

Title	語彙学習のための単語が強調された生成画像における視線と記憶の検証および評価
Author(s)	落合, 卓登
Citation	
Issue Date	2026-03
Type	Thesis or Dissertation
Text version	author
URL	https://hdl.handle.net/10119/20456
Rights	
Description	Supervisor:長谷川 忍, 先端科学技術研究科, 修士(情報科学)

修士論文

語彙学習のための単語が強調された生成画像における
視線と記憶の検証および評価

落合 卓登

主指導教員 長谷川 忍

北陸先端科学技術大学院大学
先端科学技術専攻
(情報科学)

令和8年3月

Abstract

Recent advances in generative artificial intelligence, especially deep learning-based text-to-image generation, have made it possible to create highly realistic images that do not necessarily exist in the real world. Such generated images have attracted attention not only in entertainment and advertising but also in educational contexts, where they may serve as visual materials that support comprehension and memory. In vocabulary learning, the combination of verbal information and visual information has long been regarded as effective, because images can complement linguistic input and help learners build stronger associations between words and meanings. However, most conventional vocabulary learning support systems employ images that correspond to correct answers and present typical semantic representations of target words. By contrast, relatively little research has examined how images generated from learners' incorrect answers influence attention, comprehension, and memory.

This study focuses on error-emphasized generated images used for vocabulary learning. These images are generated from English sentences that include an incorrect word selected by the learner, and they are intended not to present the correct answer directly but to visualize the learner's error in a way that promotes reflection and re-interpretation. The images used in this study were generated through the L-VEIGe (Learning Vocabulary Error Image Generation) framework proposed in prior work. L-VEIGe is designed to support vocabulary learning by presenting images that reflect learners' incorrect answers and by controlling the degree to which the error is represented visually. Although previous studies have suggested that such images may help reduce repeated errors, the cognitive processes underlying this effect have not been sufficiently clarified. In particular, it remains unclear how learners visually explore these images and how such gaze behavior relates to the later recall and retention of the incorrect word and its contextual information.

The purpose of this study is to investigate how error-emphasized generated images affect learners' eye movements and memory retention in vocabulary learning. More specifically, the study examines whether differences in the saliency of generated images influence learners' visual exploration behavior, immediate recall, delayed recall, and subjective impressions. Unlike approaches that attempt to measure whether learners looked directly at a predefined error region, this study emphasizes how learners explore the overall visual context that contains the error. This perspective is based on the assumption that understanding an error may require not only looking at a specific object but also integrating surrounding contextual information represented in the generated image.

To analyze this issue, an eye-tracking experiment was conducted with university students who were native speakers of Japanese and learners of English as a second language. Sixteen participants took part in the experiment. The stimulus set consisted of ten English fill-in-the-blank sentences, each associated with an incorrect word that had been assumed or selected in advance. For each sentence, sixteen error-emphasized generated images were prepared by manipulating four saliency-related parameters at high and low levels: error concept consistency, contextual consistency, surface ratio of the emphasized concept, and color difference of the emphasized concept. These four parameters correspond to two higher-order dimensions. The first is recallability, which concerns how accurately the image represents the incorrect concept and its context. The

second is impressiveness, which concerns how strongly the emphasized concept stands out visually. By combining the high and low values of these four parameters, sixteen image conditions were created for each sentence, resulting in a total of 160 generated images.

In the experiment, each participant viewed ten images, one for each sentence, so that across sixteen participants all sixteen conditions were covered for each sentence. The order of image presentation was randomized across participants in order to reduce sequence effects. Before each image was displayed, the participant read the corresponding sentence containing the incorrect word and advanced to the image presentation by pressing a key. Each image was then displayed for ten seconds against a black background, with the square stimulus centered on the screen. Eye movements were recorded during image presentation using a Tobii Pro Fusion 250 eye tracker. After each trial, participants answered a short impression questionnaire concerning image readability and the ease with which they could notice where the emphasized incorrect word was reflected in the image. After the viewing phase, participants completed an immediate free-recall test in which they were asked to recall the emphasized word and contextual information from the image. A delayed recall test using the same procedure was conducted one week later to evaluate memory retention.

The analysis focused on gaze exploration over the entire image rather than direct fixation on a predefined semantic target. To achieve this, the displayed image area was divided into equal-sized Areas of Interest (AOIs). The main analysis used a 3×3 AOI grid, while a supplementary analysis used a 4×4 AOI grid in order to examine robustness and reduce the effect of central bias. Gaze samples that fell outside the image area, such as those on the black background, were excluded. Based on these AOIs, several indices were calculated, including AOI distribution entropy and average pupil diameter. AOI distribution entropy was used as a measure of how broadly or narrowly gaze was distributed across the image. A larger entropy value indicates a more exploratory viewing pattern in which attention is spread over multiple regions, whereas a smaller value indicates more concentrated viewing. Average pupil diameter was treated as a supplementary measure related to cognitive effort or processing load.

The results were organized from two perspectives: recallability and impressiveness. From the recallability perspective, images with high error concept consistency and high contextual consistency tended to yield more distributed gaze behavior, as reflected in higher AOI entropy. In particular, the condition in which both parameters were high showed the largest entropy values among the recallability conditions, suggesting that learners explored the whole image more actively when the emphasized concept and its surrounding context were represented coherently. These conditions also tended to produce higher immediate recall, delayed recall, and retention rates. In addition, average pupil diameter was larger in the condition where both recallability-related parameters were high, and significant differences were observed between this condition and several other recallability conditions. This result suggests that learners may have engaged in deeper cognitive processing when the image provided both a clear representation of the incorrect concept and a meaningful contextual structure.

From the impressiveness perspective, images with a higher surface ratio of the emphasized concept and greater color difference generally led to stronger subjective impressions and better noticeability of the emphasized element. AOI entropy in these conditions also tended to be higher, indicating that visual emphasis did not merely attract gaze to one local point but could encourage broader exploration of the image. Moreover, immediate and delayed recall tended to be relatively higher for highly impressive conditions, suggesting that visually salient emphasis may function as a memory cue. At the same time, average pupil diameter also tended to be larger in conditions with stronger impressiveness, implying that visual emphasis may increase cognitive effort as learners attempt to interpret the emphasized content.

Taken together, the results suggest that error-emphasized generated images can support vocabulary learning not simply by drawing attention to a wrong element, but by encouraging learners to explore the broader visual context in which the error is embedded. Recallability appears to contribute mainly to semantic understanding and memory retention by making the incorrect concept and its context interpretable, whereas impressiveness appears to contribute mainly to noticeability and subjective salience by making the emphasized element easier to detect. The findings therefore indicate that effective image design for vocabulary learning should not rely solely on visual emphasis; rather, it should balance semantic interpretability and perceptual salience.

This study has several limitations. First, the number of participants was limited, and some results were interpreted as tendencies rather than conclusive effects. Second, AOI-based analysis provides a spatial approximation of gaze behavior but does not necessarily correspond to semantic units in the image. Third, pupil diameter is sensitive to multiple factors, including lighting, blinking, and missing data, and should therefore be interpreted cautiously. Future work should increase the sample size, conduct more comprehensive statistical testing, refine semantic region analysis, and examine interaction effects between recallability and impressiveness in greater detail.

Despite these limitations, this study contributes foundational knowledge about how learners visually and cognitively respond to error-emphasized generated images in vocabulary learning. By integrating generated images, eye-tracking analysis, impression evaluation, and memory tests, the study provides empirical evidence that the design of error-emphasized visual materials can influence not only what learners notice but also how they explore, interpret, and remember incorrect vocabulary. These findings offer a useful basis for the future design of adaptive vocabulary learning systems that employ generative images as reflective and cognitively effective educational materials.

目次

第1章 はじめに	1
第2章 関連研究	3
2.1 視線行動と記憶形成の関係	3
2.2 視線行動と脳活動に基づく記憶研究.....	3
2.3 語彙学習における視覚情報利用と誤答生成画像	4
2.4 想起性と印象性の定義.....	5
2.5 本研究の位置づけと学術的意義.....	6
第3章 誤答画像生成手法.....	8
3.1 誤答画像生成手法の概要	8
3.2 誤答生成画像の生成	8
3.3 誘目性を構成するパラメータ	8
3.3.1 誤り概念の一致度	10
3.3.2 コンテキストの一致度	10
3.3.3 誤り概念表層比率	10
3.3.4 誤り概念色差.....	10
第4章 実験・評価.....	11
4.1 実験の構成とデータセット	11
4.2 誘目性の設計.....	12
4.3 被験者.....	14
4.4 刺激提示方法および被験者への課題内容	14
4.4.1 刺激提示の構成と順序	14
4.4.2 画像提示環境および提示アプリケーション	15
4.4.3 被験者への課題	16
4.5 実験装置および視線計測	16
4.6 実験手順.....	17
4.7 取得データおよび評価指標	18
4.8 視線データ解析方法	20
4.8.1 本研究における AOI 設計.....	21
4.8.2 解析方法	21

第5章 結果	23
5.1 想起性観点における結果	23
5.1.1 想起性観点における視線行動（AOI 分布エントロピー）	23
5.1.2 想起性観点における視線行動（平均瞳孔径）	24
5.1.3 想起性観点における記憶保持	25
5.1.4 想起性観点における印象アンケート	27
5.2 印象性観点における結果	28
5.2.1 印象性観点における視線行動（AOI エントロピー）	28
5.2.2 印象性観点における視線行動（平均瞳孔径）	29
5.2.3 印象性観点における記憶保持	30
5.2.4 印象性観点における印象アンケート	32
第6章 考察	34
6.1 視起性観点における結果	34
6.1.1 想起性観点における視線行動（AOI 分布エントロピー）	34
6.1.2 想起性観点における視線行動（平均瞳孔径）	34
6.1.3 想起性観点における記憶保持	35
6.1.4 想起性観点における印象アンケート	35
6.2 印象性観点における考察	35
6.2.1 印象性観点における視線行動（AOI 分布エントロピー）	35
6.2.2 印象性観点における視線行動（平均瞳孔径）	36
6.2.3 印象性観点における記憶保持	36
6.2.4 印象性観点における印象アンケート	37
6.3 想起性と印象性の役割の違い	37
6.4 本研究の意義	37
6.5 今後の課題	38
第7章 おわりに	39

図目次

図 2.1	L-VEIGe における問題文と選択肢[7]	5
図 2.2	誤答を含んだ生成画像[7]	5
図 3.1	想起性パラメータ	9
図 3.2	印象性パラメータ	10
図 4.1	想起性パラメータの強弱	13
図 4.2	印象性パラメータの強弱	13
図 4.3	提示画面	15
図 4.4	印象アンケート	16
図 4.5	3×3 分割	19
図 4.6	4×4 分割	20
図 5.1	AOI 分布エントロピー（想起性 ABCD）	24
図 5.2	瞳孔径（想起性 ABCD）	25
図 5.3	直後想起テスト（想起性 ABCD）	26
図 5.4	遅延想起テスト（想起性 ABCD）	26
図 5.5	印象アンケート（想起性 ABCD）	28
図 5.6	AOI 分布エントロピー（印象性 abcd）	29
図 5.7	瞳孔径（印象性 abcd）	30
図 5.8	直後想起テスト（印象性 abcd）	31
図 5.9	遅延想起テスト（印象性 abcd）	31
図 5.10	印象アンケート（印象性 abcd）	33

表目次

表 2.1	4つの評価軸	6
表 4.1	実験で用いた英文	11
表 5.1	誘目性パラメータの略表	23
表 5.2	想起テスト（想起性 ABCD）	27
表 5.3	想起テスト（印象性 abcd）	32

第1章 はじめに

近年、人工知能技術、特に深層学習を基盤とする生成 AI の発展により、現実世界には存在しない画像を高い写実性で生成することが可能となっている。このような生成画像は、エンタテインメントや広告分野にとどまらず、教育分野においても教材としての応用が期待されている。とりわけ語学学習の分野では、単語や文章と視覚的情報を組み合わせることにより、学習者の理解促進や記憶定着を支援する試みが数多く行われてきた[1][2][3]。

語彙学習において、視覚情報が記憶形成に有効であることは、認知心理学および教育工学の分野で広く報告されている[2]。視覚的な画像と語彙情報を併用することで、言語情報のみを提示した場合と比較して、語彙の想起率や保持率が向上することが示されている。この背景には、視覚情報と言語情報が異なる符号化経路を通じて記憶に保持されるという理論的枠組みが存在する。そのため、語彙学習において適切な画像を提示することは、学習効果を高める有効な手段の一つであると考えられる。

一方で、従来の語彙学習支援において提示される画像は、多くの場合、正解語に対応した典型的な意味表現を持つものであった[2][3][4]。これに対し、杉田らは、学習者が誤答した単語を含む英文から生成された画像を提示することで、誤りへの気づきを促進し、理解を深める手法を提案している[5][6][7]。このような手法では、誤答箇所が生成画像内で視覚的に強調される点に特徴があり、学習者に違和感や印象を与えることで、誤答の理解や修正を促進することを目的としている。

しかしながら、誤答が強調された生成画像が、学習者の認知過程にどのような影響を与えているのかについては、十分に明らかにされていない。特に、学習者が画像を閲覧する際の視線行動と、その後の記憶保持との関係については、定量的な検証が不足しているのが現状である。誤答箇所が強調された画像に対して、学習者が実際にどの領域をどの程度注視しているのか、またその注視行動が誤答の想起や記憶成績とどのように関連しているのかを明らかにすることは、生成画像を用いた語彙学習支援の有効性を評価する上で重要な課題である。

近年、アイトラッカーを用いた視線計測技術の発展により、画像閲覧時の注視位置、注視回数、停留時間、視線移動パターンといった視線行動を高精度に取得することが可能となっている[9][10]。これらの視線指標は、学習者の注意配分や

情報処理の深さを反映する客観的な指標として用いられており、記憶形成過程との関連性も報告されている[11][12][13][14]. そのため、誤答が強調された生成画像を提示した際の視線行動を計測・分析することは、学習者の認知状態を定量的に把握する有効な手段であると考えられる.

本研究では、語彙学習において誤答した単語を含む英文から生成された「誤答が強調された生成画像」に着目し、これらの画像が学習者の視線行動および記憶保持に与える影響を検証することを目的とする. 具体的には、生成画像における誤答箇所の強調度合い、すなわち誘目性の違いが、誤答箇所への注視行動および誤答内容の想起・記憶成績にどのような変化をもたらすのかを明らかにする.

そのため、誤答生成画像を閲覧する際の視線データをアイトラッカーによって取得し、注視回数や停留時間といった指標を用いて視線行動を分析する. さらに、画像提示後に実施する自由再生テストを通じて、誤答内容の記憶保持を評価する. これらの分析により、誤答が視覚的に強調された生成画像が、学習者の注意喚起および記憶定着にどのような影響を及ぼすのかを定量的に検討する.

本研究の成果は、生成画像を用いた語彙学習支援において、誤答強調の適切な設計指針を示す基礎的知見を提供することが期待される. また、視線行動と記憶成績との関係を明らかにすることで、学習者の認知特性に基づいた効果的な学習支援システムの設計に寄与することを目指す.

第2章 関連研究

本章では、本研究に関連する先行研究について整理する．具体的には、(1) 視線行動と記憶形成の関係に関する研究、(2) 語彙学習における視覚情報利用に関する研究、(3) 誤答を反映した生成画像を用いた語彙学習支援に関する研究を概観し、本研究の位置づけを明確にする．

2.1 視線行動と記憶形成の関係

人間の視線行動は、注意の配分や情報処理過程を反映する重要な指標であり、近年ではアイトラッカーを用いた視線計測により、これらを定量的に評価する研究が数多く行われている[9][10][11][12]．特に、注視 (fixation) や停留時間といった視線指標は、視覚情報が実際に認知処理されている状態を示すものとされ、記憶形成との関連性が指摘されている．

Fehlman らは、視覚探索中の視線行動とその後の記憶想起能力との関係を検証するため、アイトラッキングと自由再生課題を組み合わせた実験を行った[14]．この研究では、被験者に複雑な視覚刺激を提示し、探索中の注視回数、注視持続時間、まばたきの長さ、視線間距離などの指標を取得した．その後、数日後に自由再生テストを実施し、記憶成績との関係を分析した結果、注視頻度が高い被験者ほど、再生される情報量が有意に多いことが示された．この結果は、視線行動が単なる視覚的反応ではなく、記憶形成そのものに寄与する重要な要因であることを示唆している．

さらに、Fehlmann らは、視線誘導を用いて探索パターンを操作する実験も行っており、特定の領域に視線が誘導された条件では、その領域に対応する情報の記憶成績が向上することを報告している[14]．この知見は、視線行動と記憶との間に因果的な関係が存在する可能性を示すものであり、視線指標を用いた学習効果評価の妥当性を裏付けるものである．

2.2 視線行動と脳活動に基づく記憶研究

視線行動と記憶形成の関係は、脳活動計測と組み合わせた研究においても検証されている．Hannula らは、視線行動と海馬活動との関連に着目し、アイトラッキングと脳活動計測を同時に行う実験を実施した [13]．

この研究では、被験者に複数のシーン画像を提示し、その後、再認課題を用いて記憶の成否を評価した。提示時の視線行動を分析した結果、後に正しく記憶されたオブジェクトに対しては、提示中の注視回数および注視時間が有意に長いことが明らかとなった。さらに、これらの注視行動と同時に、記憶形成に参与する海馬領域の活動が強く活性化していることが報告されている [13]。

この結果は、視線行動が記憶の結果を単に反映する指標ではなく、記憶形成過程そのものと密接に関係していることを示している。そのため、学習時の視線行動を計測・分析することは、記憶定着の程度を評価する上で有効な手法であると考えられる。

2.3 語彙学習における視覚情報利用と誤答生成画像

語彙学習において、視覚情報を併用することの有効性は、教育工学および第二言語習得研究の分野で広く認識されている。視覚的な画像と言語的なテキスト情報を同時に提示することで、語彙の理解や記憶保持が促進されることが示されてきた。

これらの研究を背景として、杉田らは、語彙学習における新たなアプローチとして、学習者が誤答した単語を含む英文から生成された画像を提示する語彙学習支援システム L-VEIGe (=Learning Vocabulary Error Image Generation) を提案している [7]。この手法では、学習者の誤答を反映した生成画像を提示することで、誤りへの気づきを促し、理解の深化を図ることを目的としている。L-VEIGe における問題文を図 2.1 に、誤答した場合に生成され、表示される画像例を図 2.2 に示す。

L-VEIGe の先行研究では、誤答生成画像の提示によって、繰り返しの誤答が有意に減少することが示されている一方で、誤答内容そのものの理解に関しては十分な改善が見られないケースも報告されている [5]。この要因として、生成画像において誤答理解に必要な情報が欠落する「特徴消失問題」が指摘されている。この問題に対処するため、杉田らは、誤答生成画像を定量的に評価する枠組みとして、後述の想起性および印象性を導入している。これらの指標は、生成画像が誤答理解と注意喚起の双方をどの程度満たしているかを評価するための基礎的尺度であり、誤答生成画像の品質評価および制御に用いられる。

しかしながら、杉田らの研究では、誤答生成画像に対する学習者の視線行動が、これらの評価指標や実際の記憶保持とどのように対応しているのかについては、

十分な実証が行われていない。特に、誤答が強調された生成画像において、学習者がどの領域を注視し、その注視行動が誤答内容の想起や記憶成績にどのように影響するのかについては、明確な検証が求められている。

Question	A large () and clocktower building sits in front of an ocean with ferry and sailboats riding past. (1) cane (2) camera (3) cathedral
	

図 2.1 L-VEIGe における問題文と選択肢[7]

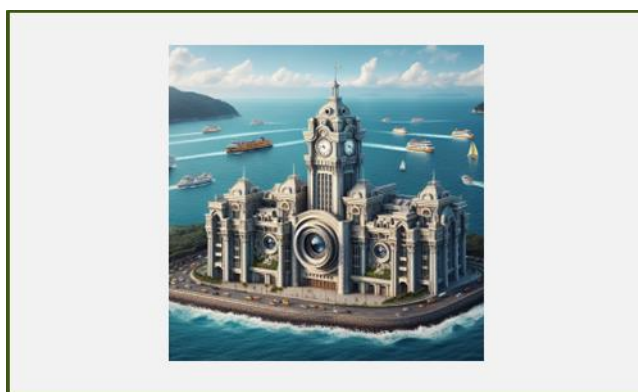


図 2.2 誤答を含んだ生成画像[7]

2.4 想起性と印象性の定義

杉田らは、誤答生成画像の学習効果を評価するために、「想起性」と「印象性」という二つの概念を定義している[7]。これらは、生成画像において誤答箇所がどの程度学習者の注意を引き付けるかを表す概念であり、誤答理解および記憶定着に影響を与える重要な要因であるとされている。想起性とは、誤答生成画像が、学習者に対して誤答内容をどの程度正確に想起させる情報を有しているかを示す指標である。これは主に物理的忠実性の観点から定義され、生成画像に含まれる視覚情報と、誤答を含む元の文章との一貫性に着目する[7]。杉田らは、想起性を構成する要素として、「誤り概念の一致度」と「コンテキストの一致度」を挙げている。これらは、「誤り概念」および「文脈性」という軸で要素を分解したものである。対応を表 2.1 に示す。

「誤り概念の一致度」とは、生成画像が学習者の誤答語彙（誤り概念）をどの程度正確に視覚的に表現しているかを示す指標であり、「コンテキストの一致度」とは、誤答が生じた文脈や状況が画像内でどの程度再現されているかを表す指標である。これらが高い場合、生成画像は学習者にとって誤答内容を理解しやすく、後の想起を促進する可能性が高いと考えられる。

一方、印象性とは、生成画像が学習者に対して誤りへの気づきをどの程度強く与えるかを示す指標であり、学習者が誤りに気づきやすい可視化表現の観点から定義される[7]。印象性は、学習者の注意を誤答箇所に向けさせ、誤りを印象付ける役割を担う。杉田らは、印象性を構成する要素として、「誤り概念表層比率」と「誤り概念色差」を設定している。こちらも「誤り概念」および「文脈性」の軸で要素を分解したものであり、対応を表 2.1 に示す。

「誤り概念表層比率」は、生成画像全体に対して誤答に対応する視覚要素が占める割合を示す指標であり、誤り概念が画像内でどの程度顕在化しているかを表す。また、「誤り概念色差」は、誤答に対応する領域とその他の領域との色彩的差異を表す指標であり、誤答箇所の視覚的強調度合いを反映する。これらの値が大きいほど、学習者は誤りに気づきやすくなる一方で、過度な強調は認知的負荷を高める可能性も指摘されている[15][16]。

本研究では誘目性をこれら 4 つから構成し、指標とし評価していく。

表 2.1 4 つの評価軸

	誤り概念	文脈性
想起性	誤り概念の一致度	コンテキストの一致度
印象性	誤り概念表層比率	誤り概念色差

2.5 本研究の位置づけと学術的意義

これまでの先行研究により、視線行動が記憶形成と強い関連を持つこと、また誤答生成画像が語彙学習支援に有効である可能性が示されてきた。しかし、誤答が強調された生成画像に対する視線行動と記憶保持の関係を、同一実験内で同時に定量評価した研究は、これまでに十分に報告されていない。

本研究は、L-VEIGe において定義された想起性および印象性の概念を前提とし、誤答生成画像における誘目性の違いが学習者の視線行動および記憶保持に与える影響を実験的に検証する点に特徴がある。具体的には、誤答強調生成画像

の提示，アイトラッカーによる注視・停留・視線移動の計測，ならびに直後に自由再生テストで，誤答強調，視線集中，記憶定着という一連の認知プロセスを定量データに基づいて明らかにする．これにより，従来は設計指標として提示されるにとどまっていた想起性・印象性が，実際の学習者の視線行動および記憶成績とどのように対応するのかを検討し，語彙学習支援における生成画像設計の妥当性および設計指針に対する工学的知見を提供することを目的とする．

第3章 誤答画像生成手法

本章では、先行研究において提案されている誤答画像生成手法について述べる。本研究では、誤答が強調された生成画像を用いた語彙学習において、生成画像が持つ誘目性の違いが学習者の視線行動および記憶保持に与える影響を検証することを目的とする。そのために、誤答生成画像の生成方法、誘目性を構成するパラメータの定義について説明する。

3.1 誤答画像生成手法の概要

本研究では、英語語彙空所補充問題を基盤とし、学習者が誤答したと想定された単語を含む英文から生成された「誤答生成画像」を提示する。誤答生成画像は、杉田ら[7]が提案した語彙学習支援システム L-VEIGe を通して生成されるものであり、学習者の誤答を視覚的に可視化することで、誤りへの内省を促すことを目的としている。

3.2 誤答生成画像の生成

本研究で用いる誤答生成画像は、学習者が誤答したと想定される単語を含む英文を入力として生成される。この画像は、正答を示すものではなく、あくまで誤答を反映した視覚表現として提示される。

誤答生成画像は、L-VEIGe において構築された評価モデルにより分析され、各画像に対して誘目性を構成する 4 つのパラメータが付随情報として付与される。本研究では、生成画像そのものの生成アルゴリズムや内部モデルの詳細については扱わず、既存の L-VEIGe の仕組みを前提として利用する。つまり、本研究は画像生成技術の提案ではなく、生成された画像が学習者に与える認知的影響の評価に焦点を当てる。

3.3 誘目性を構成するパラメータ

本研究では、生成画像における誤答箇所の誘目性を操作するため、先行研究で定義された「誤り概念の一致度」、「コンテキストの一致度」、「誤り概念表層比率」、「誤り概念色差」の 4 つのパラメータを用いる。図 3.1 に示す想起性パラメータは「誤り概念の一致度」、「コンテキストの一致度」に分けられ、図 3.2 に示

す印象性パラメータは「誤り概念表層比率」,「誤り概念色差」に分けられる[7].
本節では, これらのパラメータを本研究においてどのように扱うかを説明する.

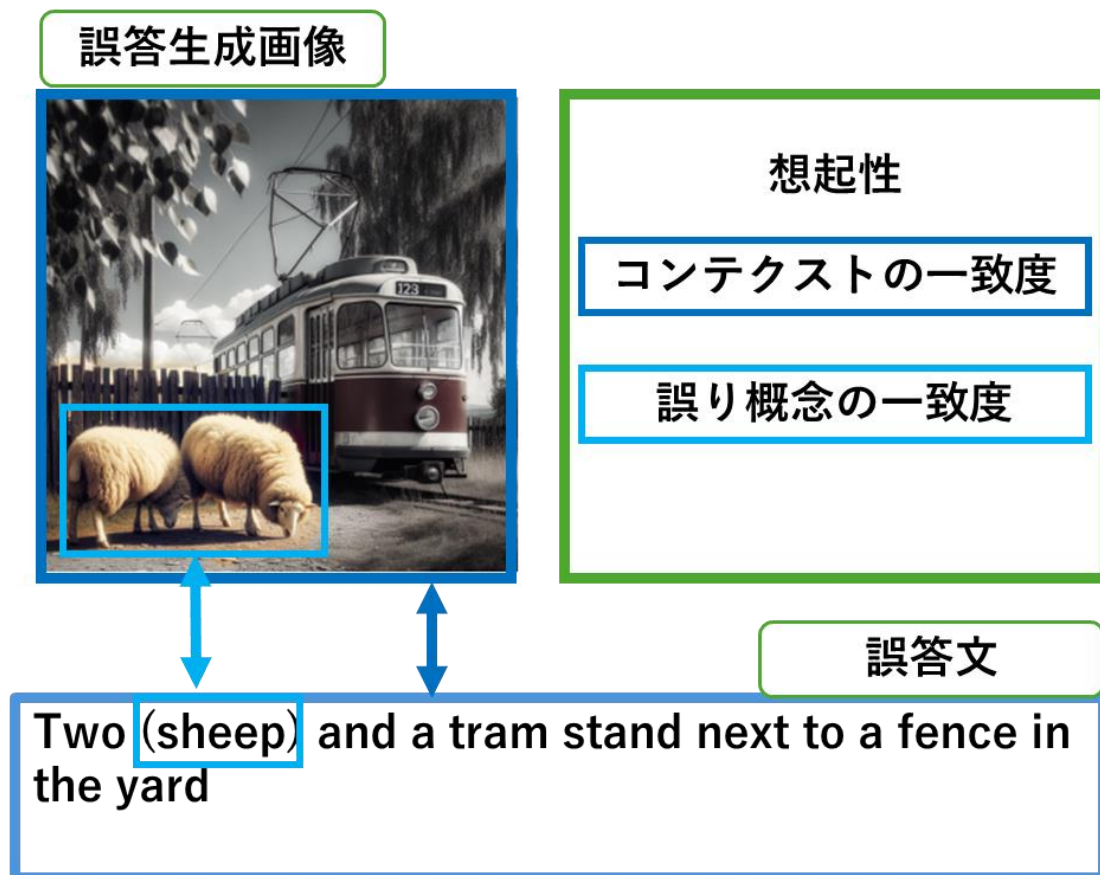


図 3.1 想起性パラメータ

誤答生成画像



印象性

誤り概念表層比率

誤り概念色差

図 3.2 印象性パラメータ

3.3.1 誤り概念の一致度

誤り概念の一致度とは、生成画像が学習者の誤答語彙（＝誤り概念）をどの程度正確に視覚的に表現しているかを示す指標である。この値は 0 から 1 の範囲を取り、値が大きいほど、画像内に誤り概念が明確に反映されていることを示す。

3.3.2 コンテキストの一致度

コンテキストの一致度とは、誤答が生じた英文の文脈や状況が、生成画像内でどの程度再現されているかを示す指標である。誤り概念の一致度と同様に、0 から 1 の範囲の値を取り、値が高いほど文脈的整合性が高いことを意味する。

3.3.3 誤り概念表層比率

誤り概念表層比率とは、生成画像全体に対して、誤答に対応する視覚要素が占める割合を示す指標である。この値は 0 から 100 の範囲を取り、誤り概念に該当する領域の量的な顕在化の度合いを表す。

3.3.4 誤り概念色差

誤り概念色差とは、誤答に対応する領域と、それ以外の領域との色彩的差異を示す指標である。この値は 0 以上 1 未満の範囲を取り、値が大きいほど、誤答箇所が視覚的に強調されていることを示す。

第4章 実験・評価

本章では、本研究で実施する実験の設計および方法について詳細に述べる。本研究は、誤答が強調された生成画像における誘目性のパラメータの違いが、学習者の視線行動および記憶保持に与える影響を検証することを目的としている。そのため、本章では、実験の構成、誘目性パラメータの設計、被験者、実験手順、取得データおよび評価指標を示す。本実験は、ライフサイエンス委員会の承認（承認番号：人 07-041）を得た上で実施している。

4.1 実験の構成とデータセット

本研究で用いる実験刺激は、杉田らが提案した語彙学習支援システム L-VEIGe を通して生成された誤答強調生成画像である[7]。刺激生成の基となる英文は、英語語彙空所補充問題から選定された 10 文であり、各英文に対して、学習者が誤答したと想定される語彙を事前に空所に挿入することで誤答強調文を構成している。その英文の一覧を表 4.1 に示す。() で囲われた語彙が、本来の正答ではない誤答の単語で、強調された単語である。これらの語彙について、名詞が 6 つ、動詞が 2 つ、形容詞が 2 つの構成となっている。

誤答強調画像の生成手法の特徴は、誤答強調生成画像に対して、強調箇所の誘目性を定量的に表す複数のパラメータを付与し、それらを段階的に制御した画像を用いる点にある。本研究では、誘目性を構成する要素として、「強調概念の一致度」、「コンテキストの一致度」、「強調概念表層比率」、「強調概念色差」の 4 つのパラメータに着目する。これらのパラメータの組み合わせによって生成された画像に対する視線行動および記憶成績を比較することで、生成画像における誘目性設計の妥当性を検証する。

表 4.1 実験で用いた英文

1	a table covered with bananas and a (cocoa)
---	--

2	A large tray is filled with broccoli and (honk) of cheese.
3	A sheep standing on top of a (dock).
4	Legs on a skateboard going down a small (camp).
5	A cup with a bird on it and a cup with a (golf) on it.
6	Two sheep and a (dam) stand next to a fence in the yard.
7	Three stuffed bears (hanging) and sitting on a blue pillow.
8	A man with glasses and sideburns (frightening) his necktie that has lights on it.
9	Many motorcycles are displayed in the store with (prescriptive) signs posted near them.
10	An (infernal) baseball game is held in an open field.

4.2 誘目性の設計

本研究では、生成画像における強調箇所の誘目性を構成する要因として、以下の4つのパラメータを用いる。

- ・ 強調概念の一致度
- ・ コンテキストの一致度
- ・ 強調概念表層比率
- ・ 強調概念色差

想起性パラメータについて、図 4.1 に”A large () and clocktower building sits in front of an ocean with ferry and sailboats riding past.”を例に挙げて示す。これは、正答”cathedral”に対し、誤答”camera”を入れた場合に出力された画像で、各パラメータの強弱が異なる画像の比較である。

印象性パラメータについて、同様に図 4.2 に” A church with a () glassed window of the virgin mary.”を例に挙げて示す。これは、正答”stain”に対し、誤答”stair”を入れた場合に出力された画像で、各パラメータの強弱が異なる画像の比較である。

研究では、1つの英文に対して、誘目性を構成する4つのパラメータについて、それぞれ「高い」「低い」の2段階を設定する。これにより、

- ・ 強調概念の一致度（高・低）
- ・ コンテキストの一致度（高・低）
- ・ 強調概念表層比率（高・低）
- ・ 強調概念色差（高・低）

の組み合わせから、1つの英文につき合計16通りの強調生成画像を用意する。本研究で用いる刺激画像の総数は、これら全10文からなる160枚となる。

ただし、「高い」「低い」の判定は、パラメータの絶対値によって一律に定義されるものではない。

これにより、強調生成画像の誘目性を多角的に操作した条件を構築し、単一の要因ではなく、複数の要因が視線行動および記憶保持に与える影響を総合的に検証することが可能となる。

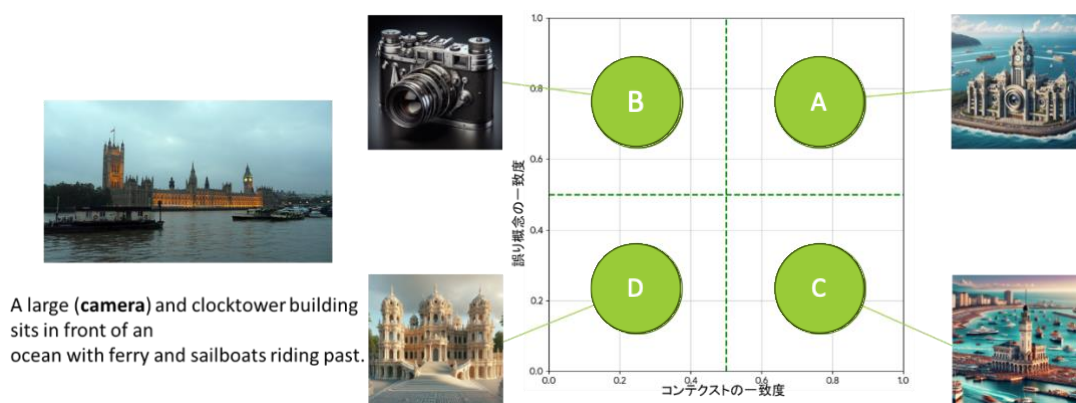


図 4.1 想起性パラメータの強弱

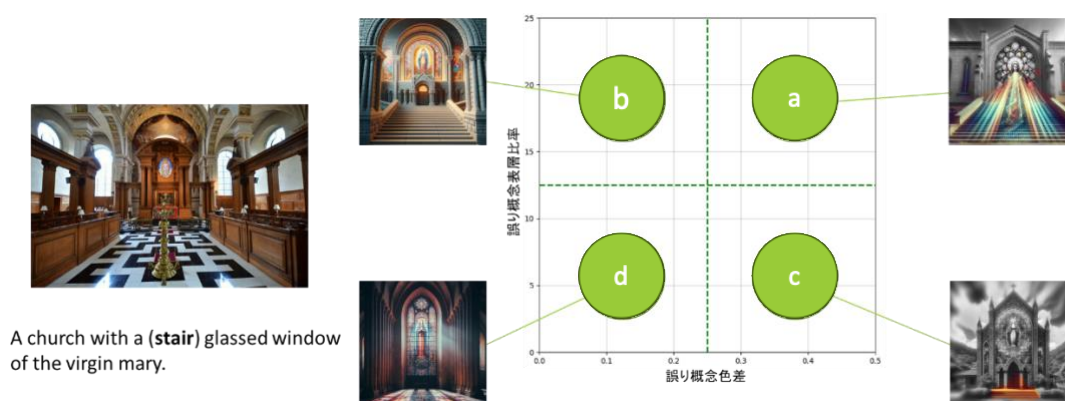


図 4.2 印象性パラメータの強弱

なお、本研究で用いる強調生成画像は、推論モデルによって生成されており、モデルの特性や推論精度の限界に起因して、すべての条件において完全に一貫した視覚表現が得られるとは限らない。そのため、本研究で用いるデータセットは、すべての生成画像について人手による確認を行い、16通りの条件に対応する画像が妥当であるかを検証した上で構築されている。

この過程において、一部の画像には、推論モデルの特性により強調概念が十分に視覚化されないケースや、パラメータ値が他の画像と比較して逸脱する外れ

値が含まれる場合がある。特に、強調概念の一致度が低い条件では、強調概念が視覚的に表象されにくく、強調概念表層比率や強調概念色差の算出が困難となる場合が確認されている。

これらの理由から、本研究では、誘目性パラメータの高低を機械的な固定閾値によって一律に定義するのではなく、データ分布を参照しながら英文ごとに相対的な高低を決定する方針を採用する。このような処理により、推論モデルの限界や外れ値の影響を考慮しつつ、誘目性条件の比較を行うことを可能とする。

4.3 被験者

本研究の被験者は、日本語を母語とし、第二言語として英語を学習している大学生とする。対象者は 16 名であり、全員が英語学習経験を有する者であった。被験者選定の理由は、第二言語としての英語語彙学習という研究目的に適合した対象群を統一的に確保するためである。

被験者からは、年齢、性別、英語学習経験年数、TOEIC スコアの基礎属性情報を事前アンケートにより収集した。

4.4 刺激提示方法および被験者への課題内容

本節では、刺激画像を被験者にどのように提示し、その際に被験者に与えた課題や収集したデータについて説明する。本研究では、画像提示の順序や条件による偏りを最小化しつつ、学習者の自然な視線行動および記憶形成過程を観察することを目的として実験設計を行った。

4.4.1 刺激提示の構成と順序

各被験者には、10 文の英文それぞれに対応する強調生成画像を 1 枚ずつ提示し、被験者 1 名あたりの閲覧枚数は合計 10 枚であった。同一被験者に対して同一画像を複数回提示することはない。

各英文には、4 つのパラメータの高低の組み合わせで 16 条件パターンの画像が存在するが、本実験では被験者数を 16 名とし、被験者間で条件を割り当てることで、1 英文あたり 16 条件が網羅されるように構成した。

提示順序は被験者ごとに変化させ、特定の英文や条件が連続することによる順序効果を抑制した。提示順の生成方法は、乱数によるシャッフルを用いる。

4.4.2 画像提示環境および提示アプリケーション

画像提示および視線計測には、実験用に自作した画像提示アプリケーションを使用した。本アプリケーションは、Python を用いて開発されており、アイトラッカーと連携することで、画像提示と同時に視線データを取得できる構成となっている。

画像提示および視線計測には自作アプリケーションを用いた。アプリケーションは Python で実装し、画面をフルスクリーン表示した上で黒背景を基調とし、正方形の刺激画像を画面中央に提示する。モニタのアスペクト比の都合で左右に余白が生じるため黒で埋め、視線解析の対象は刺激画像領域のみとした。

画像提示の直前に英文を画面に提示し、被験者は内容を任意の時間で確認したのち、キー入力により画像提示へ進む。画像提示は 10 秒間固定で行い、提示中の視線データを連続取得する。実際に提示した画面を図 4.3 に示す。

画像の表示時間は 1 枚あたり 10 秒間固定とし、表示開始と終了のタイミングはアプリケーション側で制御・記録される。

具体的なシステム内容は、まず黒一色の背景画像のみを表示し、被験者が任意のタイミングでスペースキーを押すと、画像が 10 秒間表示される仕組みとなっている。



図 4.3 提示画面

4.4.3 被験者への課題

各画像提示終了後には、簡単な事後アンケートを挟み、主観的な印象（画像の見やすさ、誤答強調箇所への気づきやすさ）について回答を求めた。これを印象アンケートとし、「画像は見やすかったか」および「誤っている単語が画像のどこに反映されているか気づくことができたか」を 5 段階で回答させた。これらはそれぞれ、画像の見やすさは想起性、強調箇所への気づきやすさは印象性に対応した問となっており、視線行動データと主観評価との対応関係を補助的に分析できるようにした。事後アンケートについて、図 4.4 に示す。これらの問を英文 10 文分用意した。

さらに、視線計測が終了した後に「直後想起テスト」を実施し、強調箇所の記憶保持を評価した。これらの記憶課題では、視線計測時に見せた画像をもう 1 度見せ、強調単語および文章の想起をさせる。加えて、1 週間後に再度同様の想起テストである「遅延想起テスト」を実施し、想起率および保持率を算出する。

Q1：画像は見やすかったか、

1. 見やすかった, 2. 少し見やすかった, 3. 普通, 4. 少し見づらかった, 5. 見づらかった

Q2：間違っている単語が画像のどこに反映されているか気付くことができたか、

1. すぐに気付くことができた,
2. 少し時間はかかったが気付くことができた,
3. 時間はかかったが気付くことができた,
4. 気付くのに時間いっぱい使ってしまった,
5. 気付かなかった

図 4.4 印象アンケート

4.5 実験装置および視線計測

使用する実験装置を表 4.3 にまとめる。視線計測には、アイトラッカーの Tobii Pro Fusion 250[17]を用いる。被験者は実験室内に設置された PC モニターの前に着席し、画像提示中の視線位置がリアルタイムで記録される。

表 4.3 実験装置

アイトラッカー	Tobii Pro Fusion 250
---------	----------------------

(Eye Tracker)	
サンプリングレート (Sampling Rate)	120Hz
計測技術 (Measurement Technology)	瞳孔角膜反射法 (Pupil core Corneal Reflection)(ビデオベース)
精度 / 精密密度 (Accuracy / Precision)	0.3° / 0.04° RMS(最適な条件下)
計測距離 (Operating Distance)	65cm に固定
モニター(Monitor)	WQHD 液晶モニター(解像度:2560×1440px), 27 インチ

4.6 実験手順

実験は以下の手順で実行した。

1. 実験説明および同意取得

実験の目的および内容を説明し，同意書への署名を得る（約 5 分）。

2. 事前アンケート

年齢，性別，英語学習歴，TOEIC スコアなどを記入する（約 5 分）。

3. 視線計測予備実験

アイトラッカーの調整および例示画像を用いた練習を行う（約 10 分）。

4. 視線計測実験

強調生成画像と対応する英文を 10 題提示する。各画像は 10 秒間表示され，その間の視線データを取得する。各画像提示後には，短い事後アンケートを挟み，休憩を兼ねる（約 10 分）。

5. 休憩

約 15 分間の休憩を設ける。

6. 直後想起テスト

先ほどと同じ画像をランダムな順で提示し，記憶されている英文または関連単語を回答させる（約 20 分）。

7. 遅延想起テスト

直後想起テストと同様の手法で回答させる。（約 20 分）

被験者 1 名あたりの総所要時間は，説明，休憩を含めて約 60 分である。

4.7 取得データおよび評価指標

本研究で取得するデータは以下の通りである．

- 視線計測データおよび AOI (=Area of Interest) 指標 (AOI 滞在時間・割合, エントロピー, 遷移指標等)
- 想起テストの回答内容と正答率・想起率
- 事後アンケートによる 5 段階の主観評価 (画像の見やすさ, 強調箇所への気づきやすさ) (図 4.4 参照)
- 被験者属性情報

視線データは, 刺激画像領域内の視線サンプルを対象として分析する．本研究では強調箇所を事前に領域指定するのではなく, 刺激画像を等面積の AOI グリッドに分割し, 画像全体の文脈に対する視線探索の偏りを定量化する．主分析では図 4.5 のような 3×3 分割を用い, 中央偏りの影響を検討する補助分析としては図 4.6 のような 4×4 分割も併用する．

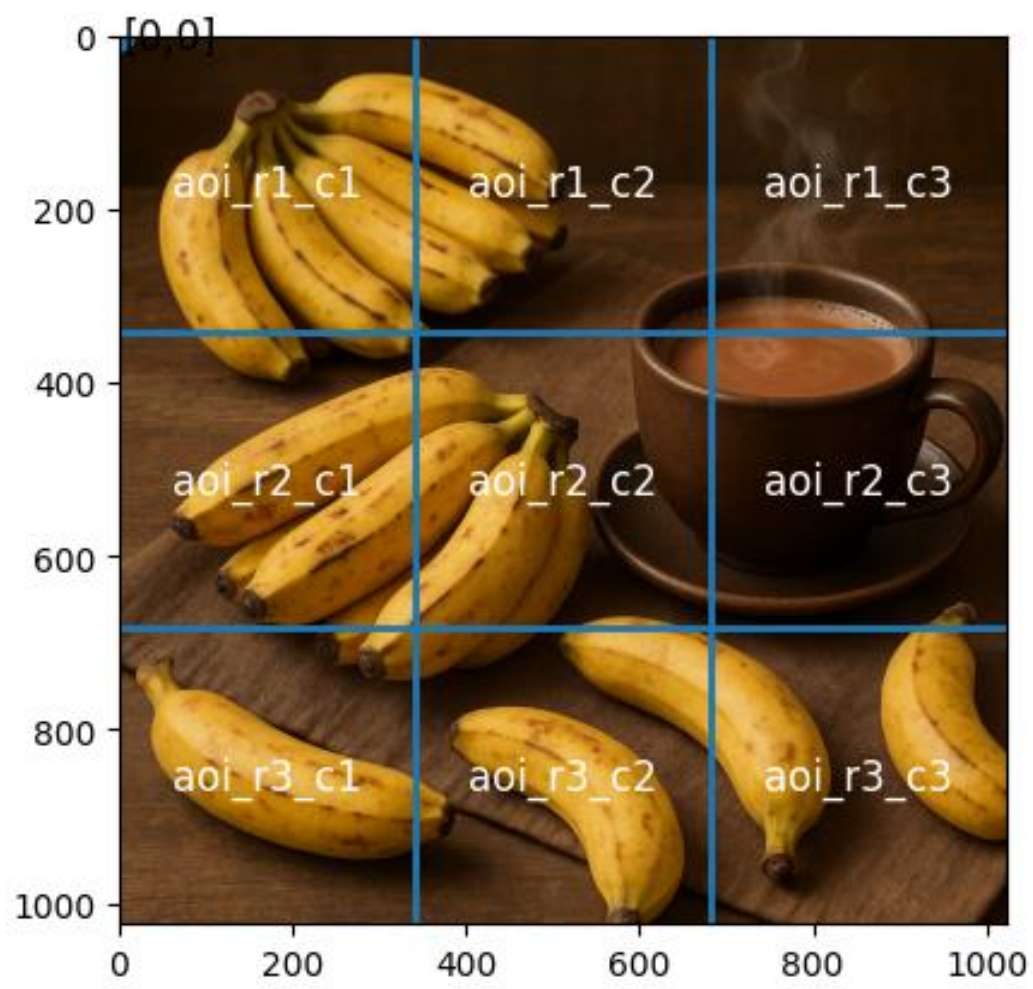


图 4.5 3×3 分割

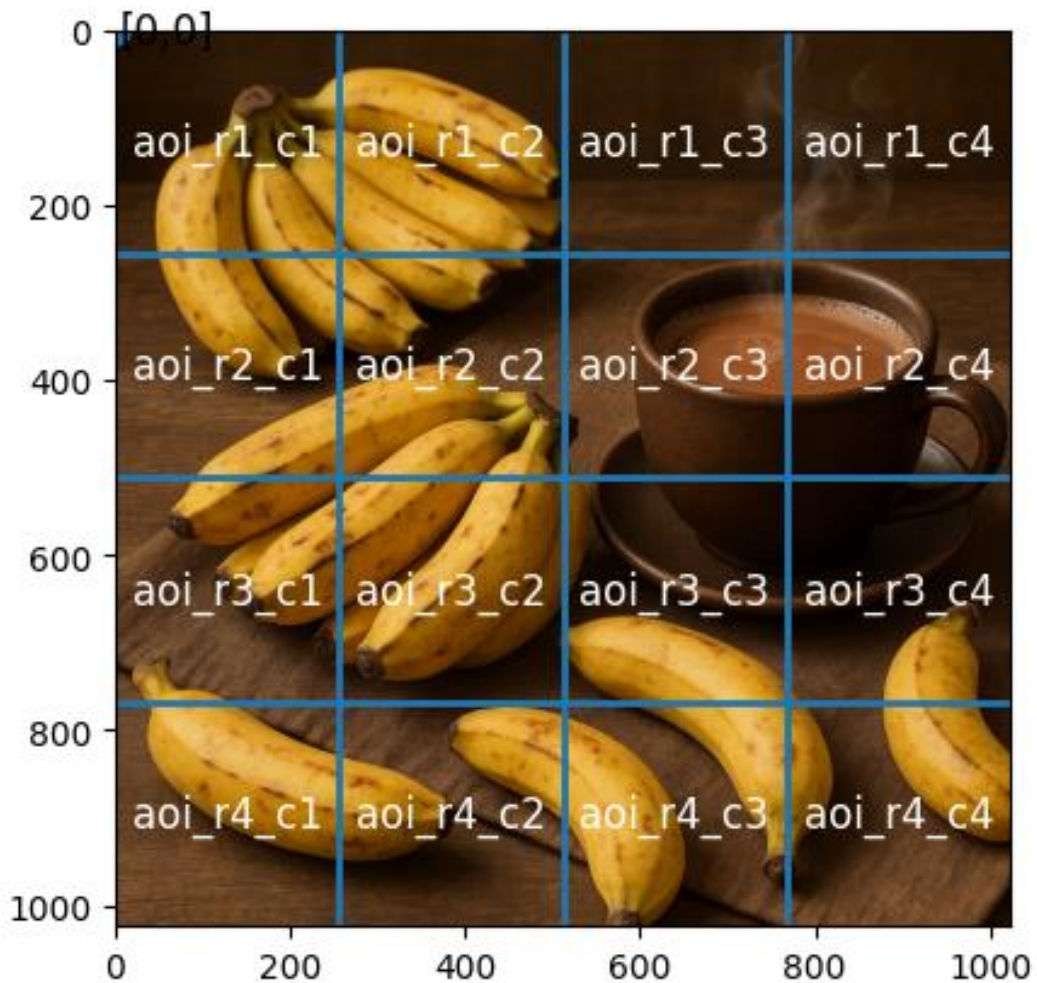


図 4.6 4×4 分割

4.8 視線データ解析方法

本研究では、強調箇所そのものの注視量ではなく、強調箇所を含む刺激画像全体に対して学習者がどのように視線探索を行うかを評価するため、AOI (=Area of Interest) グリッドに基づく解析を採用する。AOI とは、刺激画像や画面領域を複数の部分領域に分割し、各領域に対する視線行動を定量的に分析するための枠組みである。AOI を用いることで、視線位置データを空間的に整理し、どの領域にどの程度注意が向けられていたかを比較可能な指標として扱うことができる。

視線研究においては、AOI 内への注視時間、注視頻度、注視遷移などが、注意配分や情報処理過程を反映する代表的な指標として用いられている[9][18]。

AOI に基づく分析は、特定の意味単位や注視対象を事前に仮定せずに、視線

行動の空間的特徴を捉えられる点に利点がある。特に、強調生成画像のように、注視すべき領域が必ずしも一意に定まらない刺激に対しては、AOI を用いた分布的分析が有効であると考えられる。

4.8.1 本研究における AOI 設計

本研究の目的は、強調単語そのものではなく、強調箇所を含む画像全体の文脈に対する視線探索の様相を捉えることである。

このように異なる解像度の AOI 分割を比較する手法は、視線分布の頑健性を検証するために有効であることが報告されている[18]。

4.8.2 解析方法

視線データは Tobii アイトラッカーから取得した視線データを用い、提示中の画像ごとに CSV として保存する。gaze_x/gaze_y は画面上の正規化座標(0~1)で取得されるため、画像提示時と同一のリサイズ・中央配置条件に基づき、視線座標を刺激画像のピクセル座標へ変換する。黒背景領域に落ちたサンプルは解析対象から除外する。

AOI 分布エントロピー：AOI 滞在割合ベクトルから情報量を算出し、AOI 数で正規化したエントロピー (0~1) も併せて求める。値が大きいほど視線が画像全体に分散していることを示す。AOI 滞在割合を p_i とすると以下の数式で表すことができる。

$$H = - \sum_i p_i \log p_i$$

平均瞳孔径：補助的指標として、AOI ごとの平均瞳孔径および画像全体の平均瞳孔径を算出する[19]。

各誘目性条件における AOI 滞在時間・滞在割合を算出し、条件間で比較を行う。主分析 (3×3) では画像全体の注意分布の傾向を把握し、補助分析 (4×4) では中央 AOI への偏りが結果解釈に与える影響を検討する。結果は、AOI ごとの滞在割合を示す図表として整理する。

想起テストの集計について、強調単語のみの回答は正誤のみ判定、文章や文脈の自由記述の回答は 3 段階で集計し、基準は以下とした。

- ・ 2：句レベル（例 “a cup with cocoa” など、関係が説明できる）

- ・ 1：関連語レベル (cup, drink など方向性は合っている)
- ・ 0：無関係/空欄

なお、句レベルまたは関連語レベルかは実験者が判定を行い、句レベルで想像したものが画像と一致していれば 2, 語のみの羅列や、例えば原文”two sheep”に対して”sheep”のみの回答であれば 1 とした。想起スコアは、これらの平均値とした。

また、統計的有意差を確認するために、Welch の t 検定を行った。想起性および印象性の指標は各 4 つあるため、Holm 補正を行った。補正後、 $p < 0.05$ となる組み合わせを有意とした。

第5章 結果

本章では、強調生成画像に付与された誘目性パラメータと、学習者の視線行動、記憶保持、および主観的印象との関係について示す。結果は、想起性の観点と印象性の観点に分けて整理する。想起性の観点では、強調概念の一致度とコンテキストの一致度の組み合わせにより 図 4.1 で示した A～D の 4 条件を設定し、印象性の観点では、強調概念表層比率と強調概念色差の組み合わせにより図 4.2 で示した a～d の 4 条件を設定した。

被験者 1 名につき 10 枚の画像を提示し、16 名で合計 160 試行を実施した。画像提示時には視線計測を行い、その後、想起テストおよび印象アンケートを実施した。

また誘目性パラメータは、表 5.1 に示す変数で表すこととする。

表 5.1 誘目性パラメータの略表

画像の パラメータ	コンテキストの一致度	強調概念の一致度	強調概念表層比率	強調概念色差
高	conH	errH	srH	ctH
低	conL	errL	srL	ctL

5.1 想起性観点における結果

5.1.1 想起性観点における視線行動（AOI 分布エントロピー）

図 5.1 に、想起性条件 A～D における AOI 分布エントロピーの平均値を示す。AOI 分布エントロピーは、視線が画像全体の複数領域にどの程度分散していたかを表す指標であり、値が高いほど、学習者が画像全体を探索的に見ていたことを意味する。

結果を見ると、A 条件が最も高く、続いて B, D, C の順となった。すなわち、強調概念の一致度とコンテキストの一致度がともに高い条件で、視線がより広く分散する傾向が見られた。一方で、両者の少なくとも一方が低い条件では、視線の分散は相対的に小さくなった。

このことから、想起性が高い画像では、学習者が画像全体を用いて強調内容を探索し、文脈的に処理していた可能性がある。ただし、統計的に有意であるかについてはデータ数を増やした今後の検証が必要である。

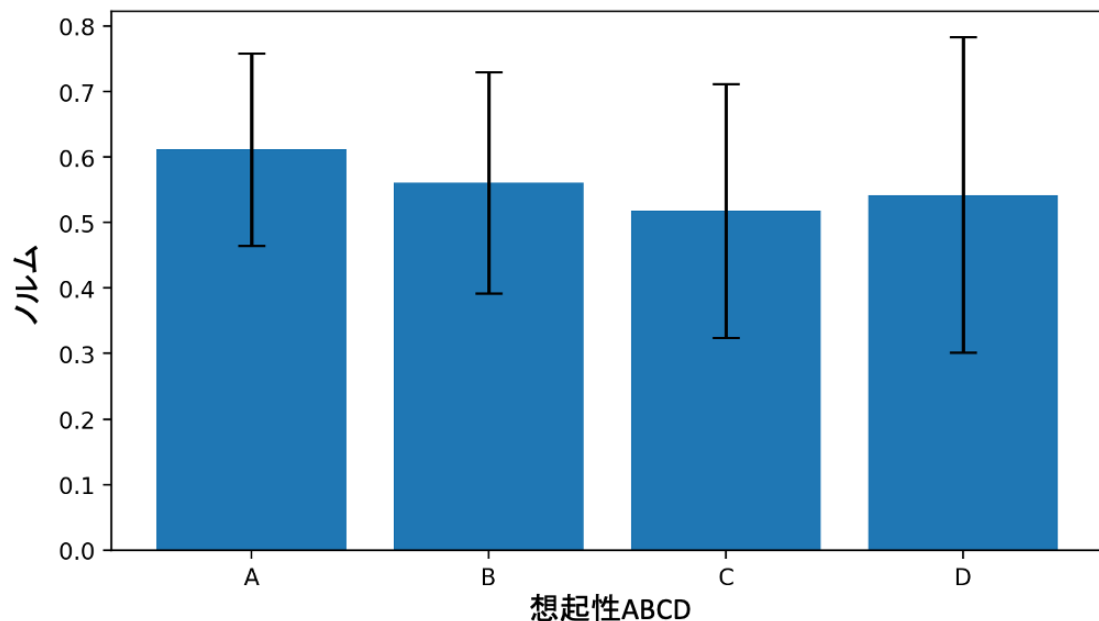


図 5.1 AOI 分布エントロピー (想起性 ABCD)

5.1.2 想起性観点における視線行動 (平均瞳孔径)

図 5.2 に、想起性条件 A～D における平均瞳孔径を示す。平均瞳孔径は認知負荷を反映する補助指標として扱った。

その結果、A 条件の平均瞳孔径が最も大きく、A 条件において、B 条件 ($p=0.0077$)、C 条件 ($p=0.0174$)、D 条件 ($p=0.0129$) の間で統計的有意差が確認された。これは、強調概念と文脈との整合性がともに高い画像において、学習者が強調内容と画像全体との関係をより深く処理していた可能性を示している。一方、その他の組み合わせでは有意差は確認されなかった ($p>0.05$)。

B, C, D 条件では大きな差は見られなかったが、全体として想起性の高い条件ほど、画像理解に伴う認知的処理が生じていたことが示唆される。

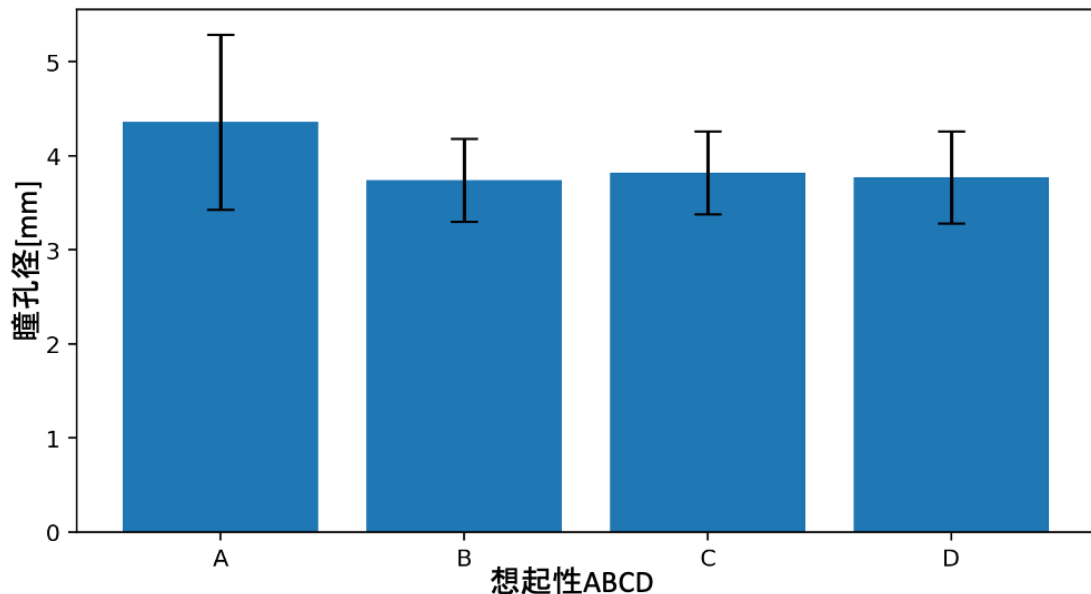


図 5.2 瞳孔径（想起性 ABCD）

5.1.3 想起性観点における記憶保持

想起性観点における記憶保持の評価として、直後想起テストおよび 1 週間後の遅延想起テストにおける強調単語の想起率、ならびに保持率を比較した。想起性条件 A～D における直後想起率を図 5.3 に、遅延想起率を図 5.4 に、保持率を表 5.2 に示す。

その結果、想起性の高い条件では、直後想起率および遅延想起率がともに高い傾向が見られた。特に、強調概念とコンテキストの双方が高い条件では、画像内の情報が強調語彙および文脈の理解に寄与しやすく、後続の記憶保持にも結び付きやすい結果となった。一方、想起性の低い条件では、強調語彙と画像内容との対応が不十分であるため、想起率および保持率が低くなる傾向が見られた。

以上より、強調生成画像が強調概念と文脈をどの程度適切に表現しているかが、記憶保持に関与している可能性がある。ただし、統計的に有意であるかについてはデータ数を増やした今後の検証が必要である。

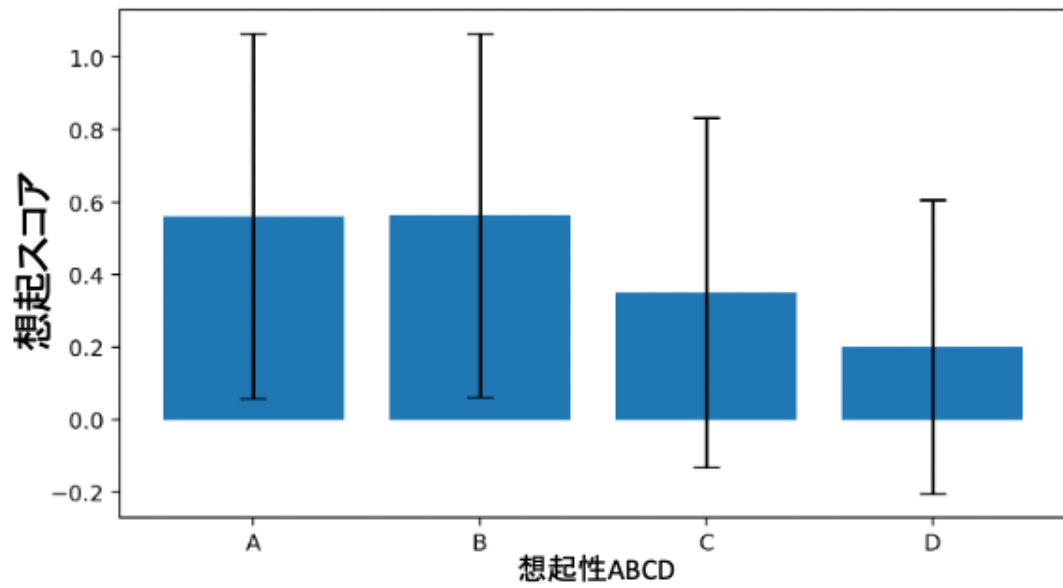


図 5.3 直後想起テスト (想起性 ABCD)

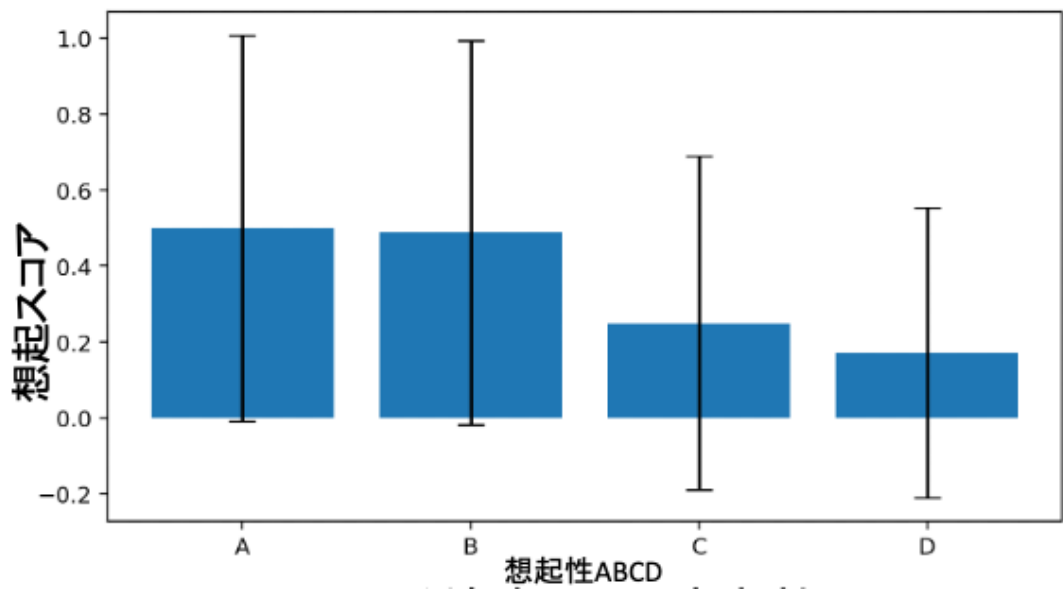


図 5.4 遅延想起テスト (想起性 ABCD)

表 5.2 想起テスト（想起性 ABCD）

想起性(ABCD)	直後想起率	遅延想起率	正解数/ 直後正解数	保持率
A	0.5625	0.5000	16/18	0.8889
B	0.5641	0.4872	19/22	0.8636
C	0.3500	0.2500	10/14	0.7143
D	0.2000	0.1714	6/7	0.8571

5.1.4 想起性観点における印象アンケート

想起性観点における主観評価として、画像提示後にアンケートを実施した。本アンケートでは、画像の見やすさについて、5段階で回答を求めた。評価尺度は以下の通りである。

- 1：見やすかった
- 2：少し見やすかった
- 3：普通
- 4：少し見づらかった
- 5：見づらかった

図 5.5 では、これらの回答値の平均値を条件別に示している。

その結果、想起性の高い条件では、画像が見やすい傾向が示された。特に、強調概念の一致度とコンテキストの一致度がともに高い条件では、主観評価が比較的良好であり、学習者が画像を意味的に処理しやすかった可能性がある。逆に、双方が低い条件では、画像の内容理解や強調語彙との対応が困難であるため、評価が低くなる傾向が見られた。ただし、統計的に有意であるかについてはデータ数を増やした今後の検証が必要である。

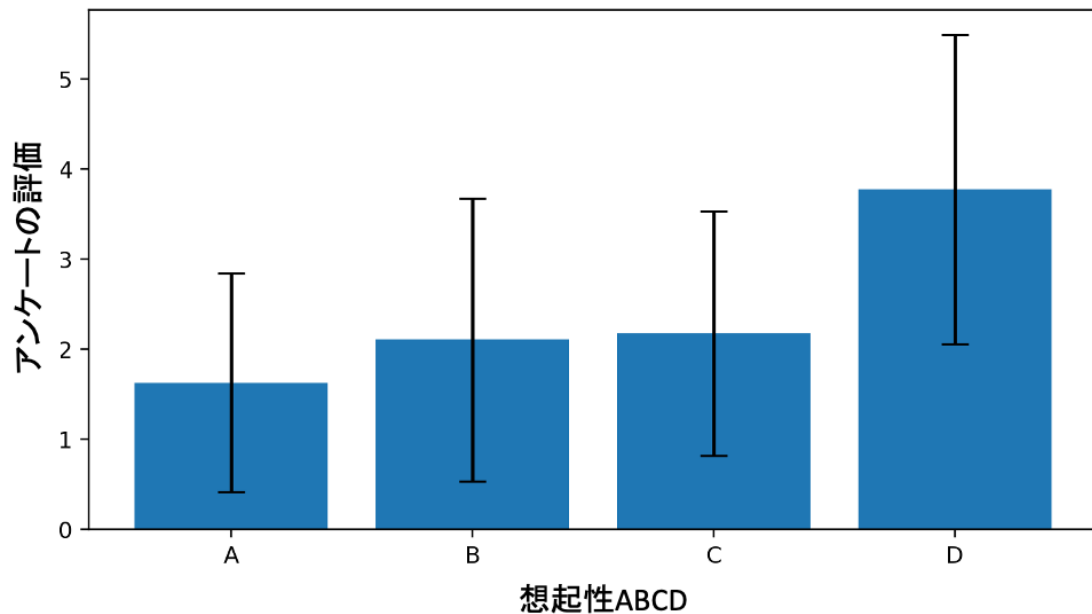


図 5.5 印象アンケート（想起性 ABCD）

5.2 印象性観点における結果

5.2.1 印象性観点における視線行動（AOI エントロピー）

印象性観点では、「強調概念表層比率」および「強調概念色差」の高低の組合せにより、a～d の 4 条件を設定した。各条件における AOI 分布エントロピーを比較することで、視線探索の分散度の違いを評価した。

その結果、c 条件において、a 条件 ($p=0.0041$)、b 条件 ($p=0.0023$) の間で統計的有意差が確認された。一方、その他の組み合わせでは有意差は確認されなかった ($p>0.05$)。

印象性が高い条件では、AOI 分布エントロピーが比較的高く、学習者が画像全体を見渡しながら強調の反映箇所を探索する傾向が見られた。特に、強調概念表層比率と色差がともに高い条件では、画像中での強調要素の存在感が大きく、視線が特定領域に留まるだけでなく、その周辺や背景との関係性を探索する行動が促されていた。

視線分布の偏りが大きく、画像内で強調要素を見つけにくい状態が反映されている可能性がある。

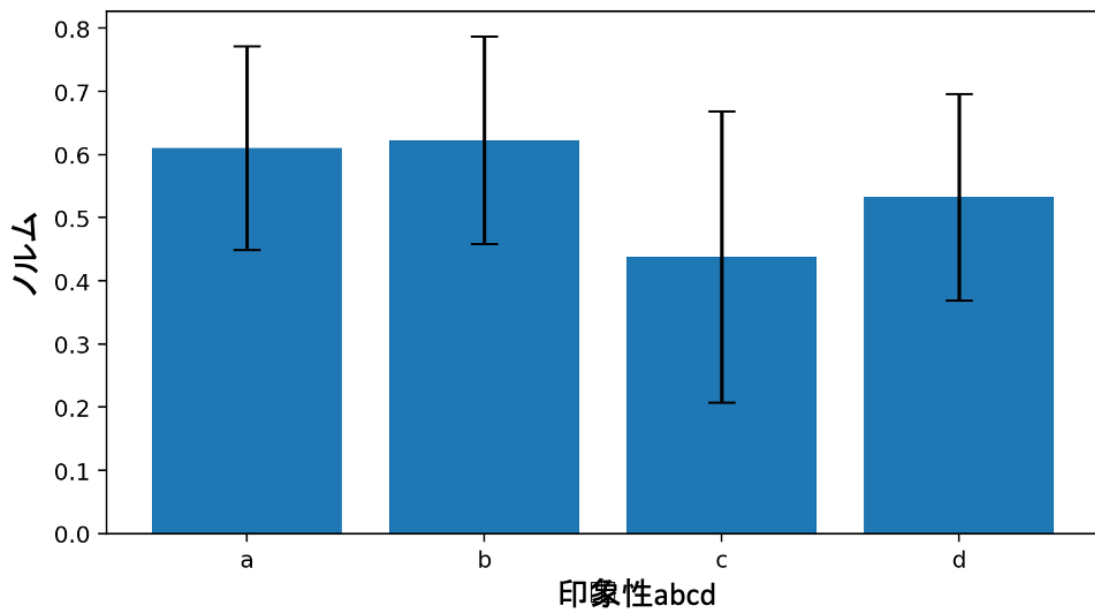


図 5.6 AOI 分布エントロピー (印象性 abcd)

5.2.2 印象性観点における視線行動 (平均瞳孔径)

印象性観点における平均瞳孔径を比較した結果、強調概念表層比率および色差が高い条件で、相対的に大きい値を示す傾向が見られた。図 5.7 に示す。b 条件において、a 条件において、b 条件 ($p=0.0137$)、d 条件 ($p=0.0191$)、c 条件において、b 条件 ($p=0.0191$)、d 条件 ($p=0.0191$) の間で統計的有意差が確認された。一方、その他の組み合わせでは有意差は確認されなかった ($p>0.05$)。

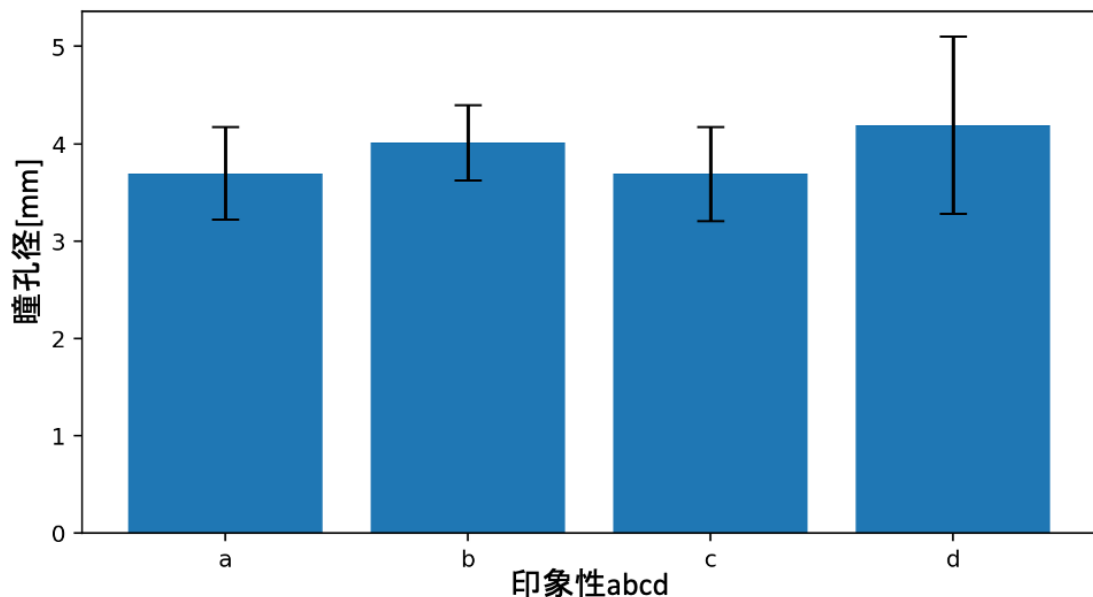


図 5.7 瞳孔径 (印象性 abcd)

5.2.3 印象性観点における記憶保持

印象性観点における記憶保持として、直後想起率、遅延想起率、および保持率を条件別に比較した。印象性条件 a~d における直後想起率を図 5.8 に、遅延想起率を図 5.9 に、保持率を表 5.3 に示す。その結果、印象性の高い条件では、直後想起率が高いだけでなく、1 週間後の保持率も比較的高い傾向が確認された。

特に、強調概念表層比率または色差のいずれかが高い条件では、視覚的に強調箇所が印象づけられやすく、その効果が記憶保持にも反映されていた。一方で、両パラメータが低い条件では、画像中の強調要素が目立ちにくく、記憶に残りにくい結果となった。保持率は全条件で比較的高く、特に d 条件では直後正解者の多くが 1 週間後も正解している。ただし、d 条件では直後想起率が低いため、画像から単語を読み取ることは難しいと考えられる。

このことから、単語を視覚的に強調する設計は、短期的な気づきだけでなく、遅延後の保持にも関与する可能性が示された。ただし、統計的に有意であるかについてはデータ数を増やした今後の検証が必要である。

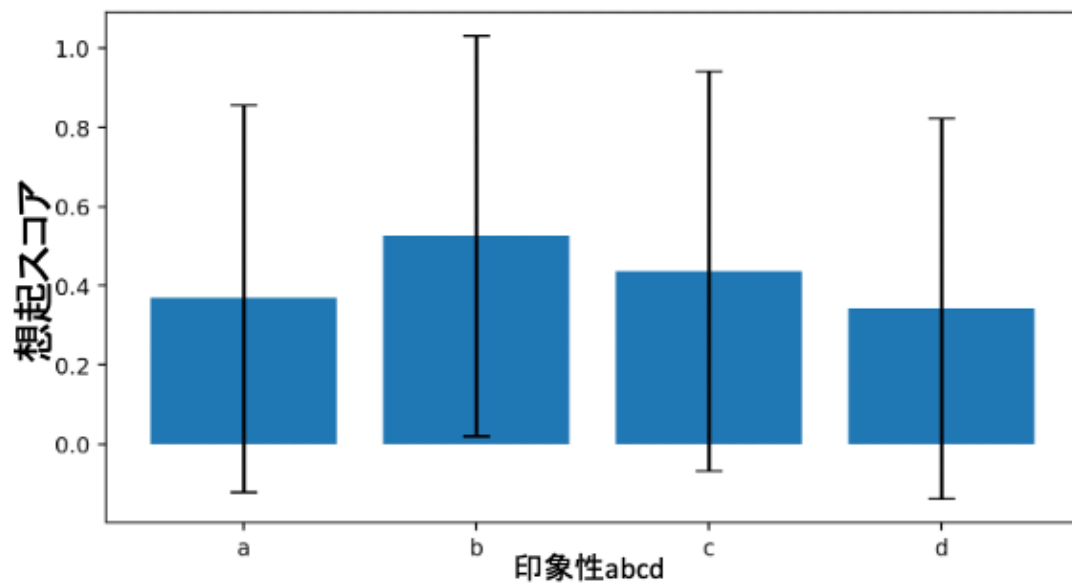


図 5.8 直後想起テスト (印象性 abcd)

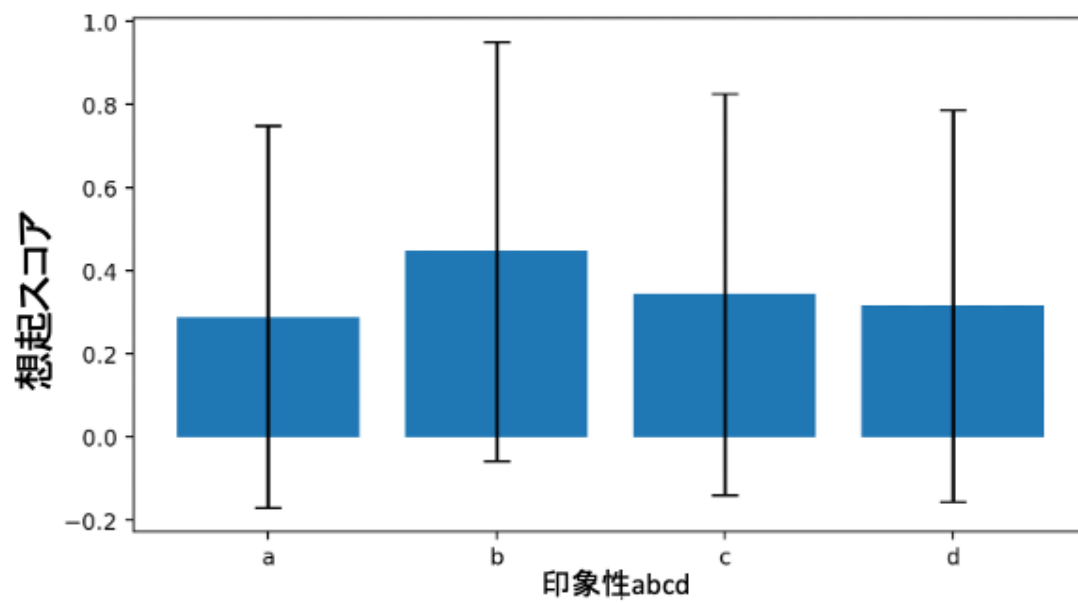


図 5.9 遅延想起テスト (印象性 abcd)

表 5.3 想起テスト（印象性 abcd）

印象性(abcd)	直後想起率	遅延想起率	正解数/ 直後正解数	保持率
a	0.368	0.289	11/14	0.786
b	0.526	0.447	17/20	0.850
c	0.438	0.344	11/14	0.786
d	0.342	0.316	12/13	0.923

5.2.4 印象性観点における印象アンケート

印象性観点における主観評価として、画像提示後にアンケートを実施した。本アンケートでは、強調箇所への気づきやすさについて、5段階で回答を求めた。評価尺度は以下の通りである。

- 1：すぐに気づくことができた
- 2：少し時間はかかったが気づくことができた
- 3：時間はかかったが気づくことができた
- 4：気付くのに時間いっぱい使ってしまった
- 5：気づかなかった

図 5.10 では、これらの回答値の平均値を条件別に示している。

その結果、画像の見やすさについては条件間の差が比較的小さい一方で、「誤っている単語が画像中のどこに反映されているか気づくことができたか」という項目では、印象性パラメータが低い条件ほど評価が低く、強調要素に気づきにくい傾向が明確に表れた。逆に、強調概念表層比率や色差が高い条件では、学習者が強調要素の存在に早く気づきやすい結果となった。

この結果は、印象性パラメータが視覚的な気づきやすさに強く関与していることを、主観評価の側面から支持するものである。ただし、統計的に有意であるかについてはデータ数を増やした今後の検証が必要である。

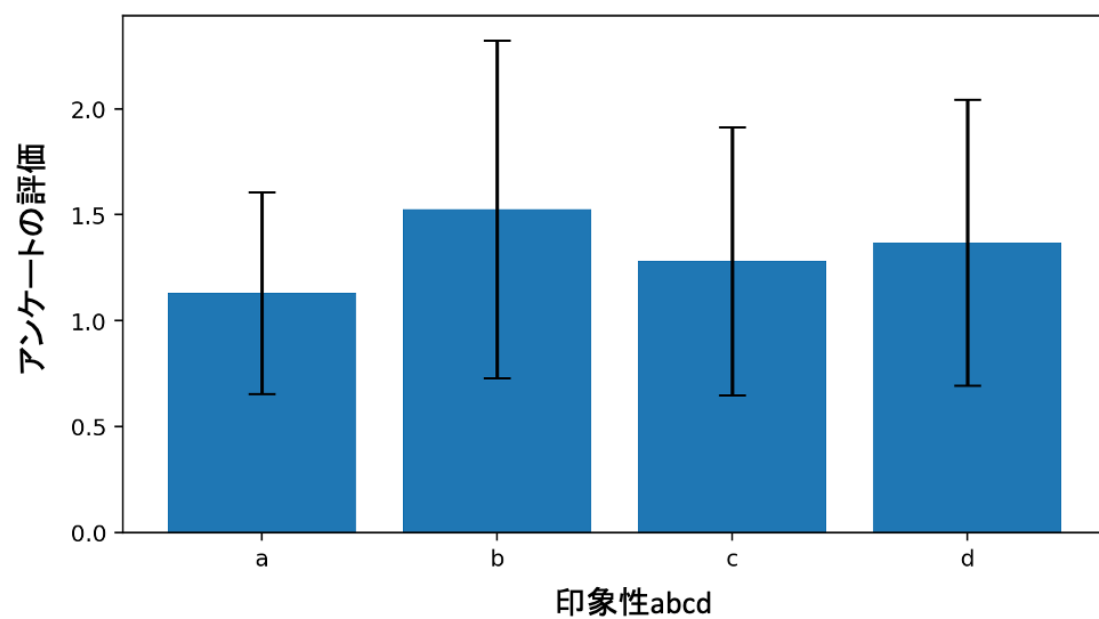


図 5.10 印象アンケート (印象性 abcd)

第6章 考察

6.1 想起性観点における結果

6.1.1 想起性観点における視線行動（AOI 分布エントロピー）

想起性観点における AOI 分布エントロピーの結果から、強調概念の一致度およびコンテキストの一致度が高い条件では、学習者の視線が画像内の複数領域に分散する傾向が見られた。これは、強調生成画像が強調語彙と文脈の双方を適切に表現している場合、学習者が局所的な視覚要素にとどまらず、画像全体の関係性を探索しながら意味づけを行っている可能性を示している。

本研究の目的は、強調箇所そのものをどれだけ見たかを評価することではなく、強調を含む画像全体の文脈を学習者がどのように探索するかを明らかにすることにある。この観点からみると、想起性の高い条件で AOI エントロピーが高くなる傾向は、強調単語を想起するために、学習者が周辺要素や背景情報も含めた文脈探索を行っていたことを示唆する。

一方で、想起性の低い条件では、強調語彙や文脈との対応が弱いため、学習者が画像全体を意味的に探索する手がかりが不足し、視線行動が限定的になった可能性が考えられる。したがって、強調生成画像においては、強調概念と文脈がある程度明確に表現することが、学習者の探索的視線行動を引き出す上で重要であると考えられる。

6.1.2 想起性観点における視線行動（平均瞳孔径）

想起性観点における平均瞳孔径の結果は、強調概念や文脈との整合性が高い条件で比較的大きな値を示す傾向を示した。瞳孔径は認知的負荷や処理努力を反映する補助的指標として位置づけられるため、この結果は、想起性の高い条件では、学習者が画像を単に眺めるのではなく、強調語彙と画像内容の関係を意味的に照合しながら処理していた可能性を示している。

ただし、瞳孔径は照明や瞬目、欠損値の影響を受けやすく、また処理負荷の増大が必ずしも学習効果の向上に直結するわけではない。そのため、瞳孔径の結果は、AOI エントロピーや記憶保持の結果と併せて慎重に解釈する必要がある。それでも、想起性の高い画像条件において、視線分布が広がると同時に瞳孔径も増加する傾向が見られたことは、画像全体を探索しつつ意味づけを行う認知的処理が生じていた可能性を支持する。

6.1.3 想起性観点における記憶保持

想起性観点では、強調概念の一致度とコンテキストの一致度の双方が高い条件において、直後想起率、遅延想起率、および保持率が高い傾向が見られた。これは、強調語彙と画像との対応が明確であり、かつ強調が置かれた文脈が画像内で適切に表現されている場合、学習者が画像から得られる情報を記憶検索の手がかりとして利用しやすくなるためと考えられる。

この結果は、第2章で述べた二重符号化理論の観点とも整合する。すなわち、視覚情報と言語情報が相互に補完し合うことで、学習内容がより強固に記憶に保持される可能性がある。本研究では、強調生成画像が強調概念と文脈の双方を含んでいる場合に、学習者が視覚情報から強調語彙を再構成しやすくなり、そのことが直後だけでなく遅延後の保持にも寄与した可能性が示された。

一方で、想起性の低い条件では、画像から得られる意味の手がかりが不十分であるため、強調内容の記憶保持も弱くなる傾向が見られた。したがって、強調生成画像の学習効果を高めるためには、強調概念そのものの表現だけでなく、それが置かれた文脈をどの程度忠実に再現できているかが重要であると考えられる。

6.1.4 想起性観点における印象アンケート

想起性観点における事後アンケートでは、強調概念の一致度およびコンテキストの一致度が高い条件ほど、画像の意味内容が理解しやすく、強調語彙との対応にも気づきやすい傾向が見られた。これは、想起性の高い画像が、学習者にとって「見て理解できる画像」として機能していた可能性を示している。

この結果は、記憶保持の結果とも整合的である。すなわち、学習者が主観的に「理解しやすい」と感じた画像ほど、後の想起や保持にも結びつきやすかったと解釈できる。したがって、想起性は単なる画像評価指標にとどまらず、学習者の主観的理解と客観的な記憶成績の双方に関与する重要な要因であると考えられる。

6.2 印象性観点における考察

6.2.1 印象性観点における視線行動（AOI 分布エントロピー）

印象性観点における AOI 分布エントロピーの結果では、強調概念表層比率および強調概念色差が高い条件において、視線が画像内の複数領域に分散する傾向が見られた。これは、視覚的に目立つ強調要素が存在することで、学習者がそ

の要素を起点に画像全体を探索し、周辺との関係を確認していた可能性を示している。

一般に、印象性の高い画像は、強調要素への気づきを促進する一方で、過度な強調はかえって視線を局所に固定させる可能性も考えられる。しかし、本研究の結果では、印象性が高い条件であっても単一 AOI への固定ではなく、画像全体への探索が見られたことから、適度な視覚的強調が文脈探索を促す方向に働いた可能性がある。

6.2.2 印象性観点における視線行動（平均瞳孔径）

印象性観点における平均瞳孔径の結果では、強調概念表層比率や色差が高い条件で大きい値を示す傾向が見られた。このことは、視覚的に強調された要素を検出し、それがどのように強調語彙と関係するかを判断する際に、追加的な認知的処理が必要であった可能性を示している。

一方で、印象性が過度に高い条件では、視覚的な目立ちや違和感そのものが処理負荷を増加させた可能性もある。そのため、瞳孔径の増大は必ずしも望ましい学習状態だけを示すものではなく、「気づきやすさの向上」と「認知的負荷の増加」の両面を持つ可能性がある。したがって、印象性を高める設計は有効である一方で、その程度を適切に制御する必要があると考えられる。

6.2.3 印象性観点における記憶保持

印象性観点における記憶保持の結果からは、強調概念表層比率や色差が高い条件において、直後想起率や遅延保持率が比較的高い傾向が見られた。これは、視覚的に強調された強調要素が、学習者に強い印象を与え、そのことが記憶の定着に寄与した可能性を示している。

特に、印象性の高い条件では、学習者が「どこが強調がされているか」に気づきやすくなり、その気づきが強調語彙の再認や想起の手がかりとして機能したと考えられる。一方で、印象性の低い条件では、強調要素が画像内で目立たず、視覚的な手がかりとして十分に機能しなかったため、記憶保持も相対的に低くなった可能性がある。

ただし、印象性のみが高く、想起性が低い画像では、学習者が強い印象を受けても、その内容を意味的に十分理解できない場合があり得る。そのため、印象性による記憶保持の効果は、想起性との相互作用を考慮して解釈すべきである。

6.2.4 印象性観点における印象アンケート

印象性観点における主観評価では、画像の見やすさには大きな差が見られなかった一方で、「誤っている単語が画像のどこに反映されているか気づくことができたか」という項目では、印象性パラメータが高い条件ほど良好な評価が得られた。これは、印象性が学習者の主観的な「気づきやすさ」に直接的に作用していることを示す結果である。

この結果は、印象性という概念が、理論上の設計指標としてだけでなく、実際の学習者の主観的経験にも対応していることを支持する。すなわち、強調生成画像における視覚的強調は、学習者に強調を気づかせるうえで有効に機能していたと考えられる。

6.3 想起性と印象性の役割の違い

本研究の結果から、強調生成画像における誘目性の設計は、強調箇所への単純な注目を促すだけでなく、学習者の視線探索行動そのものを変化させる可能性が示された。すなわち、強調生成画像は、「強調を一点で示す教材」ではなく、「強調を含む文脈を探索させる教材」として機能し得る。

この観点から見ると、想起性や印象性といった設計指標は、強調箇所への直接的な注視量よりも、視線分布の広がりや探索の様式を通じて学習効果に影響を与えていると考えられる。本研究は、こうした設計指標と学習者の探索行動を結びつける定量的根拠を提供する点で意義を持つ。

6.4 本研究の意義

本研究の意義は、強調生成画像の学習効果を、視線行動、記憶保持、主観評価という複数の観点から統合的に検証した点にある。特に、従来は設計指標として提案されるにとどまっていた想起性および印象性について、学習者の実際の視線探索行動や記憶成績との対応関係を示したことは、強調生成画像の設計妥当性を検証する上で重要な知見である。

また、本研究では、強調箇所そのものへの直接的注視ではなく、画像全体の文脈探索に着目した AOI ベースの分析を行った。これにより、強調理解が単なる局所的注視ではなく、画像全体の関係情報の探索によって支えられている可能性を示した点にも意義がある。

6.5 今後の課題

本章では、想起性および印象性の観点から、強調生成画像が学習者の視線行動、記憶保持、および主観評価に与える影響について考察した。想起性の高い画像は、画像全体の探索的視線行動と高い記憶保持に結びつきやすく、印象性の高い画像は、強調要素への気づきやすさと保持率の向上に関与する可能性が示された。これらの結果は、強調生成画像の設計において、意味的理解を支える想起性と、視覚的気づきを支える印象性の双方を考慮することの重要性を示している。

第7章 おわりに

本研究では、英語学習における強調理解を支援するための強調生成画像に着目し、その学習効果を視線行動および記憶保持の観点から検討した。特に、強調生成画像の誘目性を「想起性」と「印象性」の二つの観点から整理し、それぞれの設計パラメータが学習者の視線探索行動、記憶保持、および主観評価にどのような影響を与えるかを実験的に検証した。

実験では、強調生成画像に対して想起性および印象性のパラメータを設定した複数条件を作成し、学習者の視線データをアイトラッカーによって取得した。視線行動の分析には AOI 分布エントロピーおよび平均瞳孔径を用い、さらに直後テストおよび 1 週間後の遅延テストによって記憶保持を評価した。また、画像提示後には印象アンケートを実施し、学習者の主観的評価についても併せて分析した。

結果として、想起性の高い条件では、学習者の視線が画像内の複数領域に分散する傾向が見られ、画像全体を探索しながら強調語彙と文脈との関係を理解しようとする行動が観察された。また、これらの条件では直後想起率および遅延想起率が比較的高く、強調概念と文脈が適切に表現された画像が記憶保持の向上に寄与する可能性が示された。一方、印象性の高い条件では、強調要素への気づきやすさが向上し、視覚的な印象づけを通じて記憶保持にも一定の効果が見られた。さらに、印象アンケートの結果からも、想起性の高い画像は理解しやすいと評価され、印象性の高い画像は強調要素に気づきやすいと評価される傾向が確認された。

これらの結果から、強調生成画像の学習効果には、強調概念と文脈との意味的対応を支える想起性と、強調要素への視覚的気づきを促す印象性の双方が関与している可能性が示された。すなわち、強調生成画像の設計においては、単に強調内容を視覚的に強調するだけでなく、強調概念と文脈の関係を適切に表現することが重要であると考えられる。本研究は、これらの誘目性パラメータを実験的に検証し、視線行動および記憶保持との関係を示した点において、強調生成画像の設計指針を検討するための基礎的知見を提供するものである。

一方で、本研究にはいくつかの課題も残されている。第一に、被験者数が限定的であるため、本研究で得られた結果は探索的な傾向として解釈する必要がある。今後は被験者数を増やし、統計的検定を含めたより厳密な検証を行うことが

望まれる。第二に、本研究では AOI グリッドによる空間的な近似を用いて視線行動を分析したが、AOI は必ずしも意味的領域と一致するものではない。今後は意味単位に基づく領域分割や画像解析技術を併用することで、より精緻な視線分析が可能になると考えられる。さらに、本研究では想起性と印象性を独立に整理して検討したが、実際には両者が相互に影響し合う可能性がある。今後は両パラメータの組合せ効果を体系的に検証し、より効果的な強調生成画像の設計指針を明らかにすることが重要である。

以上より、本研究は強調生成画像の誘目性パラメータと学習者の視線行動および記憶保持との関係を明らかにすることで、強調理解を支援する視覚教材設計の基礎的知見を示した。今後は、本研究で得られた知見を基に、より多様な学習環境や学習者特性を考慮した教材設計の検討を進めることで、強調を活用した学習支援のさらなる発展につながることが期待される。

謝辞

本研究を進めるにあたり，多くの方々から多大なるご指導とご支援を賜りました．ここに深く感謝の意を表します．

まず，本研究全体を通して終始懇切丁寧なご指導を賜りました長谷川先生に，心より御礼申し上げます．研究の方向性や論文構成に関する的確なご助言は，本研究をまとめ上げる上で大きな支えとなりました．

また，本研究で用いた誤答強調生成画像データセットおよび関連する知見をご提供くださった杉田さんには，心より感謝いたします．学習システムやデータセットや生成モデルの特性に関するご説明は，本研究の実験設計および解釈において不可欠なものでした．

さらに，実験の実施にあたりご協力いただいた皆様，ならびに被験者として実験に参加してくださった方々に厚く御礼申し上げます．皆様のご協力なくして，本研究を遂行することはできませんでした．

最後に，本研究を支えてくださった長谷川研究室の皆様に感謝いたします．日常的な議論や助言を通じて，多くの刺激を受けることができました．

ここに改めて，本研究にかかわったすべての方々に，深く感謝の意を表します．

参考文献

- [1] Clark, J. M., & Paivio, A. (1991). "Dual Coding Theory and Education", *Educational Psychology Review*, 3, 149–210,
<https://nschwartz.yourweb.csuchico.edu/Clark%20&%20Paivio.pdf>
- [2] Plass, J. L., Chun, D. M., Mayer, R. E., & Leutner, D. (1998). "Supporting Visual and Verbal Learning Preferences in a Second-Language Multimedia Learning Environment ",*Journal of Educational Psychology*, 90(1), 25-36.,https://www.academia.edu/15362842/Supporting_visual_and_verbal_learning_preferences_in_a_second_language_multimedia_learning_environment
- [3] Hasnine, M. N., et al. (2019). Vocabulary learning support system using automatic image captioning. In D. Leal Filho (Ed.), *Learning and collaboration technologies* (pp. 344–356). Springer
- [4] Nassaji, H. (2003). "L2 Vocabulary Learning From Context: Strategies, Knowledge Sources, and Their Relationship With Success in L2 Lexical Inferencing", *TESOL Quarterly*, 37(4), 645–670.,
<https://doi.org/10.2307/3588216>
- [5] Sugita, K., et al. (2023). Error-driven learning: Vocabulary learning support using error by image generation. *IIAI Letters*, 3, 1.
<https://doi.org/10.52731/iil.v003.078>
- [6] 杉田一樹, et al. (2023). 語彙学習支援システム：画像の示唆性に対する指標. *教育システム情報学会論文誌*, 38(2), 118–124
- [7] 杉田一樹, et al. (2024). 語彙学習支援システム：誤答生成画像に対する定量評価. 第 49 回教育システム情報学会全国大会, 147–148
- [8] Horiguchi, T., & Hirashima, T. (2002). Simulation-based learning for error-correction. *Transactions of the Japanese Society for Artificial Intelligence*, 17, 462–472
- [9] Kenneth Holmqvist, Marcus Nyström, Richard Andersson, Richard Dewhurst, Halszka Jarodzka, Joost van de Weijer (2011), "Eye Tracking: A Comprehensive Guide to Methods and Measures", Oxford University Press. <https://research.ou.nl/en/publications/eye-tracking-a-comprehensive-guide-to-methods-and-measures/>
- [10] Andrew T. Duchowski (2007), "Eye Tracking Methodology: Theory and Practice (2nd ed.)", Springer. <https://link.springer.com/book/10.1007/978-1-84628-609-4>
- [11] Keith Rayner (1998), "Eye movements in reading and information

- processing: 20 years of research", *Psychological Bulletin*, 124(3), 372–422. <https://pubmed.ncbi.nlm.nih.gov/9849112/>
- [12] Marcel A. Just, Patricia A. Carpenter (1980), "A Theory of Reading: From Eye Fixations to Comprehension", *Psychological Review*, 87(4), 329–354. (doi:10.1037/0033-295X.87.4.329)
- [13] Deborah E. Hannula, Charan Ranganath (2009), "The eyes have it: Hippocampal activity predicts expression of memory in eye movements", *Neuron*, 63(5), 592–599. <https://pubmed.ncbi.nlm.nih.gov/19755103/>
- [14] Bernhard Fehlmann, David Coynel, Nathalie Schicktz, Annette Milnik, Leo Gschwind, Pascal Hofmann, Andreas Papassotiropoulos, Dominique J-F de Quervain (2020), "Visual Exploration at Higher Fixation Frequency Increases Subsequent Memory Recall", *Nature Communications*, 11, 4489. <https://pmc.ncbi.nlm.nih.gov/articles/PMC8153053/>
- [15] Sweller, J. (1988). "Cognitive load during problem solving: Effects on learning", *Cognitive Science*, 12(2), 257–285. https://doi.org/10.1207/s15516709cog1202_4
- [16] Plass, J. L., Moreno, R., & Brünken, R. (2010). "Cognitive Load Theory", Cambridge University Press. <https://doi.org/10.1017/CBO9780511844744>
- [17] <https://www.tobii.com/ja/products/eye-trackers/screen-based/tobii-pro-fusion>(2026 年 1 月 11 日 アクセス)
- [18] Tatler, B. W. (2007), The central fixation bias in scene viewing, *Journal of Vision*, 7(14), 1–17, <https://doi.org/10.1167/7.14.4>
- [19] van der Wel, P., & van Steenbergen, H. (2018). Pupil dilation as an index of effort in cognitive control tasks: A review. *Psychonomic Bulletin & Review*, 25(6), 2005–2015. <https://doi.org/10.3758/s13423-018-1432-y>