

推薦論文

# オノマトペ CAPTCHA の開発と評価

滋野 莉子<sup>1</sup> 山田 道洋<sup>2,a)</sup> 菊池 浩明<sup>1</sup> 坂本 真樹<sup>3</sup>

受付日 2017年11月27日, 採録日 2018年6月8日

**概要:** CAPTCHA は、コンピュータと人間を区別するための完全自動化された公開チューリングテストであり、コンピュータエージェントによる不正行為を防止する重要な役割を果たしている。しかし、コンピュータの AI 技術が発展した現在、既存の CAPTCHA のいくつかが破られていることが大きな課題となっており、機械判別が不可能な新しい CAPTCHA の生成が求められている。そこで、本研究では、擬態語や擬音語であるオノマトペを応用した「オノマトペ CAPTCHA」と5つの問題形式を提案する。計3回外国人63名を含む合計377名の評価実験を通じて、オノマトペ CAPTCHA の精度と課題を明らかにし、ユーザビリティの向上を研究目的とする。既存の代表的な CAPTCHA であるパズル型、画像認証型と性能を比較する。

**キーワード:** CAPTCHA, 認証, オノマトペ, 合成語, 辞書攻撃

## Evaluation and Development of Onomatopoeia CAPTCHA

RIKO SHIGENO<sup>1</sup> MICHIMIRO YAMADA<sup>2,a)</sup> HIROAKI KIKUCHI<sup>1</sup> MAKI SAKAMOTO<sup>3</sup>

Received: November 27, 2017, Accepted: June 8, 2018

**Abstract:** CAPTCHA (Completely Automated Public Turing test to tell Computers and Humans Apart) is a type of challenge-response test used in determining whether or not when a given access is made by human. However, as AI technologies are improved, most existing CAPTCHAs have been broken. Hence, new CAPTCHA that is secure against machine is required. In this paper, we propose five new schemes, called “Onomatopoeia CAPTCHA” based on hardness to interpret unknown onomatopoeias. We conduct some experiments involved with total of 377 human subjects including 63 foreigners to show that our proposed CAPTCHAs improve usability and the accuracy. We also show a performance comparison of our scheme with some representative CAPTCHA schemes, puzzle and image-recognition styles.

**Keywords:** CAPTCHA, authentication, onomatopoeia, synthesized word, dictionary attack

### 1. はじめに

CAPTCHA (Completely Automated Public Turing Test To Tell Computers and Humans Apart) は、コンピュー

タと人間を区別するための完全自動化された公開チューリングテストである。コンピュータエージェントによる不正が深刻な課題となっているなか、セキュリティ上で重要な役割を果たしている。CAPTCHA の必要条件は、

- (1) 人間が容易に解ける、
- (2) コンピュータには解くことができない、
- (3) 問題を自動生成できる、

ことである。しかし、コンピュータによる深層学習が発展した現在、既存の CAPTCHA の多くが条件 (2) を満

<sup>1</sup> 明治大学総合数理学部  
School of Interdisciplinary Mathematical Science, Meiji University, Nakano, Tokyo 164-8525, Japan

<sup>2</sup> 明治大学大学院先端数理科学研究科  
Graduate School of Advanced Mathematical Sciences, Meiji University, Nakano, Tokyo 164-8525, Japan

<sup>3</sup> 電気通信大学大学院情報理工学研究科  
Graduate School of Informatics and Engineering, The University of Electro-Communications, Chofu, Tokyo 182-8585, Japan

a) cs172001@meiji.ac.jp

本論文の内容は 2017 年 6 月のマルチメディア、分散、協調とモバイル (DICOMO2017) シンポジウムにて報告され、セキュリティ心理学とトラスト研究会主査より情報処理学会論文誌ジャーナルへの掲載が推薦された論文である。

たさなくなっているという問題点がある。Bursztein らは 2011 年に当時の商用サイトで利用されていた視覚型と聴覚型の大部分の CAPTCHA に対して、機械学習を用いたロボットによる攻撃に成功した [1], [2]。また、最近の研究 [3], [4], [5] では、Bursztein らの報告では比較的頑強であった GoogleCAPTCHA に対しても、高い確率で攻撃に成功したとの報告がされている。そこで、機械判別が困難な新しい CAPTCHA の生成が求められている。藤田らは、3DCG 画像を用いた非現実画像 CAPTCHA [6] を提案している。山口らは機械合成文の不自然度相対識別問題に基づく CAPTCHA [7] を提案している。

本研究では、擬態語や擬音語であるオノマトペ (Onomatopoeias) を応用した新たな「オノマトペ CAPTCHA」を提案する。オノマトペは普段人間が無意識に使用している概念であり、音や響きから意味を推測できるので、人間にとって解くのが簡単である。一方で、その解釈の原理はよく知られておらず、機械的に解くのは困難と予測される [8]。

本稿では、提案方式が CAPTCHA の必要条件を満たすか明らかにするために、出題方式を 5 形式提案する。また既存のオノマトペと合成オノマトペを使用した問題の人間による正しく受理される確率などを明らかにする。幅広い被験者で評価するため、クラウドソーシングサイトを用いて被験者を集めて実験を行った。これにより、人間にとって容易なオノマトペ CAPTCHA の形式を明らかにするとともに、既存の CAPTCHA との性能比較も行う。またオノマトペ CAPTCHA が機械を誤って受け入れる確率を算出し、提案方式の安全性を明らかにする。さらに、言語的な背景知識への依存が小さいマンガと組み合わせたマンガオノマトペ CAPTCHA を提案し、外国人の被験者を募集し実験を行い、日本人との正しく受理される確率に差が生じるかを検討する。

本稿の構成は次のとおりである。2 章で、CAPTCHA とオノマトペについての基本概念を定義し、3 章では提案方式について、4 章では問題 4 形式を用いた実験、5 章ではマンガオノマトペ CAPTCHA について記述する。

## 2. 基本定義

### 2.1 CAPTCHA について

Capy による「パズル型 CAPTCHA (図 1) [11]」や google による「画像認証型 CAPTCHA (図 2) [12]」など様々な CAPTCHA が使用されている。パズル型 CAPTCHA は、図 1 のようにパズルの 1 ピースを絵や写真に適切にはめこむテストである。画像認証型 CAPTCHA は、与えられた記述に該当する写真をすべて選択するテストである。図 2 では、「ワッフルの画像をすべて選択してください」と出題されており、9 枚の写真の中から、ワッフルの写真をすべて (1 行 3 列, 2 行 2 列, 2 行 3 列) の 3 枚を選べば正解である。

true success

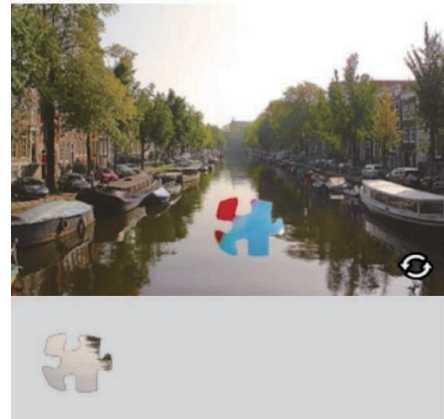


図 1 パズル型 CAPTCHA [11]

Fig. 1 Puzzle-style CAPTCHA.



図 2 画像認証 CAPTCHA [12]

Fig. 2 Re-CAPTCHA.

### 2.2 CAPTCHA の評価

本研究では、CAPTCHA が人間を受け入れる確率を人間受入率  $HAR$  (Human Acceptance Rate) と定義する。一方、CAPTCHA が機械を受け入れる確率を機械受入率  $MAR$  (Machine Acceptance Rate) と定義する。 $HAR$  が高いことは、(1) 人に容易に解けることを、 $MAR$  が低いことは、(2) 機械には解けないことを表している。 $HAR$  が高く、 $MAR$  が低いとき精度が高い CAPTCHA といえる。 $HAR$  は本実験の正解率と、実験結果から算出した理論値と定義する。 $MAR$  は攻撃者の持つ能力に依存する。本研究ではブルートフォース攻撃、google 翻訳攻撃、google 検索攻撃と辞書攻撃に対する安全性を評価する。

#### 2.2.1 ブルートフォース攻撃

各選択肢を一樣な確率でランダムに選択する攻撃者をブルートフォース攻撃と呼ぶ。その攻撃者の受入率を  $MAR_b$  と定める。たとえば、選択肢が 4 つあるときは、 $1/4$  となる。

#### 2.2.2 google 翻訳攻撃

$MAR_g$  は、google 翻訳でオノマトペの意味が正しく解釈されるときに、CAPTCHA が機械を誤って受け入れる確率とする。ただし、複数の選択肢の一部のみ翻訳されるこ

とも考慮して厳密に算出する必要がある。詳細は、4.3.2 項で述べる。

### 2.2.3 google 検索攻撃

背景知識として、検索エンジンにアクセスできる攻撃者を与える。この攻撃では、形式 1 のような選択問題に対して、意味の単語と選択肢の各々について AND 検索を行い、検索ヒット数が最も多くなる選択肢を選ぶ。詳細は、4.3.2 項で述べる。

### 2.2.4 辞書攻撃

攻撃者がオノマトベ辞書を入手した場合の受入率を  $MAR_d$  とする。辞書のオノマトベを選択肢に使用している場合、攻撃者はオノマトベの意味を知ることができるため正解することができる。

## 2.3 オノマトベについて

オノマトベには、ものの音や声をまねた擬態語（ざわざわ、じょきじょきなど）、状態をまねた擬音語（てきぱき、きらきらなど）、擬声語（わんわん、おぎゃー）がある。一般語彙と比べると臨場感に溢れ、繊細な表現を可能としているという特徴がある。また、日本語は他の言語と比べてオノマトベの種類が圧倒的に多いといわれている [8]。小松らは、オノマトベの持つイメージを数値化し、オノマトベの持つイメージを客観的に表現するようなシステムを生成する研究をしている [9]。清水らは、ユーザの入力した印象評価値に適合した音韻と形態を持つオノマトベ表現から、ユーザの感性的印象に適合したオノマトベを生成するシステムを開発している [8]。このようにオノマトベには様々な先行研究があるが、いまだにオノマトベの明確な規則性は明らかにはされていない。

## 3. 提案方式

オノマトベ CAPTCHA はオノマトベの意味を正しく判断する問題を使用した CAPTCHA である。普段人間は正確な語でなくても、音のリズムや響きからそのオノマトベの意味を正しく理解することができる。しかしオノマトベには明確な規則性がないために、機械に意味を解釈するのは困難である。ゆえにオノマトベは人間に容易（条件 1）で、機械には困難（条件 2）という点で CAPTCHA に適している。

### 3.1 オノマトベ辞書

オノマトベ CAPTCHA の問題生成するにあたって、まずオノマトベ辞書を作成した。オノマトベ辞書には、日本語オノマトベ辞典 [10] の意味分類の中から 20 項目を選出し、全 977 語を登録した。表 4 にその一部を示す。オノマトベ辞書の意味に対するオノマトベ数は一定ではない。その統計量を表 3 に示す。

本研究では、人間や機械について、オノマトベ CAPTCHA

表 1 オノマトベの解釈難易度（0 が易、2 が難）の調査結果  
Table 1 Survey results on interpretation difficulty of onomatopoeia.

難易度	意味	人間 1 評価		人間 2 評価		Google 翻訳	
		語数	%	語数	%	語数	%
0	既知	359	37	228	23	262	27
1	未知	280	29	371	38	168	17
2	誤解	338	34	378	39	547	56
	計	977	100	977	100	977	100

表 2 オノマトベ辞書

Table 2 Onomatopoeia dictionary.

タイプ	人間の予測	Google 翻訳	語数
A	0	1,2	210
B	1,2	0	112
C	1	1	59
その他 1	0	0	150
その他 2	2	2	190
その他 3	1	2	195
その他 4	2	1	61

表 3 実験に用いた辞書の意味に対するオノマトベ数

Table 3 Number of onomatopoeia for a meaning of the dictionary used in the experiment.

	語数
平均	48.85
最大	74
最小	11
総数	977

の解釈の正しさや解答時間にどれだけ差が生じるかを調べるために、全オノマトベの解釈難易度を人間予測と google 翻訳で評価した。人間予測は、人間にとって既知、未知、誤訳の 3 段階で評価した。google 翻訳は、正しく翻訳される語を 0（既知）、ローマ字で表示される語を未知 1、誤って翻訳される語を 2（誤解）とした。その結果を表 1 に示す。また人間 1, 2 は著者ら（滋野, 山田）が評価した。google で正しく翻訳できるかどうかという観点（1 と 2 の評価を等しいとしたとき）では評価者 2 人の予想は 71% の一致率だった。

調査結果に基づいて、オノマトベを表 2 で定義される A, B, C の 3 タイプに分類した。A, B, C それぞれの例を表 4 に示す。たとえば、A に分類した「めらめら」は、人間は正しい意味「燃える」と理解したが、Google 翻訳では異なった意味「Glittering: 輝く」が表示された。A は人間には理解でき、機械には解けないという CAPTCHA にとって最良な語である。逆に、B は、人間には理解できず、機械には解けるという CAPTCHA にとって最悪な語である。

表 4 オノマトペ辞書の一部

Table 4 Part of onomatopoeia dictionary.

意味 <i>S</i>	<i>A</i>	<i>B</i>	<i>C</i>
燃える	めらめら	ばちばち	ぼっ
走る	てくてく	どたん	さっさ
笑う	うはうは	ころころ	うふっ
元気がない	がびーん	とほとほ	くしゅん
はやい	びゅん	ずんずん	さっさ
ゆれる	ぶらぶら	へらへら	ゆさりゆさり
やわらかい	ぶよぶよ	なよなよ	くにゅくにゅ

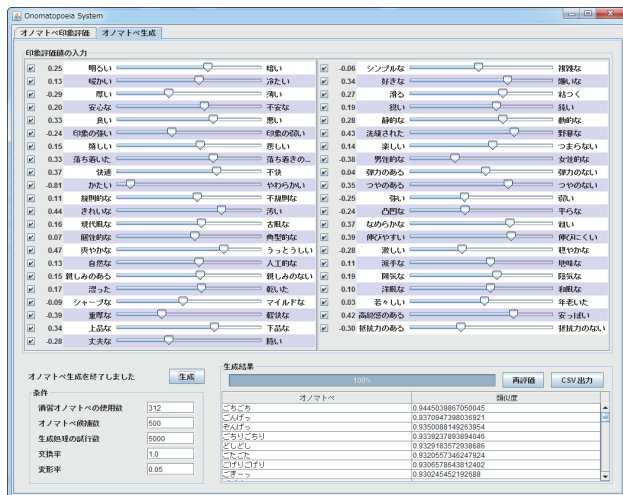


図 3 オノマトペ合成の実行情例 [8]

Fig. 3 Example of onomatopoeia synthesis.

表 5 実験のために合成したオノマトペの例

Table 5 Synthesized onomatopoeia used in our experiments.

印象語	合成オノマトペ				
かたい	ごげりごげり	ぞんげっ	ごんげっ	ごきーっ	
きれいな	さんあり	しゃんあり	さっあり	ひんあ	
滑る	きゅう	しゅり	しゅえ	きゅら	
湿った	ねべ	にゅちよ	ぬっちよ	ねびゃ	
脆い	ひうひう	きろ	しゃろり	しわり	
柔らかい	ふんにゅり	ふーにゃ	ふにゅっ	ふにーり	

3.2 オノマトペ合成

ユーザの感性的印象に適合したオノマトペを合成するシステム [8] を用い、オノマトペを合成した。図 3 は本研究で合成したオノマトペの例である。図 3 では「かたい」、「粗い」、「うっとうしい」のパラメータを大きく設定し「ごんげっ」、「ごちごち」、「ごちりごちり」などのオノマトペが生成されている。本オノマトペ生成システムは、与えた複数の印象評価値に適合した音韻と形態を持つオノマトペを自動合成する。合成アルゴリズムは文献 [8] を参照されたい。表 5 に、本システムに与えた印象語とその出力された合成オノマトペの例をあげる。これらは、すべて google 翻訳では解釈できない語であった (タイプ *C*, または、その他 2~4)。

3.3 問題形式

本研究では、問題形式による *HAR* と、*MAR* の変化を調査し、最適なオノマトペ CAPTCHA の形式を明らかにするため 4 種類の問題形式のオノマトペ CAPTCHA を検討する。5 章でもう 1 つの問題形式について述べる。それぞれの問題形式とその例を以下に示す。

形式 1 (4 択問題)

意味を表す最適なオノマトペを 1 つ選択する。

(例) 「歩く」にあてはまるオノマトペは?

- 1. ばくばく 2. ぼそぼそ 3. とことこ 4. ぎゃはは

(答え) とことこ

形式 2 (オノマトペ選択)

オノマトペから類似のオノマトペを選択する。

(例) 「ぐずぐず」と同じ意味のオノマトペは?

- 1. ほくほく 2. へらへら 3. ほろほろ 4. たらたら

(答え) たらたら

形式 3 (複数選択)

問題の意味にあてはまるオノマトペをすべて選択する。

(例) 「喜ぶ」にあてはまるオノマトペをすべて選べ

- 1. ばくばく 2. うきうき 3. ごによごによ 4. らんらん

(答え) うきうき, らんらん

形式 4 (画像選択)

辞書の意味をもとに画像検索をして画像に合うオノマトペを選択する

(例) 画像の意味と同じ意味のオノマトペを選べ



- ・べらべら
- ・きよろり
- ・どんちゃん
- ・るんるん

(答え) るんるん

3.4 出題方法

問題の出題方式を *A, B, C, RA, RB, RC, SK, SKAll* の 8 種類用意した。 *A* は選択肢のオノマトペがすべて *A* であり、 *B, C* も同様である。 *RA* は正解のオノマトペが *A* であり、ほか 3 つの選択肢は *A, B, C* 1 つずつである。 *RB, RC* も同様である。 *A, B, C, RA, RB, RC* を総じて辞書形式と呼ぶ。

*SK* は合成オノマトペが正解であり、ほか 3 つの選択肢は *A, B, C* 1 つずつである。 *SKAll* は、すべての選択肢が合成オノマトペである。問題形式 3 では、与える語はすべて *A* で正解の数を 1, 2, 3, 4 それぞれ 3 問ずつ用意した。問題形式 4 のタイプ *A* では、問題画像の数を 1 枚から 3 枚をそれぞれ 3 問ずつ用意した。各形式について、辞書と合成語の比較を行う。ただし、形式 1 は、合成語の影響を詳しく見るため *SKAll* を追加し、形式 3 では合成語を省略した。

オノマトベ CAPTCHA の問題を自動生成するアルゴリズムを以下に示す。ここで、 $D_{s,A}$  は、意味  $S$  を持つタイプ  $A$  のオノマトベ集合とする。

**出題アルゴリズム 1 形式 1—出題方式 A**

- (1) 意味  $S$  と語  $x \in D_{s,A}$  をランダムに選ぶ。
- (2)  $S \neq S1, S2, S3$  となる意味  $S1, S2, S3$  と選択肢  $y \in D_{s1,A}, z \in D_{s2,A}, w \in D_{s3,A}$  をランダムに選ぶ。
- (3) 意味  $S$  とランダムな順で  $x, y, z, w$  を出題する。  
 $B, C, SKALL$  の出題方式も同様である。

**出題アルゴリズム 2 形式 1—出題方式 RA**

- (1) 意味  $S$  と語  $x \in D_{s,A}$  をランダムに選ぶ。
- (2)  $S \neq S1, S2, S3$ , 選択肢  $y \in D_{s1,A}, z \in D_{s2,B}, w \in D_{s3,C}$  からランダムに選ぶ。
- (3) 意味  $S$  とランダムな順で  $x, y, z, w$  を出題する。  
 $RB, RC, SK$  の出題方式も同様である。

**出題アルゴリズム 3 形式 2**

- (1) 意味  $S$  と  $x_0, x \in D_{s,A}$  をランダムに選び、 $x_0$  を問題、 $x$  を正答にする。
- (2) (2) はアルゴリズム 1 と同じ。
- (3) 出題語  $x_0$  と選択肢  $x, y, z, w$  をランダムに出題する。

**出題アルゴリズム 4 形式 3**

- (1) 意味  $S$  とその補集合  $\bar{S}$  について、 $1/2$  の確率で、 $x \in D_{s,A}$  と  $x \in D_{\bar{s},A}$  どちらか選ぶ。
- (2)  $y, z, w$  に対しても (1) 同様とする。

**出題アルゴリズム 5 形式 4**

- (1) 意味  $S$  で検索した画像上位 10 枚からランダムに 1 枚選ぶ。
- (2)  $x \in D_{s,A}$  をランダムに選ぶ。
- (3)  $S \neq S1, S2, S3$ , 選択肢  $y \in D_{s1,A}, z \in D_{s2,A}, w \in D_{s3,A}$  からランダムに選ぶ。
- (4) 意味  $S$  とランダムな順で  $x, y, z, w$  を出題する。

**3.5 理論値**

提案方式の選択肢の数や問題形式を変更したときの HAR を見積もるため、理論値を算出する。 $Q$  と  $R$  をクエリとレスポンスの確率変数とし、 $m$  (マッチ) か  $w$  (非マッチ) のどちらかの値をとる (表 6)。このとき、意味の正しいオノマトベ ( $Q = m$ ) を正しい ( $R = m$ ) と判断する条件付き確率を

$$p = Pr(R = m|Q = m),$$

正しくないオノマトベ ( $Q = w$ ) を正しくない ( $R = w$ ) と判断する条件付き確率を

$$q = Pr(R = w|Q = w)$$

とする。

本研究では、形式 1 を正しく回答する理論値を形式 3 か

表 6 マッチと非マッチ

Table 6 Match and miss-match.

$R \setminus Q$	マッチ $m$	非マッチ $w$
マッチ $m$	$p$	$1 - q$
非マッチ $w$	$1 - p$	$q$

ら算出した  $p$  と  $q$  を用いて、

$$\begin{aligned} P(\text{形式 1 を正解する}) \\ &= P(\text{マッチを正当 OR 非マッチを正当} \times 3) \\ &= 1 - (1 - p) \cdot (1 - q^3) \end{aligned}$$

と求められる。

**4. 評価実験**

**4.1 実験目的**

本実験目的は次のとおりである。

実験 1 人間の属性による結果の変化

実験 2 出題形式による人間受入率 HAR と機械受入率 MAR の差を明らかにする

実験 3 既存のオノマトベと合成語や認知度による精度の差を明らかにする

実験 4 既存の CAPTCHA と提案したオノマトベ CAPTCHA の精度の比較をする

**4.2 実験方法**

本研究では、クラウドソーシングサイトで募集した 54 名と菊池研究室の学生 9 名、合計 63 名を被験者とした。被験者には、実験用サイトで 4 形式についての計 84 問の問題を解答してもらい、その解答内容と解答時間を記録した。また、実験終了後に以下の 5 項目についてアンケート調査を行った。漫画はオノマトベを多く含み、本は多くの言葉に触れるという点から、それぞれ読んでいる数が多いほど正解率が高くなるのではないかと考え、アンケート項目に

- (1) 性別
- (2) 年齢 (10 代から 60 代まで)
- (3) 文系か理系か
- (4) この 3 カ月で読んだ漫画の数
- (5) この 3 カ月で読んだ本の数

**4.3 実験結果**

**4.3.1 実験 1 被験者の属性の比較**

アンケート結果に基づく平均 HAR と平均解答時間を表 7 に記す。表から、30 代 40 代が最も平均 HAR が高く、平均解答時間も短かった。逆に 10 代の HAR が最も低かった。性別、文理、漫画と本の読んだ数の差による実験結果の差は、各々、0.032, 0.001, 0.021, 0.004 でいずれも有意な大きさではなかった。

4.3.2 実験 2 出題形式の差

問題形式 1, 2, 3, 4 による実験結果と MAR を表 8 に示す。実験結果からの人間受入率を平均 HAR, 3.5 節で述べた理論値による人間受入率を HAR\* と表す。

MAR<sub>g</sub> は、表 2 のオノマトベ辞書で B に分類される単語が、google 翻訳で正しく意味が解釈されるときに、CAPTCHA が機械を誤って受け入れると仮定して算出した。すなわち、機械が正解する確率が google 翻訳による MAR である。たとえば、形式 1 で選択肢がタイプ A, B, C, A で答えが A の場合、4 択の中で B が答えでないことは分かるので正解する確率は 1/3 となる。形式 1 と形式 2 はすべての出題形式の HAR を算出し、その平均を表 8 に示している。形式 3 では、すべてのオノマトベにタイプ A を使用しているため、ブルートフォースと google 翻訳の MAR は変わらない。すなわち、MAR<sub>b</sub> = MAR<sub>g</sub> である。形式 4 では問題が画像のため、オノマトベの意味を理解できても画像の意味を理解していないと正解することはでき

ないため、google 翻訳はブルートフォースと同じ MAR とした。

google 検索攻撃者の受入率を MAR<sub>gs</sub> とする。たとえば、形式 1 の例題では、「歩く、ばくばく」458,000 件、「歩く、ぼそぼそ」329,000 件、「歩く、とことこ」422,000 件、「歩く、ぎゃはは」9,500 件なので、最もヒット数が多い「ばくばく」を選ぶ、形式 4 では問題に画像を使っているためこの攻撃はできない。

表 8 より、最も HAR が高かったのは、問題形式 4 の画像から意味を読み取り、オノマトベを答える形式であった。問題形式 1 も 0.8 を超える HAR で、問題形式 2 と問題形式 3 はどちらも 0.6 後半の HAR だった。平均解答時間は問題形式による差はほとんどみられなかった。

問題形式 3 では、複数選択の正解の選択肢の数を 1 つから 4 つまで 3 問ずつ用意した。正解の選択肢の数ごとの平均 HAR と、平均解答時間を表 9 に示す。また不正解の解答時に、被験者が実際の正解の選択肢の数に対していくつの選択肢を選択したかを表 10 に示す。表から正解の選択肢の数が 1 つのときの HAR が 0.540 で最も低い。また正解の選択肢の数が 1 つのときに、選択肢を 2 つ選んでしまう被験者がきわめて多かった。

表 11 は形式 4 の問題画像を変化させた際の平均 HAR と平均解答時間である。表から、画像の枚数が少ないほど平均 HAR が高くなり、平均解答時間も短くなった。

4.3.3 実験 3 合成語と辞書

辞書形式 A, B, C, RA, RB, RC ごとの結果を表 12 に示す。最も HAR が高いのは、RB であった。A, B, C では、平均 HAR は高い順に C, A, B となり、解答時間は短い順に C, B, A となった。人間には困難だと思われ、Google 翻訳で正しい意味で翻訳された B が 0.759 で圧倒的に HAR が低かった。

既存のオノマトベと合成オノマトベとの平均 HAR と平均解答時間をそれぞれ表 8 に示す。SK が選択肢に合成語

表 7 アンケート結果

Table 7 Questionnaire results.

	被験者数 N	平均正解数	平均 HAR	平均解答時間
男	29	64.03	0.760	9:47
女	34	66.56	0.792	9:53
10代	8	60.88	0.725	11:27
20代	11	65.55	0.780	10:47
30代	26	66.35	0.790	9:37
40代	15	66.53	0.792	11:02
50代	2	63.00	0.750	10:50
60代	1	63.00	0.750	18:38
文系	43	65.60	0.780	10:45
理系	20	65.40	0.779	10:12
漫画 (n ≥ 10)	22	64.27	0.765	10:44
漫画 (n < 10)	41	66.00	0.786	10:29
本 (n ≥ 10)	18	65.17	0.776	10:21
本 (n < 10)	45	65.49	0.780	10:40

表 8 形式の違いによる受け入れ率の実験結果

Table 8 Experimental result of acceptance ratios in some styles.

形式	1			2		3	4	
	辞書	SK	SKALL	辞書	SK	辞書	辞書	SK
オノマトベの種類								
被験者	日本人			日本人		日本人	日本人	
被験者数 N	63			63		63	63	
出題数 M	18	3	3	18	3	12	24	3
MAR <sub>b</sub>	1/4	1/4	1/4	1/4	1/4	1/16	1/4	1/4
MAR <sub>g</sub>	4/9	1/3	1/4	3/7	1/4	1/16	-	-
MAR <sub>gs</sub>	0.400	0.333	0	0.277	0.333	-	-	-
MAR <sub>d</sub>	1	1/4	0	1	1/4	1	1	1/4
平均 HAR	0.891	0.757	0.656	0.705	0.444	0.680	0.894	0.513
HAR*	0.979	-	-	-	-	表 9	-	-
標準偏差	0.084	0.229	0.290	0.107	0.259	0.162	0.063	0.280
平均解答時間 (秒)	00:07	00:08	00:08	00:08	00:09	00:07	00:07	00:09

表 9 形式 3 の解答数による HAR

Table 9 HAR for number of answers of format 3.

正解数	1	2	3	4
出題数 $M$	3	3	3	3
HAR*	<u>0.704</u>	0.699	0.694	0.688
平均 HAR	0.540	0.709	0.720	<u>0.751</u>
平均解答時間 (秒)	00:07	00:07	00:07	00:08
最大解答時間 (秒)	00:51	00:59	00:51	02:44

表 10 形式 3 間違えた問題数

Table 10 Number of mistaken questions in style No.3.

解答個数 \ 正答個数	1	2	3	4
総不正解数 $N$	<u>87</u>	<u>55</u>	<u>53</u>	<u>47</u>
1	0	21	4	6
2	78	10	36	11
3	7	23	4	30
4	2	0	9	0

表 11 形式 4 画像の枚数ごとの正解率

Table 11 HARs for number of images in style No.4.

画像の枚数	平均時間 (秒)	平均 HAR
1	5	0.937
2	10	0.884
3	11	0.783

表 12 A, B, C による HAR

Table 12 HAR for word types A, B and C.

	A	B	C	RA	RB	RC
出題数 $M$	15	9	9	9	9	9
平均 HAR	0.823	0.759	<u>0.856</u>	<u>0.885</u>	0.872	0.864
平均解答時間 (秒)	00:08	00:07	00:06	00:08	00:07	00:06
最大解答時間 (秒)	02:56	01:12	00:58	03:47	05:22	01:59

オノマトペを 1 つ使用した問題で、SKALL が選択肢のすべてが合成語オノマトペの問題である。合成オノマトペを使用した形式 1, 2, 4 では既存のオノマトペと比較し、どの形式も平均 HAR は低く、平均で 0.25 低い。また、表 8 より 4 つの選択肢すべてに合成オノマトペを使用した問題 (SKALL) は、HAR が 0.656 と選択肢の 1 つに合成オノマトペを使用した問題 (SK) より、0.101 低い。以上より、合成オノマトペは MAR を下げるが同時に、HAR も下げることが示された。

#### 4.3.4 実験 4 既存の CAPTCHA 方式の比較

既存の代表的な CAPTCHA と提案オノマトペ CAPTCHA との平均 HAR と平均解答時間の比較結果を表 13 に示す。既存の CAPTCHA であるパズル型 CAPTCHA と画像認証型 CAPTCHA は 2016 年 8 月に 200 名の被験者に対して評価した。表 13 は、これら既存の CAPTCHA の実験結果と本研究の問題形式 1 の平均 HAR と平均解答時間である。

表 13 既存の CAPTCHA との比較結果

Table 13 Comparison result with the existing CAPTCHAs.

	平均解答時間 (秒)	標準偏差	平均 HAR (%)	標準偏差
オノマトペ	7	9.92	<u>89.1</u>	0.092
パズル	9.44	5.23	88	0.346
画像	18.07	11.65	81	0.388

表 14 補正された既存方式との性能比較

Table 14 Modified performance comparison with the existing CAPTCHAs.

	MAR <sub>b</sub>	補正後		
		問題数 $x$	補正 HAR	解答時間 (秒)
オノマトペ	1/4	4.5	0.595	31.5
パズル	1/90	1.38	<u>0.838</u>	13
画像	1/512	1	0.810	18.07

実験結果から、オノマトペ CAPTCHA はほか 2 つの CAPTCHA と比べて平均 HAR は高く、解答時間も短く見える。しかしコンピュータのブルートフォース攻撃による受入率 (MAR<sub>b</sub>) が異なるため、公平な評価ではない。たとえば、オノマトペ CAPTCHA は 4 つの選択肢、画像認証 CAPTCHA は 9 つの画像の複数選択なので、 $2^9 = 512$  通りから選ぶ。よって MAR<sub>b</sub> = 1/4 > 1/512 であり、画像 CAPTCHA よりはるかに機械に対する安全性が低い。そこで、機械に対する安全性を揃えるために、各問題を  $x$  回繰り返すことを考える。3 方式の MAR<sub>b</sub> を等しくするには、オノマトペを  $x_o$  回、画像を 1 回、パズルを  $x_p$  回解く必要があるので、

$$\left(\frac{1}{4}\right)^{x_o} = \left(\frac{1}{90}\right)^{x_p} = \frac{1}{512}$$

を解いて、

$$x_o = \log_4 512 = 4.5,$$

$$x_p = \log_{90} 512 = 1.38$$

である。したがって、等価な安全性を得るにはオノマトペ CAPTCHA を 4.5 回繰り返す必要があり、8 秒 × 4.5 の平均解答時間と  $0.891^{4.5} = 0.595$  の HAR に補正される。以上の補正結果を表 14 に示す。

それぞれの CAPTCHA のブルートフォース攻撃の確率の求め方は以下に示す。

(1) オノマトペ CAPTCHA のブルートフォース攻撃

表 14 のオノマトペ CAPTCHA は形式 1 の 4 択問題の場合である。4 つの選択肢から答えを 1 つ選択するため MAR<sub>b</sub> は 1/4 となる。

(2) パズル型 CAPTCHA のブルートフォース攻撃

パズル型 CAPTCHA はパズルのピースをはめこむ CAPTCHA である。マスが 30 × 27 あり、正解の位置前後 9 マスが受け入れられる。よってパズル型 CAPTCHA の MAR<sub>b</sub> は、1/90 となる。

表 15 複数選択問題における正しい選択数

Table 15 Number of correct selections in multiple selection question.

形式 3 すべての問題					
選択肢番号	1	2	3	4	計
選択肢数 $M$	378	567	441	504	1,890
選択した数	346	520	405	468	1,739
形式 3 で選択肢が 1 つのもののを除いた問題					
選択肢数 $M$	378	504	378	441	1,701
選択した数	346	457	342	405	1,550

表 16 複数選択問題における間違った選択数

Table 16 Number of wrong selections in by multiple selection question.

形式 3 すべての問題					
選択肢番号	1	2	3	4	計
選択肢数 $M$	378	189	315	252	1,134
選択した数	254	177	307	252	990
形式 3 で選択肢が 1 つのもののを除いた問題					
選択肢数 $M$	189	63	189	126	567
選択した数	143	63	189	126	521

(3) 画像認証型 CAPTCHA のブルトフォース攻撃

画像認証型 CAPTCHA が 9 枚の画像から正しいものをすべて選択するため、 $MAR_b = 2^{-9} = 1/512$  となる。

4.4 形式 1 の HAR\*

形式 3 は複数選択問題のため、各々のオノマトベに対して正しいか正しくないかを確認している。表 15 では形式 3 で正しいオノマトベを解答した数、表 16 では誤ったオノマトベを選択した数を示している。この割合から

$$p = 1550/1701 = 0.911,$$

$$q = 521/567 = 0.918$$

を得て、3.5 節で述べた理論値を見積もる。

ただしここで、解答数が 1 つの問題はきわめて HAR が低かったことを考慮して、形式 3 の答えが 1 つの問題を抜いた数値を使用している。複数問題と表示されているため答えが 1 つではないという心理的影響を受けたと考えられる。

4 択問題では 4 つの選択肢から 1 つを選択するため、形式 1 の HAR の理論値は、

$$\begin{aligned} P(\text{形式 1 を正解する}) \\ &= 1 - (1 - 0.911) \cdot (1 - 0.918^3) \\ &= 0.979 \end{aligned}$$

と予測できる。

理論値 HAR\* は 0.979 となったが、本実験の形式 1 の辞書、SK, SKAll の平均 HAR は 0.845 となった。理論値 HAR\* と比較して HAR は 0.134 低い。形式 3 ではタイプ

A のオノマトベのみを使用している。一方、形式 1 では A, B, C のほかに合成オノマトベを混入している。表 8 より合成オノマトベを抜いた形式 1 の人間受入率は 0.891 であり、人間に理解しやすいオノマトベが多かったため、この差が生まれたと考えられる。

形式 3 の正解数を変化させた場合の HAR\* を表 9 に示す。HAR\* では正解数を増やすほど正答率は低くなっていくが、実験結果では正解数を増やすほど HAR は高くなっている。

同様に、形式 2 と 4 についても、対応する複数選択問題の実験結果に基づいて、理論値を算出することができる。たとえば、形式 2 (オノマトベ選択) については、「わくわく」と合致するオノマトベをばくばく、うきうき、ごによごによ、らんらんの中からすべて選べ、という形式で実験し、

$$p' = Pr(R = m|Q = m(\text{わくわく, うきうき})),$$

$$q' = Pr(R = w|Q = w(\text{わくわく, ばくばく}))$$

の条件付き確率を求めれば、3.5 節の式によって、4 択の形式 2 の理論値が与えられる。形式 4 の理論値も同様である。ただし、この複数選択形式の実験を新たに追加するには本文で大規模な変更が必要であり、本稿では割愛する。

4.5 解答時間のヒストグラム

年代別の総解答時間と HAR の分布をそれぞれ図 4, 図 5 に示す。どちらの分布も、年代による大きな差は示されなかった。いずれも、平均値 590[s], 0.776 の正規分布に従っている。

4.6 問題・出題形式の差の確率検定

出題・問題形式の違いによる解答時間・HAR への影響を確かめるために t 検定を行った結果を表 17, 表 18 にそれぞれ示す。有意な P 値に\*をつけている。形式 1, A をそれぞれ基準、 $p < 0.05$  を有意水準とした場合、すべての問題・出題形式について平均 HAR については形式 1, A に対してそれぞれ有意差が見られ、問題・出題形式による平均 HAR は独立であるという帰無仮説は棄却される。すなわち、問題形式によって難易度が異なっていたということが示された。

4.7 考察

本実験より、人間にとってオノマトベは年齢や性別など属性に関係ないことが明らかになった。また問題形式 1~4 の結果から、問題形式 4 の HAR が 0.894 で 4 つの形式の中で最も高いことが示された。したがって、人間はオノマトベを形式 1~3 のような明確な意味ではなく、形式 4 のようなイメージで把握しているからではないかと推定されるが、その差は微量であり、今後、新たな実験を積み重ね



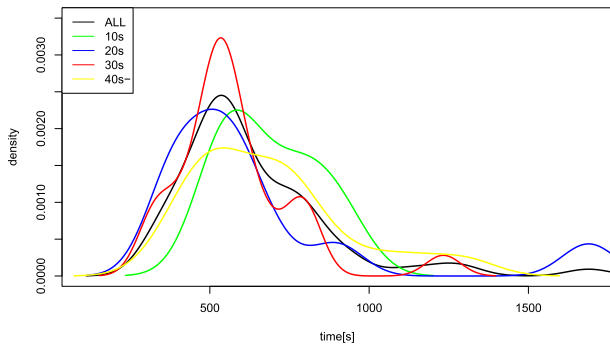


図 4 年代ごとの総解答時間の分布

Fig. 4 Distribution of total solution time in ages of subjects.

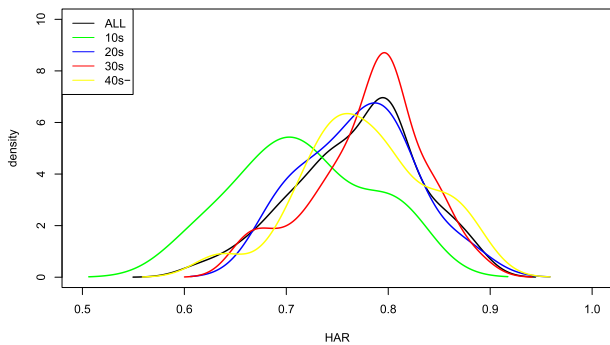


図 5 年代ごとの HAR の分布

Fig. 5 Distribution of HAR in ages of subjects.

表 17 問題形式についての仮説検定

Table 17 Hypothesis test on question type.

問題形式	平均	statistic	p-value
解答時間についての検定			
(intercept)	00:07	-	-
2	00:08	-2.169	0.030*
3	00:07	-0.717	0.472
4	00:07	-0.767	0.442
HAR についての検定			
(intercept)	0.891	-	-
2	0.668	9.734	6.90e-17*
3	0.680	6.904	5.05e-10*
4	0.894	-0.477	6.33e-01*

表 18 検査語についての仮説検定

Table 18 Hypothesis test on independent of class of words.

検査語	平均	statistic	p-value
解答時間についての仮説検定			
(intercept)	00:08	-	-
B	00:07	2.637	0.008*
C	00:06	2.977	0.002*
HAR についての仮説検定			
(intercept)	0.823	-	-
B	0.759	2.637	1.89e-04*
C	0.856	2.977	1.48e-01*

て明らかにしていく必要がある。

また、問題形式 3 で解答個数が 1 つのときに、HAR が低く、2 個以上選択してしまう被験者が多かったのは、「すべて選べ」という設問文による心理的影響が大きいと考えられる。

オノマトペ辞書や合成オノマトペの結果を比較したところ、タイプ B や合成オノマトペの HAR が低かったことから、普段馴染みのないオノマトペは人間にはイメージしにくく少し困難である。

既存の CAPTCHA に比べてオノマトペ CAPTCHA は HAR が高く、解答時間も短かった。しかし、ブルートフォースの耐性を考慮すると、提案オノマトペ CAPTCHA の精度や解答時間はまだ十分でないと考えられる。また、以上の実験をオノマトペの発達している日本語で日本人についてのみ行っているため、普遍的検査方法であるか疑問が残る。そこで次章ではさらなる改良を加える。

## 5. 日英比較

### 5.1 提案方式

#### 5.1.1 マンガオノマトペ CAPTCHA

実験 1 から画像を用いたオノマトペ CAPTCHA が有効であることが分かった。またオノマトペ CAPTCHA の言語的な背景知識への依存度を小さくするために、新たにマンガオノマトペ CAPTCHA (形式 5) を提案する。英語版 (5E)、日本語版 (5J) を用意し日英の比較を行う。5E<sub>J</sub> は英語版を日本人の被験者に、5E<sub>e</sub> は外国人の被験者に対して行ったものとする。

#### 5.1.2 問題形式例

##### 形式 5J (日本語マンガオノマトペ CAPTCHA)

日本語のマンガに用いられているオノマトペを隠して与え、選択肢から適切なオノマトペを 1 つ選択する。例を図 6 に示す。

##### 形式 5E (英語マンガオノマトペ CAPTCHA)

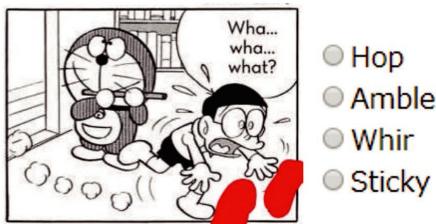
本形式は、英語に翻訳されているマンガを用いた、形式 5J と同様なテストである。例を図 7 に示す。



画像の赤く塗られたところに合いそうな語は？

図 6 形式 5J の例 (答え:「いらいら」) [13]

Fig. 6 Example of Style 5J (Answer: “iraira”).



Which word that are likely to fit in the red painted image?

図 7 形式 5E の例 (答え: “Hop”) [14]

Fig. 7 Example of Style5E (Answer: “Hop”).

表 19 日英比較の実験概要

Table 19 Outline of examples in terms of Japanese and English Difference.

	実験 5	実験 6
目的	提案方式 3 の有用性	英語圏被験者の属性分析
言語	日	英
形式	1・2・5J	5J・5E

### 5.2 実験目的

本実験の目的は以下のとおりである。

**実験 5** 新たな形式の日本語マンガオノマトベ CAPTCHA の HAR を明らかにし、従来手法 (形式 1, 形式 4) の結果と比較する。

**実験 6** 新たな英語マンガオノマトベ CAPTCHA の HAR を明らかにし、日英の比較や年代・国籍・性別など属性の違いによる HAR の変化を明らかにする。

以上の実験の概要を表 19 にまとめる。

### 5.3 実験方法

本研究では、主に、著者 (滋野) の SNS で募集した日本人 57 名、外国人の 63 名の計 120 名を被験者とした。日本人には形式 5J と 5E を各 10 問、外国人には 5E を 10 問出題した。実験はウェブ上でを行い、解答内容と解答時間を記録した。実験終了後に、アンケート調査も行った。

### 5.4 実験結果

#### 実験 5 マンガオノマトベ CAPTCHA

マンガオノマトベ CAPTCHA の実験結果と MAR を表 20 に示す。表 8 の形式 1, 形式 4, 形式 5J の結果より、形式 5J の日本語マンガオノマトベ CAPTCHA の HAR が最も高かった。3J の平均解答時間は、形式 1, 形式 4 の 7 秒の倍の 14 秒かかっている。各問題の散布図を図 8 に示す。図 8 より形式 1 や形式 4 は HAR が高く、平均解答時間も短い。一方で HAR の低い形式  $E_e$  や形式  $E_j$  は解答時間も長い。図 8 より、時間をかけたからといって理解できるわけではないことが分かる。

#### 実験 6 日英比較と人種、属性別の分析

表 8 の形式 5J と  $5E_j$ ,  $5E_e$  の実験結果より、日本語マンガオノマトベ CAPTCHA と比較して英語オノマトベ

表 20 形式 5 の受け入れ率の実験結果

Table 20 Experimental result of acceptance ratios in style 5.

オノマトベの言語	J	$E_j$	$E_e$
被験者	日本人	日本人	外国人
被験者数 $N$	57	57	63
出題数 $M$	10	10	10
$MAR_b$	1/4	1/4	1/4
$MAR_g$	-	-	-
$MAR_{gs}$	-	-	-
平均 HAR	0.895	0.452	0.410
$HAR^*$	-	-	-
平均解答時間 (秒)	00:14	00:26	00:18

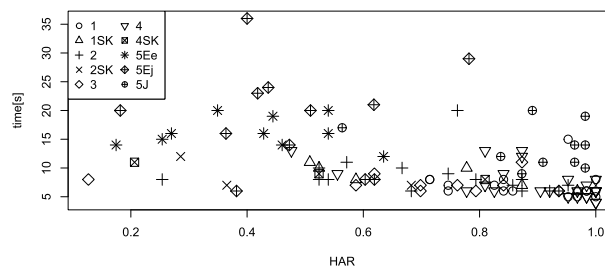


図 8 各問題の散布図

Fig. 8 Scatter plot of answering time and HARs for queries.

CAPTCHA の HAR は 0.485 低かった。日本人に対して行った英語マンガオノマトベ CAPTCHA も同様である。本結果より、提案方式は被験者の言語的知識に依存している可能性がある。その原因を次に検討する。

本実験の被験者となった外国人の国籍、人数、地域、性別、年代、国ごとの被験者数と HAR を表 21 に示す。地域ごとの HAR を表 22 に示す。表 22 より、北アメリカが最も HAR が高く、続いて高いのがヨーロッパであった。

英語力順位 [15] と読解力順位 [16]、英語を話せるか、性別、年齢など属性による分割表を表 23, 表 24 に示す。英語力順位は、2015 年に Education First の英語テストを受験した 95 万名を超える成人のデータを基にした 72 カ国の順位である。読解力順位は、2015 年に OECD の 15 歳を対象とする PISA (Programme for International Student Assessment) 調査の順位である。それぞれの順位を基に上位から被験者の出身国を 1 から 13 位まで順位づけ、ランキング外の国を 14 位とした。

英語力順位では 8 位未満に比べ 8 位以上の HAR が 0.035 高く、読解力順位では 8 位未満に比べて 8 位以上の HAR が 0.064 高い。また、英語が話せると答えた人は、話せないと答えた人に比べて HAR が 0.059 高い。したがって、HAR の差が生じた原因の 1 つは、これらの英語力などの知識である可能性がある。しかし、他の言語との差異は要因が多数存在することが容易に考えられ、翻訳精度の妥当性や被験者の言語学的背景の差異などが議論されないままでは言語的知識の問題とはいきれないため、引き続き研究が必要である。

表 21 被験者の属性別統計と形式 5 の HAR

Table 21 Demographical statistics of subject and HAR of Format5.

	被験者数 $N$	地域	英語力順位 [15]	読解力順位 [16]	性別		年代			形式 5E の HAR	
					男性	女性	10代	20代	30代		
United States	1	北アメリカ	1	6	1	0	0	1	0	0.700	
Colombia	2	南アメリカ	14	11	1	1	1	1	0	0.300	
Brazil	3		11	12	0	3	0	1	2	0.500	
Korea	15	アジア	6	1	6	9	1	10	4	0.393	
Taiwan	7		9	5	4	3	0	7	0	0.371	
China	2		10	7	2	0	1	1	0	0.450	
Israel	1		14	9	0	1	0	1	0	0.600	
Turkey	1		12	10	1	0	0	1	0	0.300	
Indonesia	1		8	13	0	1	0	0	1	0.300	
Philippines	5		4	14	5	0	0	2	3	0.300	
Saudi Arabia	1		13	14	0	1	0	1	0	0.200	
Nepal	18		14	14	13	5	1	16	1	0.394	
Germany	2		ヨーロッパ	3	2	1	1	1	1	0	0.550
Sweden	1			2	3	0	1	0	0	1	0.300
France	2	7		4	1	1	0	2	0	0.650	
Czech Republic	1	5		8	0	1	0	1	0	0.900	
total	63	-	-	-	35	28	5	46	12	0.410	

表 22 地域ごとの形式 HAR

Table 22 Regional statistics of HAR.

地域	被験者数 $N$	マンガ平均 HAR
北アメリカ	1	0.7
南アメリカ	5	0.42
アジア	51	0.38
ヨーロッパ	6	0.6
計	63	0.41

表 23 語学力, 読解力についての HAR の分割表

Table 23 Contingency table of HAR for language skill and reading skill.

	英語力国別順位 [15]		読解力国別順位 [16]		英語話せる	
	8位以上	8位未満	8位以上	8位未満	話せる	話せない
被験者数 $N$	28	35	31	32	40	23
平均 HAR	0.429	0.394	0.442	0.378	0.427	0.378

表 24 性別, 年齢についての HAR の分割表

Table 24 Contingency table of HAR for sex and age.

	性別		年齢		
	男	女	10代	20代	30代
被験者数 $N$	35	28	5	46	12
平均 HAR	0.413	0.407	0.389	0.429	0.388

## 6. おわりに

本研究で行った実験の目的と結果のまとめを表 25 に示す。本研究では「オノマトペ CAPTCHA」を提案し、人間受入率 HAR と機械受入率 MAR の性能評価を行った。4 択形式による提案方式の平均 HAR は 0.891, 平均解答時間

表 25 実験目的と実験結果

Table 25 Experiment objectives and experiment results.

実験番号	実験目的	実験結果
1	被験者の属性による結果の変化	属性の違いは実験結果に影響しない
2	出題形式による精度の差	形式 4 (画像) と形式 5J (マンガオノマトペ CAPTCHA) の人間受入率 (HAR) が高い (統計的に有意差有)
3	既存のオノマトペと合成オノマトペの比較	合成オノマトペの安全性が高く (MAR が低), しかし, 有用性も低下した (HAR が低)
4	既存 CAPTCHA との精度の比較	提案方式は平均解答時間が最短だが, 安全性を補正すると既存方式に劣る
5	マンガオノマトペ CAPTCHA の有用性	提案方式の中で最も HAR が高い
6	英語圏被験者の属性分析	英語圏被験者の受入率 (HAR) は, 非英語圏のものより高い

は 7 秒, 検索エンジンによる MAR は 0.400 であった。1 問あたりで比較するとオノマトペ CAPTCHA は人間に優しくストレスなく解くという観点では, 既存の CAPTCHA よりも優れているといえる。しかし機械に対する安全性を考えるとまだ十分な性能を持つ CAPTCHA とはいえない。

検索エンジン攻撃に対しても安全な自動合成したオノマトペを用いた 4 択形式の平均 HAR は 0.656 であり, 既知オノマトペ辞書によるものより低い。

また言語的背景知識への依存度を小さくするためにイラストやマンガを用いた新しい CAPTCHA を提案した。日

本人と外国人の被験者によるオノマトペの受入率に平均 0.485 の差があることを明らかにした。

今後は、MAR を下げ、日本語以外のオノマトペ CAPTCHA の HAR をあげることを目標とする。

#### 参考文献

- [1] Bursztein, E., Martin, M. and Mitchell, J.: Text-based CAPTCHA Strengths and Weaknesses, *Proc. 18th ACM Conference on Computer and Communications Security, CCS'11*, pp.125–138, ACM (2011).
- [2] Bursztein, E., Beauxis, R., Paskov, H.S., Perito, D., Fabry, C. and Mitchell, J.C.: The Failure of Noise-Based Non-continuous Audio Captchas, *32nd IEEE Symposium on Security and Privacy, S&P 2011*, pp.19–31 (2011).
- [3] Project Stiltwalke, available from <http://www.dc949.org/projects/stiltwalker/> (accessed 2013-05-01).
- [4] Sano, S., Otsuka, T., Itoyama, K. and Okuno, H.: HMM-based Attack on Google's ReCAPTCHA with Continuous Visual and Audio Symbols (Preprint), *IPSJ Journal*, Vol.56, No.11 (2015).
- [5] Sivakorn, S., Polakis, I. and Keromytis, A.D.: I Am Robot: (Deep) Learning to Break Semantic Image CAPTCHAs, *Proc. 1st IEEE European Symposium on Security and Privacy, EuroSP'16* (2016).
- [6] 藤田真治, 池谷勇樹, 西垣正勝: 非現実画像 CAPTCHA: 常識からの逸脱を利用した 3DCG 画像 CAPTCHA, *情報処理学会論文誌*, Vol.56, No.12, pp.2324–2336 (2015).
- [7] 山口通智, 岡本 健, 菊池浩明: 機械合成文の不自然度相対識別に基づく CAPTCHA の提案, *情報処理学会論文誌*, Vol.56, No.9, pp.1834–1835 (2015).
- [8] 清水祐一郎, 土斐崎龍一, 鍵谷龍樹, 坂本真樹: ユーザの感性的印象に適合したオノマトペを生成するシステム, *人工知能学会論文誌*, Vol.30, No.1, pp.319–330 (2015).
- [9] 小松孝徳, 秋山広美: ユーザの直観的表現を支援するオノマトペ意図理解システム, *電子情報通信学会論文誌*, Vol.J92-A, No.11, pp.752–763 (2009).
- [10] 小野正弘: 擬音語・擬態語 4500 日本語オノマトペ辞典 (2007).
- [11] Capy パズルキャプチャ, available from [https://www.capy.me/jp/products/puzzle\\_captcha](https://www.capy.me/jp/products/puzzle_captcha).
- [12] ReCAPTCHA demo, available from <https://www.google.com/recaptcha/api2/demo>.
- [13] 国立国語研究所: 日本語を楽しもう!e-Japan, 入手先 <https://pj.ninjal.ac.jp/archives/Onomatope/manga/iraira2.html>.
- [14] 藤子・F・不二雄: DORaEMON セレクション 2, 小学館, p.8 (2013).
- [15] Education Fast, English Proficiency Index, available from <http://www.efjapan.co.jp/epi/>.
- [16] 国立教育政策研究所: OECD 生徒の学習到達度調査—2015 年調査国際結果の要約, 入手先 [http://www.nier.go.jp/kokusai/pisa/pdf/2015/03\\_result.pdf](http://www.nier.go.jp/kokusai/pisa/pdf/2015/03_result.pdf).

#### 推薦文

DICOMO2017 の発表論文の中で特に評価が高かった。オノマトペを利用したキャプチャの評価というテーマに新規性があり、評価方法・結果も信頼できるものであった。

(セキュリティ心理学とトラスト研究会主査 寺田真敏)



滋野 莉子

2018 年明治大学総合数理学部先端メディアサイエンス学科卒業。現在、株式会社ベネッセコーポレーション所属。



山田 道洋 (学生会員)

2017 年明治大学総合数理学部先端メディアサイエンス学科卒業。現在、明治大学大学院博士前期課程在学中。



菊池 浩明 (正会員)

1988 年明治大学工学部電子通信工学科卒業。1990 年同大学大学院博士前期課程修了。1994 年同博士 (工学)。1990 年 (株) 富士通研究所入社。1994 年東海大学工学部電気工学科助手。1995 年同専任講師。1999 年同助教教授。2006 年同情報理工学部情報メディア学科教授。1997 年カーネギーメロン大学計算機科学学部客員研究員。2013 年明治大学総合数理学部先端メディアサイエンス学科教授。2016 年同大学院先端数理科学研究科長。WIDE プロジェクト暗号メールシステム FJPEM の開発、認証実用化実験協議会 (ICAT)、IPA 独創情報技術育成事業等に従事。暗号プロトコル、ネットワークセキュリティ、ファジイ論理、プライバシー保護データマイニング等に興味を持つ。1990 年日本ファジイ学会奨励賞、1993 年情報処理学会奨励賞、1996 年 SCIS 論文賞、2010 年度、2017 年度情報処理学会 JIP Outstanding Paper Award。2013 年 IEEE AINA Best Paper Award。2014 年情報セキュリティ文化賞。電子情報通信学会、日本知能情報ファジイ学会、IEEE、ACM 各会員。本会フェロー。



坂本 真樹 (正会員)

1998 年 3 月東京大学大学院総合文化研究科言語情報科学専攻博士課程修了。同専攻助手、電気通信大学電気通信学部准教授を経て、2015 年 4 月より同大学大学院情報理工学研究科教授。人工知能学会、VR 学会、感性工学会、認知科学会等各会員。