human-Computer Studies, vol. 59, 07 2003, pp. 185–198.

[5] R. Henderson, M. Bradley, and P. Lang, "Emotional imagery and pupil diameter," Psychophysiology, vol. 55, 12 2017, p. e13050.

[6] M. Oliva and A. Anikin, "Pupil dilation reflects the time course of emotion recognition in human vocalizations," Scientific Reports, vol. 8, 03 2018.

[7] P. Tarnowski, M. Kołodziej, A. Majkowski, and R. Rak, "Eye-tracking analysis for emotion recognition," Computational Intelligence and Neuroscience, vol. 2020, 09 2020, pp. 1–13.

[8] R. Hirabayashi, M. Shino, K. T. Nakahira, and M. Kitajima, "How auditory information presentation timings affect memory when watching omnidirectional movie with audio guide," in Proceedings of the 15th International Joint Conference on Computer Vision, Imaging and Computer Graphics Theory and Applications (VISIGRAPP 2020), vol. 2, 2020, pp. 162–169.

[9] M. Murakami, M. Shino, K. T. Nakahira, and M. Kitajima, "Effects of emotion-induction words on memory of viewing visual stimuli with audio guide," in Proceedings of the 16th International Joint Conference on Computer Vision, Imaging and Computer Graphics Theory and Applications (VISIGRAPP 2021), vol. 2, 2021, pp. 89–100.

[10] K. T. Nakahira, M. Harada, and M. Kitajima, "Analysis of the relationship between subjective difficulty of a task and the efforts put into it using biometric information," in Proceedings of the 17th International Joint Conference on Computer Vision, Imaging and Computer Graphics Theory and Applications, VISIGRAPP 2022, Volume 2: HUCAPP, Online Streaming, February 6-8, 2022, A. Paljic, M. Ziat, and K. Bouatouch, Eds. SCITEPRESS, 2022, pp. 241–248. [Online]. Available: https://doi.org/10.5220/0010906800003124

[11] Y. Kurihara, M. Shino, K. Nakahira, and M. Kitajima, "Visual behavior based on information foraging theory toward designing of auditory information," in Proceedings of the 19th International Joint Conference on Computer Vision, Imaging and Computer Graphics Theory and Applications - HUCAPP, INSTICC. SciTePress, 2024, pp. 530–537.

[12] K. T. Nakahira, M. Harada, S. Moriya, and M. Kitajima, "Local-global reaction map for classifying listeners by pupil response to sentences with emotion induction words and its application to auditory information design," International Journal on Advances in Intelligent Systems, no. 1&2, 2024, pp. 38 – 48.

[13] W. Kintsch, "The use of knowledge in discourse processing: A construction-integration model," Psychological Review, vol. 95, 1988, pp. 163–182.

[14] W. Kintsch, Comprehension: A paradigm for cognition. Cambridge, UK: Cambridge University Press, 1998.

[15] C. Wharton and W. Kintsch, "An overview of construction-integration model: A theory of comprehension as a foundation for a new cognitive architecture," SIGART Bull., vol. 2, no. 4, jul 1991, p. 169–173. [Online]. Available: https://doi.org/10.1145/122344.122379

[16] D. Kahneman, Thinking, Fast and Slow. New York: New York: Farrar, Straus and Giroux, 2011.