This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2021.3130277, IEEE Access

IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Date of publication xxxx 00, 0000, date of current version xxxx 00, 0000.

*Digital Object Identifier 10.1109/ACCESS.2017.DOI*

# Unsupervised Learning Enables Extraction of Tactile Information from Text Database

**TATSUHO NAGATOMO[1], TAKEFUMI HIRAKI[2], (Member, IEEE), HIROKI ISHIZUKA[3], AND NORIHISA MIKI[1], (Member, IEEE)**

[1]Graduate School of Science and Technology, Keio University, Yokohama 223-8522, Japan
[2]Faculty of Library, Information and Media Science, University of Tsukuba, Tsukuba 305-8550, Japan
[3]Graduate School of Engineering Science, Osaka University, Toyonaka 560-8531, Japan

Corresponding author: Tatsuho Nagatomo (e-mail: tatsuho19950307@keio.jp).

This work was supported by JSPS KAKENHI Grant Number JP20H02121 and JST CREST Grant Number JPMJCR19A2, Japan.

**ABSTRACT** In this work, we propose a new approach to tactile research using natural language processing of archival word corpus as the database. Tactile perception, or assessment of surfaces, is recognized as a language. Thus, by extracting touch-related words and sentences from a text corpus and learning their relationships, we can ultimately learn how humans perceive surfaces. We selected 6 adjectives and 42 onomatopoeias in Japanese as our tactile words. The adjectives represent physical properties, such as roughness and hardness, while onomatopoeias, such as "zara-zara" and "tsuru-tsuru," are widely used to describe surfaces in Japanese and can correspond to both physical texture cognition and affective cognition. First, using natural language processing of word corpora, we successfully mapped the onomatopoeias with respect to the 6 adjectives, which matched well with the results based on an enquete-based survey. This verified the effectiveness of natural language processing for tactile research. In addition, principal component analysis revealed new tactile dimensions based on onomatopoeias, which we presumably assessaffective tactile dimensions. The proposed approach using natural language processing of archival text databases can provide a large number of datasets for tactile research and culminate in new findings and insights.

**INDEX TERMS** Machine Learning, Unsupervised Learning, Tactile Perception, Natural Language Processing, and Onomatopoeia.

## I. INTRODUCTION

The sense of touch needs to be exploited alongside sight and hearing to develop XR applications. To this end, tactile perception must be characterized quantitatively, by determining the perceptual dimensions. Several studies attempted to extract these dimensions, which include, but are not limited to, roughness, friction, hardness, moisture, and warmness [1], [2]. These dimensions are associated with the physical properties of materials.

"Roughness" and "Friction" can correspond to the surface roughness [m] and the friction coefficient [-], respectively. "Hardness" is, in fact, the stiffness or Young's Modulus [Pa]. "Moisture" can be assessed by wettability or surface energy [N/m]. "Warmness" is associated with the temperature [K] [3]. However, these dimensions do not provide a precise quantitative characterization of the surface properties and can be insufficient to assess tactile perception. For example, the surface roughness can be described quantitatively in terms of many different forms of parameters, such as Rz, RSm, and Rq. Friction feeling is reported to be more dependent on the mean variation of the friction coefficient than the friction coefficient itself [4]. "Warmness" can be controlled by not only temperature but also thermal conductivity [5]. Therefore, when tactile perception or surface sensing is assessed with respect to these dimensions, much information can be lost.

Here, we would like to discuss another approach to selecting tactile dimensions. Tactile perception can be recognized in the form of a word [6]–[8]. In order to describe surface texture and tactile perception, we try to find the appropriate words that represent them. Adjectives, such as rough, smooth, and stiff, are good candidates. Also used

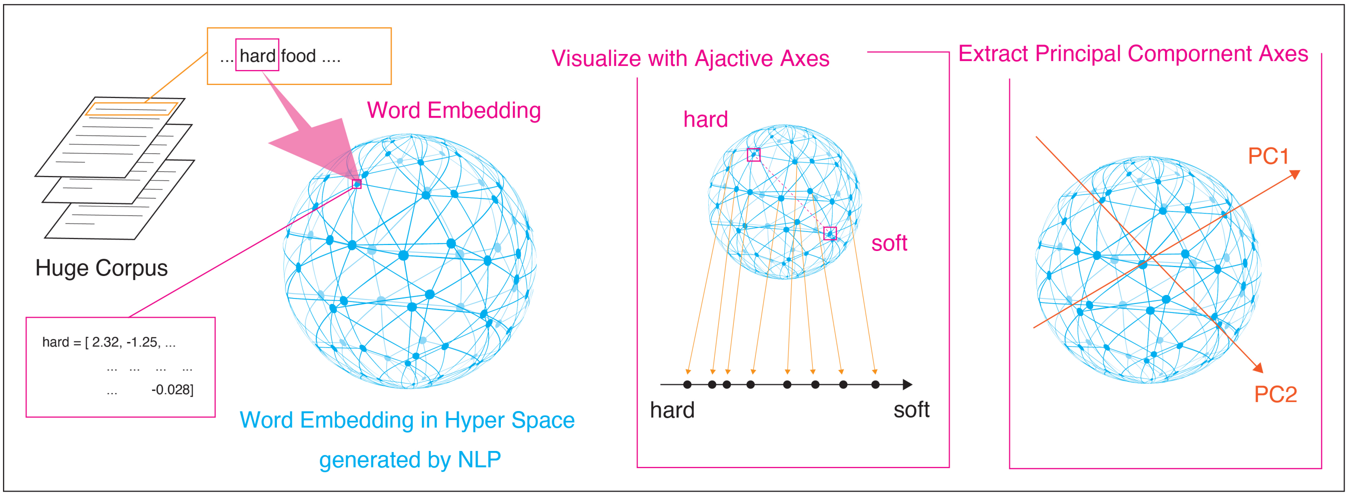


**FIGURE 1. Concept of proposed approach to derive tactile perceptional dimensions from a huge text corpus. NLP enables learning of semantic relations between words from such corpora and generation of word embeddings of tactile onomatopoeia and adjectives. It can generate onomatopoeic maps by mapping the onomatopoeias onto the adjective axis and sort them based on the principal component axis.**

frequently in Japanese are onomatopoeias, which comprise repetitions of a few letters. For example, "zara-zara" and "tsuru-tsuru" describe rough and smooth surfaces, respectively. In order to correlate onomatopoeias to physical properties, onomatopoeia maps were proposed with respect to the various tactile dimensions on the basis of a questionnaire-based survey [9]. It was reported that onomatopoeias could distinguish minute differences in materials better than adjectives, and that onomatopoeias are more effective than adjectives in understanding the differences in individuals' sensory perception of texture [10].

Onomatopoeias are also reported to be able to describe both texture cognition and affective cognition [11] [6]. An onomatopoeia map uses tactile dimensions based on physical properties, which can express texture cognition. However, affective cognition may not be able to be mapped onto the tactile dimensions. There may exist new tactile dimensions correlating to affective cognition that can be described with onomatopoeias.

A major challenge in tactile research is that tactile perception experiments are time-consuming. In many cases, participants need to go to an experimental site to touch samples or tactile displays physically and label the perception. It is difficult to find many participants and thus collect a lot of data. In the present study, instead of conducting perception experiments with human participants, we take a different approach. As discussed above, when we describe a surface texture or tactile perception, we use words that represent that texture or perception. These words are recorded in archival text corpora, such as Wikipedia. Thus, we believe that we can extract tactile dimensions composed of onomatopoeias by conducting natural language processing (NLP) on corpora containing an extremely large number of user data. We can access several large text corpora, such as Wikipedia, and analyze them effectively thanks to recent advances in natural language processing.

In this paper, we propose and demonstrate a novel approach to deriving the relationship between tactile perceptional dimensions from a large word corpus using NLP techniques. In this approach, no human labeling of tactile information is necessary. We vectorize all the candidate words, which are 6 adjectives and 42 onomatopoeias, using a distributed representation of words from the corpus. We apply several NLP methods and attempt to deduce the relationship between the words, or the onomatopoeia map. First, we verify the effectiveness of our approach by comparing the map to a previously reported map and enquete-based survey. Then, affective tactile dimensions are extracted by PCA analyses. The proposed method can automatically create onomatopoeic maps from textual information without the need for time-consuming subject experiments. In addition, this technology can be easily applied to all languages (Fig. 1).

## II. APPROACH AND METHOD

We used natural language processing to acquire tactile dimensional representations using onomatopoeia, without the need for subject experimentation. One of the basic methods in natural language processing is to represent words in a corpus as high-dimensional vectors (word embeddings) in order to learn the grammatical and semantic relationships in the text corpus. It is usually constructed using machine learning algorithms, such as Word2Vec [12] or FastText [13]. These algorithms use information about the co-occurrence of words in a text corpus. For example, if we learn from a rich corpus of word meanings, the vector representing the word "school" should have a close cosine distance to the vector representing "university" and the vector representing "education". Importantly, skip-gram-type NLP algorithms, such as Word2Vec, which learn meaning by predicting words that appear in the neighborhood of the target word, do not require human labeling, because they use neighboring words as supervised data. In contrast to conventional visualization

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2021.3130277, IEEE Access



tools for tactile information such as onomatopoeia maps [9], which are based on the results of subject experiments, our method uses archived corpora as supervisory data to create vectors of tactile words, allowing visualization of tactile information without subject experiments (i.e., without human labeling). Since tactile perception is recognized as a language [6]–[8], we think that it is possible to extract semantic information from tactile information by acquiring word embedding expressions from a set of semantic data on the language (corpus) and restricting the words to those that express tactile sensation.

### A. NATURAL LANGUAGE PROCESSING MODELS

In our proposed method, we used FastText [13], a skip-gram NLP algorithm, and BERT [14], a general-purpose natural language processing model employing skip-grams. FastText is an improved model of Word2Vec [12] that can handle unknown words. It divides each word into sub-words, extracts features for each letter's n-gram, and generates a word vector by adding each feature vector. Thus, the model can deal with word conjugations, similar words, compound words, and unknown words that do not exist in the corpus. BERT is a general-purpose NLP model, a fine-tuning model that consists of bidirectional encoding representations using Transformers [15], and can be used to pre-train word embeddings and solve NLP tasks in the same model. BERT is very powerful because it uses Transformers to identify polysemous words from the way they are used in sentences; however, it is not designed to strictly represent the relationships between words in terms of cosine distances [14].

Another powerful natural language processing model other than BERT is GPT-3 [16], a unidirectional autoregressive language model (AR model), which can generate sentences that can be mistaken for human-generated ones using a huge text dataset of 45TB and about 175 billion parameters as a model. There is also a text-to-image model called DALL-E [17], which is an application of GPT-3. This model takes a caption as input and outputs an image dataset that matches the meaning of the text. Like the images in this model, it is expected to be applied to tactile information. On the other hand, unlike BERT, it generates vectors through a unidirectional autoregressive process with context weighting, so it is not suitable for tasks where we want to estimate words from both front and back contexts. In addition, it is challenging to prepare a corpus in Japanese for constructing GPT-3, and it is computationally expensive, so we excluded it from this study. The successors to BERT [14] are XLNeT [18], RoBERTa [19], ALBERT [20], and ELECTRA [21]. XLNet is an improved version of BERT's Masked Language Model by applying an AR model. RoBERTa uses a model that improves accuracy by tuning BERT's hyper-parameters; ALBERT improves accuracy by separating the dimension of the word embedding representation from the dimension of the hidden layer and sharing the weights for each layer, making the model lighter; ELECTRA is an adversary generation Instead of masking the input, ELECTRA replaces some tokens with plausible alternatives sampled from a small generator network to achieve the same level of accuracy with about a quarter of the computational complexity of XLNet and ALBERT. Although these models have improved accuracy and computational complexity, the generated features of vectors are not expected to change significantly. Therefore, in this study, we decided to use a model pre-trained on the already existing Japanese Wikipedia to see if tactile information can be extracted from a pre-trained model trained on a generic model.

**TABLE 1.** 42 selected onomatopoeias and their English translations.

| Onomatopoeia | Word meanings |
|---|---|
| kasa-kasa | dry, rustle, bone dry |
| gasa-gasa | dry or rough feeling, rustling |
| kunya-kunya | (the impression of being) soft and flexible |
| gunya-gunya | flabby, limpness, soft |
| kunyo-kunyo | flabby, soft, flexible |
| keba-keba | garish, gaudy, ostentatious |
| kochi-kochi | tense, scared stiff, frightened |
| gotsu-gotsu | rugged, scraggy, rough, angular |
| kori-kori | crunchy, crisp (like a pickle), stiff (musculature) |
| gori-gori | scratching, hard (to the bite, to the touch), scraping |
| gowa-gowa | starchy, stiff |
| sara-sara | murmuring, fluently |
| zara-zara | gritty, granular, rough (touch, voice, etc.), coarse |
| jyari-jyari | gritty (sand, pebbles, etc.), crunchy, chunky |
| syori-syori | scratching, scrubbing, scraping |
| jyori-jyori | the feeling of touching something hard and short-hair-like |
| siwa-siwa | crumpled, wrinkled, crinkled |
| sube-sube | sleek, smooth (skin, etc.), silky |
| chiku-chiku | stinging, tingling, prickling |
| tsubu-tsubu | lumpy, grains, bumpiness |
| tsuru-tsuru | slippery, sleek, slick, smooth |
| toge-toge | prickly, thorny, spiny, stinging |
| toro-toro | syrupy, simmering, dozing |
| nyuru-nyuru | slurping, slithering |
| nume-nume | slimy, slippery, wet, glistening |
| nuru-nuru | slimy, slippery |
| necha-necha | slippery, messy (from mud, ink, etc.), sticky |
| necho-necho | prickly, sticky, stickily |
| neba-neba | stickiness |
| fuka-fuka | soft (and fluffy) (e.g., bed, bread, baked potato) |
| fusa-fusa | tufty, tufted |
| puchi-puchi | small grains, little bits |
| putsu-putsu | pimply, bumpy (e.g., of a rash), knobbly |
| funya-funya | soft, limpness, flabby |
| punyu-punyu | soft and squishy, jellylike |
| puni-puni | squishy, pillowy, pudgy |
| puru-puru | jiggle, bounce, slightly trembling |
| beta-beta | sticky, cliched, hackneyed |
| becha-becha | sticky, prattling, messy (from mud, ink, etc.), chattering |
| beto-beto | sticky, gooey, prickly |
| moko-moko | lumpy, fuzzy, soft and fluffy |
| mochi-mochi | springy, sticky, glutinous |

### B. MAPPING ONOMATOPOEIAS BY ADJECTIVE AXES

In this study, we adopted Japanese Wikipedia as data and selected 42 words used by Sakamoto *et al.* as Japanese onomatopoeias for tactile perception (Table 1).

By mapping onomatopoeias onto the tactile adjective axis, we verified whether tactile onomatopoeias generated from the corpus have tactile information. We visualize the semantic distribution of tactile words in language space using NLP, where each word is expressed as an embedding and the cosine distance of these vectors is calculated. Skip-gram-type NLP algorithms use neighboring words as supervised data, thus eliminating the need for human labeling. Since words with similar meanings often appear in similar contexts, the corresponding embeddings will also be similar, and the semantic closeness between words can be calculated by calculating the cosine distance of the embeddings [12]. Using this property, we can visualize the relationship between tactile onomatopoeias without human labeling (i.e., tactile subject experiments). Six adjectives, viz., "hard," "soft," "wet," "dry," "rough," and "smooth," are included in the NLP and used to form the three axes (lines drawn between two points), "hard" and "soft" (hardness), "wet" and "dry" (wetness), and "rough" and "smooth" (roughness). This automatically creates onomatopoeic maps with respect to the adjective axes, which represent physical properties.

The vectors that represent an adjective pair can be expressed as follows:

$$\boldsymbol{a_1} = (a_{1_1}, a_{1_2}, \cdots, a_{1_n}),\ \boldsymbol{a_2} = (a_{2_1}, a_{2_2}, \cdots, a_{2_n}), \quad (1)$$

where $\boldsymbol{a_1}$ and $\boldsymbol{a_2}$ are the vectors representing an adjective pair, and $n$ is the number of dimensions of the hidden layer that is represented by the NLP embedding. We used 300 and 768 as the $n$ values for the FastText and BERT models, respectively. These are the recommended values for these two models. The unit vector $\boldsymbol{u}$ of the axes formed by the adjective pair can be expressed as follows:

$$\boldsymbol{u} = \frac{\boldsymbol{a_1} - \boldsymbol{a_2}}{\|\boldsymbol{a_1} - \boldsymbol{a_2}\|}. \quad (2)$$

By mapping the word group $\boldsymbol{X}$ of the onomatopoeia with this unit vector, we can obtain a distributed representation mapped to the adjective axes.

$$\boldsymbol{X_{after}} = \boldsymbol{X} \cdot \boldsymbol{u} \quad (3)$$

Since the obtained distributed representation $\boldsymbol{x_{after}}$ is projected against the adjective axis, we consider that the distance between the adjective at both ends and the onomatopoeia expresses the material properties. For example, "tsuru-tsuru" is a word that means slippery, sleek, slick and smooth, but in this case, the cosine distance from smooth is very short, and the distance from rough is large. Thus, "tsuru-tsuru" can be regarded as a physical property with very low roughness.

## III. EXPERIMENTS

### A. OBTAINING DISTRIBUTED REPRESENTATIONS OF ONOMATOPOEIAS

Since Japanese onomatopoeia is diverse and each word has a distinct meaning, the onomatopoeias selected in Table 1 were embedded using the FastText and BERT pre-training models to obtain an embedded representation of the word. The details of the pre-training models for each model are described below.

#### 1) FastText

We used a publicly available pre-training FastText model in Japanese[1]. This model uses MeCab, a morphological analysis engine, and ipadic, dictionary data to separate Wikipedia articles into sub-words, 32,000 vocabulary, 300 dimensions of hidden layers, 20 processes per batch, and 10 epochs of training. The other parameters are the default settings of FastText.

#### 2) BERT

We used a publicly available pre-training BERT model in Japanese[2]. This model uses 12 layers of transformers, 768 dimensions of hidden layers, 12 heads of self-Attention, 512 tokens per process, 256 processes per batch, and 20 epochs of training. The morphological analysis engine, dictionary data and vocabulary are the same as FastText.

### B. EXPERIMENT A: VISUALIZING THE EMBEDDING OF ONOMATOPOEIAS WITH ADJECTIVES

We plotted the 42 onomatopoeias on each adjective axis, the straight line formed by the adjective pairs. Since FastText yielded better results, only the FastText results are shown.

Fig. 2 shows a onomatopoeia map with hardness and wetness axes, which were obtained from the FastText model. The horizontal axis represents hardness; the further to the right, the closer to the word "hard" (i.e., stiff), and the further to the left, the closer to the word "soft". The vertical axis represents the wetness axis; the higher up you go, the closer you are to the word "wet," and the lower down you go, the closer you are to the word "dry". Fig. 3 shows the roughness and hardness, while Fig. 4 shows the wetness and roughness.

### C. EXPERIMENT B: EMBEDDED ONOMATOPOEIA

In order to quantify the extent to which the distributed representation of an onomatopoeia mapped onto these adjective axes reflects the material properties, we conducted a user study[3]. For this study, we recruited 20 participants (17 males and 3 females, aged 22–26). It is important to note that the purpose of this study was to determine quantitatively how close the distribution obtained by this experiment was to actual cognition, and it did not affect the distribution of onomatopoeias obtained by the embedded representation. The participants were asked to describe the image of the 42 onomatopoeias learned in this study on the hardness axis "hard/soft," the wetness axis "wet/dry," and the roughness axis "roughness-smoothness." Then, they were asked to answer the questions in seven steps using the semantic differential (SD) method. No tactile sensation was given at this time, and the participants were asked to answer on the basis of their own perception and image (or memory). We averaged and normalized the evaluation values for each

[1]https://github.com/Hironsan/awesome-embedding-models
[2]https://github.com/cl-tohoku/bert-japanese
[3]Ethical approval for this study was obtained from Keio University Faculty of Science and Technology Ethics Committee Number 2020-32.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2021.3130277, IEEE Access



onomatopoeia so that they would take values between 1 and -1. We used the obtained data as the correct answer data, and the degree of agreement with the distributed representation of the onomatopoeia was calculated. This time the data were normalized so that to they would fall in the range from 1 to -1. Thus, we can define the degree of agreement for each onomatopoeia as follows:

$$P(n) = \frac{1}{2}(2 - \|x_{SD_n} - x_{embedding_n} + 1\|), \tag{4}$$

where $P(n)$ is the agreement of the $n$th onomatopoeia, $x_{SD_n}$ is the averaged and normalized evaluation value for the $n$th onomatopoeia, and $x_{embedding_n}$ is the normalized distributed representation of the $n$th onomatopoeia mapped onto each adjective axis.

Table 2 compares the tactile representation with the participant answers; the values represent the average values of the degree of agreement ($P(n)$) for each axis. BERT's agreement was 61% on the hardness axis, 65% on the wetness axis, and 73% on the roughness axis, all lower than FastText's 72%, 71%, and 75%. This is likely because BERT learns word variants that take into account context and sentence structure. For example, when outputting the sentence "You like [mask] bread," it estimates whether it is "hard" or "soft" by inferring from the context before and after [14]. However, in the proposed method, no sentences were used as input, and words were input, so that even if the words had no context and were opposite in meaning, they may have been judged to be close. On the other hand, since FastText generates a unique distributed representation for each word, it can judge words based on their simple meanings, and the score is considered to be high. Thus, BERT succeeded in extracting tactile information from the distributed representation of natural language processing.

**TABLE 2. Comparison of tactile representation derived by embeddings of onomatopoeia using NLP models with human cognition.**

| Adjective Axis | FastText Score | BERT Score |
|---|---|---|
| hardness | 0.72 | 0.61 |
| wetness | 0.71 | 0.65 |
| roughness | 0.75 | 0.73 |
| total | 0.73 | 0.66 |

## IV. RESULTS AND DISCUSSION

### A. EMBEDDED ONOMATOPOEIAS AND HUMAN COGNITION

From the FastText variance representation, "gotsu-gotsu" and "gasa-gasa" were shown as "hard" words, which agreed with the tactile phase diagram with 42 onomatopoeic words presented by Sakamoto *et al.* [9]. Other good agreements include "beta-beta" and "necho-necho" as "soft" words, "toro-toro" as "wet", "jyari-jyari" as "dry", "zara-zara" as "rough", and "tsuru-tsuru" as "smooth". Thus, it was verified that NLP can

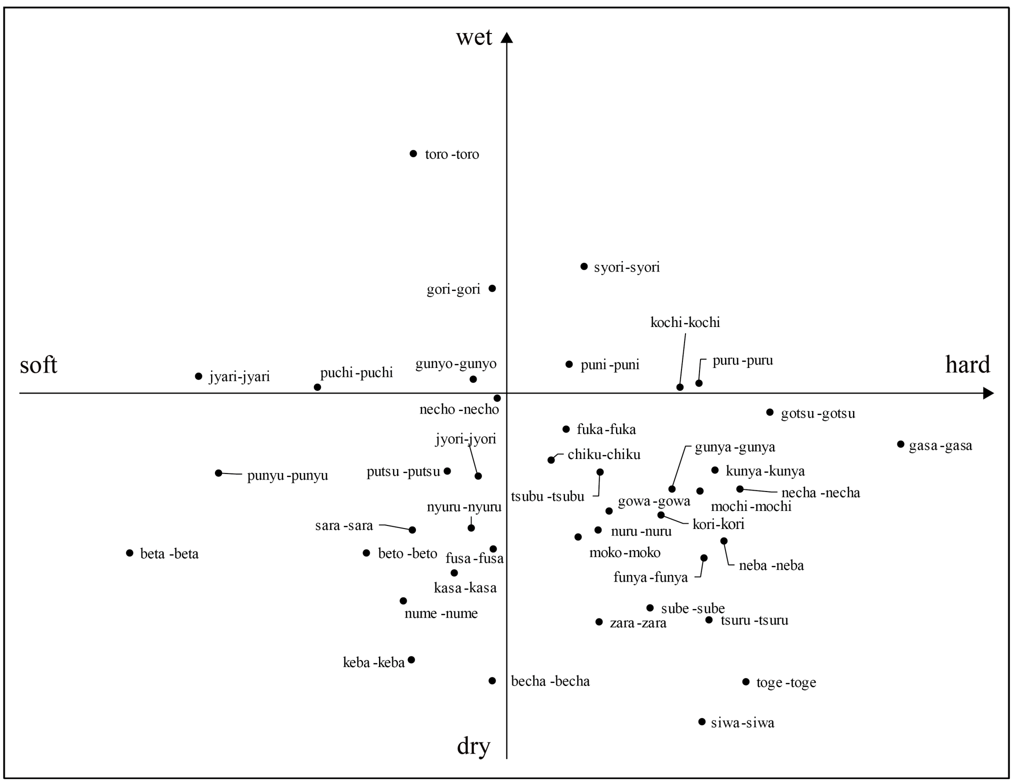


**FIGURE 2. Onomatopoeia group mapping with the hardness and wetness axes of the distributed representation obtained from the FastText model. The horizontal and vertical axes represent hardness and wetness, respectively.**

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2021.3130277, IEEE Access



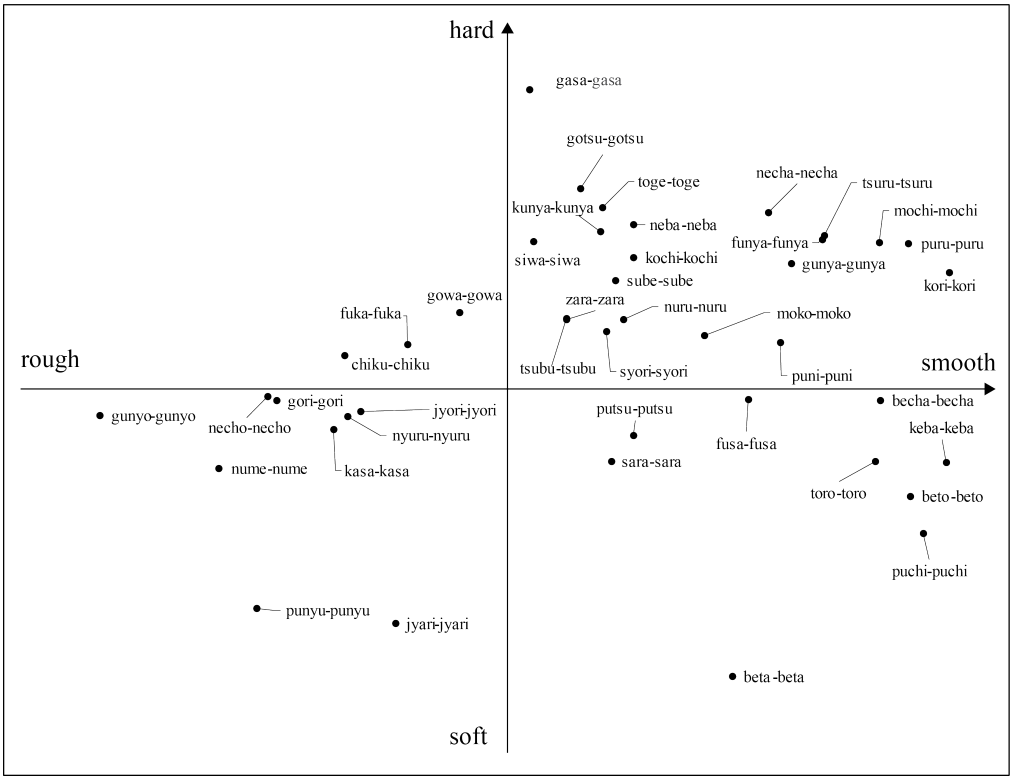


**FIGURE 3.** **Onomatopoeia group mapping with the roughness and hardness axes of the distributed representation obtained from the FastText model. The horizontal and vertical axes represent roughness and hardness, respectively.**

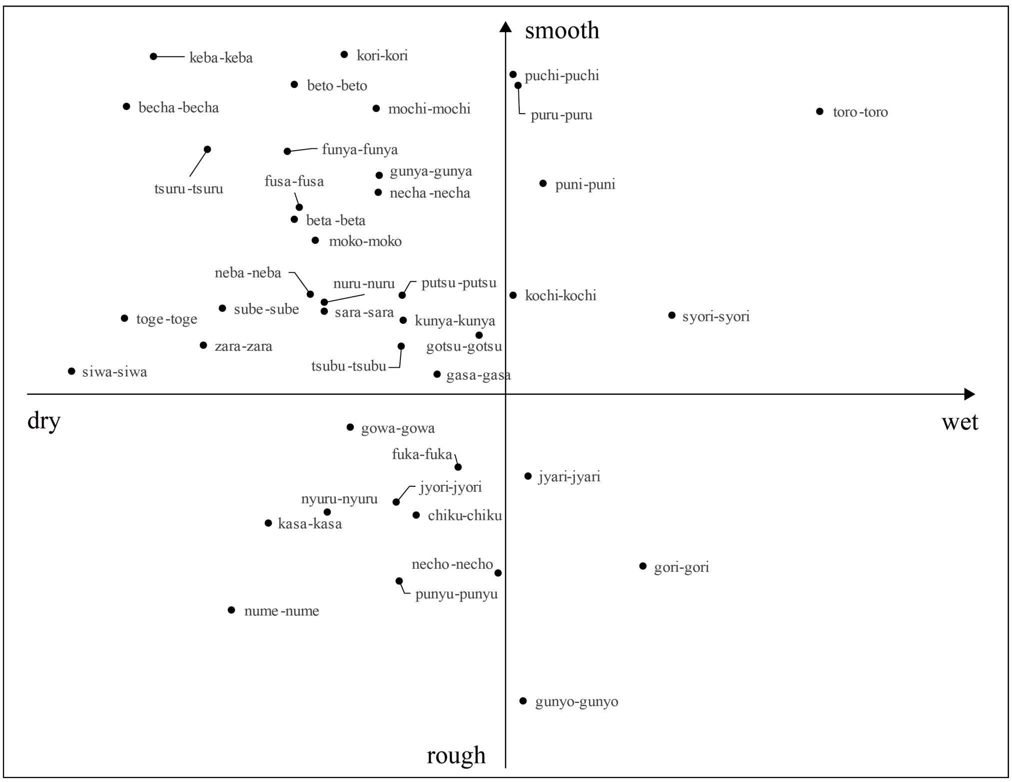


**FIGURE 4.** **Onomatopoeia group mapping with the wetness and roughness axes of the distributed representation obtained from the FastText model. The horizontal and vertical axes represent wetness and roughness, respectively.**

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2021.3130277, IEEE Access



extract tactile information as successfully as previous studies could from subject experiments.

On the other hand, even though "nyuru-nyuru", "fusa-fusa", and "kunya-kunya" are "soft" words, FastText mapped these words close to the origin of all the axes. In NLP, these words were considered not to involve tactile information.

We can conclude that NLP can successfully extract and map onomatopoeic words that contain tactile information.

The main difference between this map and the onomatopoeia map of Sakamoto et al. lies in the data set used to create this map. While Sakamoto et al. used data from 10 subjects who touched actual materials to create this map, this study automatically created it from the corpus of Wikipedia. The present study can be said to be a method that solves the difficulty of data collection in tactile research since the same tendency was obtained without preparing actual materials or subjects. In addition, by using a large data set such as wikipedia, it is possible to generate data on the average tactile cognition of Japanese people rather than individual tactile information. This data is expected to become a standard for judging cognitive differences between cultures and thresholds in tactile cognition between individuals.

### B. AFFECTIVE TACTILE DIMENSIONS: FACTOR ANALYSIS OF EMBEDDINGS OF ONOMATOPOEIA

Principal component analysis (PCA) of the embeddings of onomatopoeia can visualize the factors that humans consider important in tactile perception, which presumably corresponds to the affective tactile dimensions. We performed PCA on the embeddings of onomatopoeia. To clarify the difference between the principal component axes obtained by PCA and the conventional explanatory variables of tactile perception, we simultaneously reduced the dimensions of the six adjectives corresponding to the three elements of tactile perception. If each embedding of onomatopoeia is $\boldsymbol{x_o}$ and each adjective embedding is $\boldsymbol{x_a}$, then the vector of the data matrix of the onomatopoeias alone, $\boldsymbol{X_o}$, and the vector of the data matrix including the adjectives, $\boldsymbol{X_{o+a}}$, can be represented by two vectors as follows:

$$\boldsymbol{X_o} = \begin{pmatrix} \boldsymbol{x_{o_{11}}} & \boldsymbol{x_{o_{12}}} & \cdots & \boldsymbol{x_{o_{1n}}} \\ \boldsymbol{x_{o_{21}}} & \boldsymbol{x_{o_{22}}} & \cdots & \boldsymbol{x_{o_{2n}}} \\ \vdots & \vdots & \ddots & \vdots \\ \boldsymbol{x_{o_{m1}}} & \boldsymbol{x_{o_{m2}}} & \cdots & \boldsymbol{x_{o_{mn}}} \end{pmatrix}, \tag{5}$$

$$\boldsymbol{X_{o+a}} = \begin{pmatrix} \boldsymbol{x_{a_{11}}} & \boldsymbol{x_{a_{12}}} & \cdots & \boldsymbol{x_{a_{1n}}} \\ \boldsymbol{x_{a_{21}}} & \boldsymbol{x_{a_{22}}} & \cdots & \boldsymbol{x_{a_{2n}}} \\ \vdots & \vdots & \ddots & \vdots \\ \boldsymbol{x_{a_{l1}}} & \boldsymbol{x_{a_{l2}}} & \cdots & \boldsymbol{x_{a_{ln}}} \\ \boldsymbol{x_{o_{11}}} & \boldsymbol{x_{o_{12}}} & \cdots & \boldsymbol{x_{o_{1n}}} \\ \boldsymbol{x_{o_{21}}} & \boldsymbol{x_{o_{22}}} & \cdots & \boldsymbol{x_{o_{2n}}} \\ \vdots & \vdots & \ddots & \vdots \\ \boldsymbol{x_{o_{m1}}} & \boldsymbol{x_{o_{m2}}} & \cdots & \boldsymbol{x_{o_{mn}}} \end{pmatrix}, \tag{6}$$

where $n$ is the number of dimensions of the embeddings, $m$ is the number of onomatopoeia, and $l$ is the number of adjectives.

We performed PCA on the distributed representation of onomatopoeia and selected the principal component axes so that the variance after projection would be maximized. In this case, we calculated the covariance using only the distributed representation group of the onomatopoeia, not including the adjective words. This is because FastText and BERT learn not only the meanings of words, but also their relationships (i.e., morphological information. Japanese onomatopoeias are composed of repeated units of 2 or 3 Japanese syllables, such as zara-zara), so if there are two types of words, an adjective and an onomatopoeia, they will select axes that expand the variance morphologically. The data covariance matrix $S$ is defined as follows:

$$S = \frac{1}{m}\sum_{i=1}^{m}(\boldsymbol{x_{o_j}} - \tilde{\boldsymbol{x_o}})(\boldsymbol{x_{o_j}} - \tilde{\boldsymbol{x_o}})^T. \tag{7}$$

Assuming that the variance after projection is projected onto the first principal component axis $\boldsymbol{u_1}$ to be thought, it can be set as $\boldsymbol{u_1}^T S \boldsymbol{u_1}$. Considering that the variance after this projection is maximized, the following equation is obtained by adding the Lagrange multiplier $\lambda_1$:

$$S\boldsymbol{u_1} = \lambda_1 \boldsymbol{u_1}. \tag{8}$$

From this equation, we can know that the first principal component axis is the eigenvector of the covariance matrix $S$. By obtaining the eigenvector $\boldsymbol{u_1}$ from this equation, we can obtain the post-projection variance $\boldsymbol{u_1}^T S \boldsymbol{u_1}$ of the first principal component axis, which is obtained in the general principal component analysis.

In the proposed method, we can obtain a new covariance $S_{o+a}$ including the adjectives to visualize it by obtaining the variance after projection, including the adjective axis.

$$S_{o+a} = \tfrac{1}{m+l}\textstyle\sum_{i=1}^{m+l}(\boldsymbol{x_{o+a_j}} - \tilde{\boldsymbol{x_{o+a}}})(\boldsymbol{x_{o+a_j}} - \tilde{\boldsymbol{x_{o+a}}})^T. \tag{9}$$

By projecting this covariance onto the eigenvectors of the first principal component axis, we can obtain the principal component variance $\boldsymbol{u_1}^T S_{o+a} \boldsymbol{u_1}$ of the onomatopoeia with adjectives.

Fig. 5 shows the cumulative number of principal component axes and the cumulative contribution ratio. It was found that 16 dimensions were necessary to express a contribution rate of 80% or more. Table 3 shows the variance of the top three axes sorted in descending order, viz., the kunya-kunya/gotsu-gotsu, punyu-punyu/gunya-gunya, and punyu-punyu/gasa-gasa axes. In this method, eigenvectors are obtained for axes where the variance of only onomatopoeias is maximized. The adjectives are mapped over the axes in Table 3. Interestingly, the distance between the paired adjectives is small. In other words, physical characteristics are not considered to be the dominant factor in these major axes.

We confirmed this quantitatively by deducing how much variance (distance) each adjective pair has on each principal component axis. Table 4 shows a quantitative representation of the relationship between the adjective pairs and the 16

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2021.3130277, IEEE Access



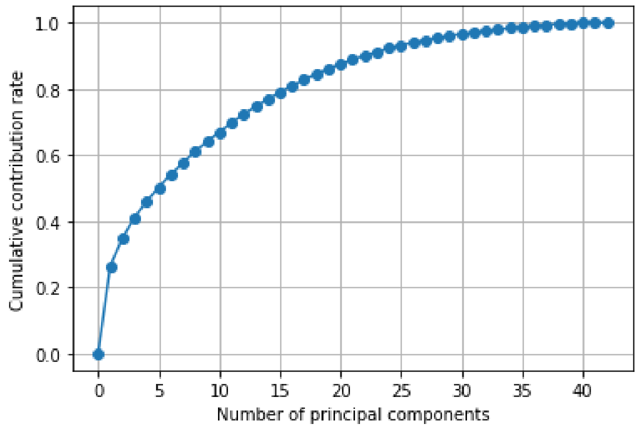


**FIGURE 5. PCA results of the embeddings of onomatopoeias generated by FastText.**

**TABLE 3. The top three principal components of the embeddings of onomatopoeias generated by FastText by principal component analysis.**

| PC1 | | PC2 | | PC3 | |
|---|---|---|---|---|---|
| kunya-kunya | -17.37 | punyu-punyu | -12.19 | punyu-punyu | -13.87 |
| funya-funya | -17.12 | putsu-putsu | -9.49 | putsu-putsu | -12.76 |
| gunya-gunya | -16.62 | kochi-kochi | -8.06 | puchi-puchi | -11.93 |
| becha-becha | -15.54 | gori-gori | -6.76 | gunya-gunya | -5.41 |
| gunyo-gunyo | -14.03 | jyori-jyori | -6.59 | jyari-jyari | -4.86 |
| necha-necha | -13.97 | gunyo-gunyo | -5.83 | puru-puru | -3.16 |
| jyori-jyori | -10.80 | necho-necho | -5.50 | puni-puni | -2.61 |
| necho-necho | -10.62 | puchi-puchi | -5.27 | nyuru-nyuru | -1.82 |
| punyu-punyu | -10.44 | fuka-fuka | -5.14 | becha-becha | -1.79 |
| gori-gori | -9.53 | gasa-gasa | -4.80 | kunya-kunya | -1.43 |
| kochi-kochi | -8.72 | moko-moko | -4.56 | kori-kori | -1.15 |
| putsu-putsu | -3.67 | keba-keba | -4.11 | tsubu-tsubu | -1.11 |
| nume-nume | -3.63 | kasa-kasa | -3.91 | beta-beta | -0.85 |
| keba-keba | -2.94 | beto-beto | -3.42 | tsuru-tsuru | -0.81 |
| beto-beto | -2.22 | syori-syori | -2.88 | sara-sara | -0.65 |
| syori-syori | -2.04 | nume-nume | -1.95 | toro-toro | -0.64 |
| chiku-chiku | -2.00 | nyuru-nyuru | -1.70 | neba-neba | -0.59 |
| gasa-gasa | -0.31 | chiku-chiku | -1.49 | funya-funya | -0.45 |
| moko-moko | -0.10 | toro-toro | -0.76 | beto-beto | -0.42 |
| nyuru-nyuru | 0.32 | jyari-jyari | -0.64 | smooth | -0.40 |
| fuka-fuka | 0.71 | mochi-mochi | -0.23 | soft | -0.30 |
| kori-kori | 1.85 | siwa-siwa | -0.18 | wet | -0.25 |
| jyari-jyari | 2.04 | beta-beta | 0.04 | sube-sube | -0.16 |
| beta-beta | 2.09 | kori-kori | 0.46 | nuru-nuru | -0.10 |
| toro-toro | 2.53 | puni-puni | 0.65 | hard | -0.05 |
| siwa-siwa | 3.55 | toge-toge | 1.53 | gotsu-gotsu | 0.04 |
| puchi-puchi | 3.66 | puru-puru | 1.69 | rough | 0.157 |
| kasa-kasa | 3.88 | sara-sara | 1.85 | dry | 0.23 |
| mochi-mochi | 4.63 | wet | 1.93 | zara-zara | 0.46 |
| puni-puni | 5.57 | neba-neba | 2.37 | gowa-gowa | 0.48 |
| puru-puru | 5.81 | tsubu-tsubu | 2.38 | fusa-fusa | 0.61 |
| sara-sara | 6.53 | rough | 2.42 | toge-toge | 1.00 |
| wet | 6.58 | fusa-fusa | 2.69 | gunyo-gunyo | 1.65 |
| fusa-fusa | 6.77 | gowa-gowa | 2.80 | mochi-mochi | 1.82 |
| tsuru-tsuru | 6.86 | soft | 2.81 | keba-keba | 1.84 |
| toge-toge | 6.86 | dry | 2.84 | necha-necha | 2.12 |
| nuru-nuru | 6.98 | tsuru-tsuru | 3.03 | chiku-chiku | 2.71 |
| neba-neba | 7.10 | gotsu-gotsu | 3.07 | moko-moko | 2.91 |
| sube-sube | 7.10 | smooth | 3.10 | syori-syori | 3.15 |
| dry | 7.25 | nuru-nuru | 3.13 | nume-nume | 3.59 |
| tsubu-tsubu | 7.57 | hard | 3.24 | kochi-kochi | 4.09 |
| hard | 7.75 | zara-zara | 3.71 | siwa-siwa | 4.58 |
| soft | 7.79 | sube-sube | 3.76 | necho-necho | 5.20 |
| rough | 7.80 | becha-becha | 7.56 | fuka-fuka | 5.46 |
| gowa-gowa | 7.88 | funya-funya | 9.10 | gori-gori | 5.62 |
| zara-zara | 7.91 | kunya-kunya | 9.38 | kasa-kasa | 5.86 |
| smooth | 8.10 | necha-necha | 9.71 | jyori-jyori | 5.94 |
| gotsu-gotsu | 8.22 | gunya-gunya | 10.26 | gasa-gasa | 8.01 |

**TABLE 4.** Relationship between each principal component axis and physical properties.

| PC Index | Explained Variance ratio | Hardness Length | Wetness Length | Roughness Length |
|---|---|---|---|---|
| 1 | 20.73% | 0.14% | 2.61% | 1.20% |
| 2 | 8.35% | 1.91% | 4.06% | 3.00% |
| 3 | 6.58% | 1.15% | 2.17% | 2.55% |
| 4 | 5.17% | 1.93% | 4.28% | 1.72% |
| 5 | 4.59% | 0.21% | 1.03% | 2.74% |
| 6 | 4.01% | 4.28% | 8.19% | 1.20% |
| 7 | 3.86% | 1.56% | 9.36% | 0.47% |
| 8 | 3.71% | 4.23% | 0.20% | 0.43% |
| 9 | 3.41% | 4.48% | 9.35% | 3.77% |
| 10 | 3.18% | 1.70% | 4.41% | 5.57% |
| 11 | 3.04% | 0.12% | 4.25% | 3.20% |
| 12 | 2.80% | 0.88% | 2.63% | 1.08% |
| 13 | 2.56% | 1.55% | 0.29% | 0.65% |
| 14 | 2.39% | 2.87% | 4.17% | 1.33% |
| 15 | 2.21% | 0.60% | 6.06% | 0.67% |
| 16 | 2.17% | 1.23% | 0.13% | 5.82% |

principal component axes by normalizing the length of the adjective pairs on each principal component axis by the length of each principal component axis. Interestingly, in all of the 16 principal component axes, the distance between the adjective pairs is smaller than 10%. Uchida *et al.* [22] have shown that onomatopoeias in Japanese contain human affective meaning. Ramachandran *et al.* [11] and Etzi *et al.* [6] have shown that tactile sensation contains affective information. Therefore, we can say that onomatopoeias are not used to represent physical properties but affective information.

Since the onomatopoeic embeddings obtained from the FastText model were shown to have physical properties in Experiment A, the respective principal component axes were visualized in the results of Experiment A in order to visualize the correspondence between the principal component axes and the physical properties. In PC1, for example, we mapped each adjective axis pseudo-physically by drawing an axis that passed through the two extreme words, "gowa-gowa" and "kunya-kunya" Fig. 6, 7 and 8.

In Fig. 6, the norm of PC1 is very small and can be regarded as an axis different from the hardness and roughness axes, which are physical properties. The norm of PC2 and PC3 is large relative to the overall distribution, but the slope is almost parallel to the hardness axis, so they can be considered as axes with roughness properties and almost no wetness properties.

In Fig. 7, the norm of PC1 is very small and can be considered as an axis different from the hardness and roughness axes, which are physical properties. The norm of PC2 and PC3 is large compared to the overall distribution, but the slope is almost parallel to the roughness axis, so it can be considered as an axis with roughness properties and little wetness properties.

In Fig. 8, the norm of PC1 is very small and it is considered to be a different axis from the hardness and roughness axes, which are physical properties. The norm of PC2 and PC3 is large in relation to the total distribution, but the slope is from the third quadrant to the first quadrant, so they can

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2021.3130277, IEEE Access



be considered as axes having both roughness and hardness properties.

Figs. 6–8 show that PC1 has properties very different from the existing physical axes, and PC2 and PC3 have properties of the hardness and roughness axes. Interestingly, PC1 has different properties from the existing physical axes even though it has a much higher contribution of 28% to the total variance than Fig. 1. This is because the existing axes ignore this 28% in the total variance, and many tactile experiments using the existing axes are considered to have information loss of this 28%.

Since PC2 and PC3 were shown to have physical properties of hardness and roughness, we visualized the physical axes in each principal component axis in Fig. 9, 10, and 11.

Fig. 9 shows PC1 on the horizontal axis and PC2 on the vertical axis. If we focus on the norm of the physical axis, we can see all three are very small compared to the total variance. Fig. 10 shows PC2 on the horizontal axis and PC3 on the vertical axis, and the norm of the physical axis is also very small compared to the total variance. Fig. 11 shows PC3 on the horizontal axis and PC1 on the vertical axis, and the norm of the physical axis is also very small compared to the total variance.

In Table 3, we mentioned the norm for each axis, similarly, for PC1 and PC2, PC2 and PC3, and PC3 and PC1, the norm for each physical axis was shown to be small relative to the overall variance. This indicates that PC2 and PC3 have physical components of hardness and roughness axes, but they are not dominant in PC2 and PC3, and can be regarded as different axes that affect the hardness and roughness axes.

Etzi *et al.* [6]. focused on the tactile adjectives rough and smooth, and showed that roughly correlated with negative words such as “dim”, “dark”, “loud” and “heavy” while smooth correlated with positive words such as “bright”, “light”, “quiet” and “lightweight”. The PC2 and PC3 obtained by this principal component analysis were found to have the characteristics of rough-smooth, which is consistent with their claim. These results can be said that our generated vectors might have not only physical properties but also affective properties.

We assume that the first principal component axis obtained in the proposed method (for the sake of convenience, we named it the “kunya-gotsu axis” since the pairs are “kunya-kunya” and “gotsu-gotsu”) can be set as a factor in the SD method, instead of adjective pairs that represent physical properties, to better express the sense of touch. For example, in the kunya-gotsu axis, the two ends of the SD method should be set to “kunya-kunya” and “gotsu-gotsu”. By using the 16 principal axes obtained this time and evaluating them with the SD method, we will be able to evaluate tactile sensation more variedly. This is because if we wanted to incorporate existing axes as before, we could use this model to re-map them into hyperspace and incorporate them into variables that can be accounted for by the methods of Exper-

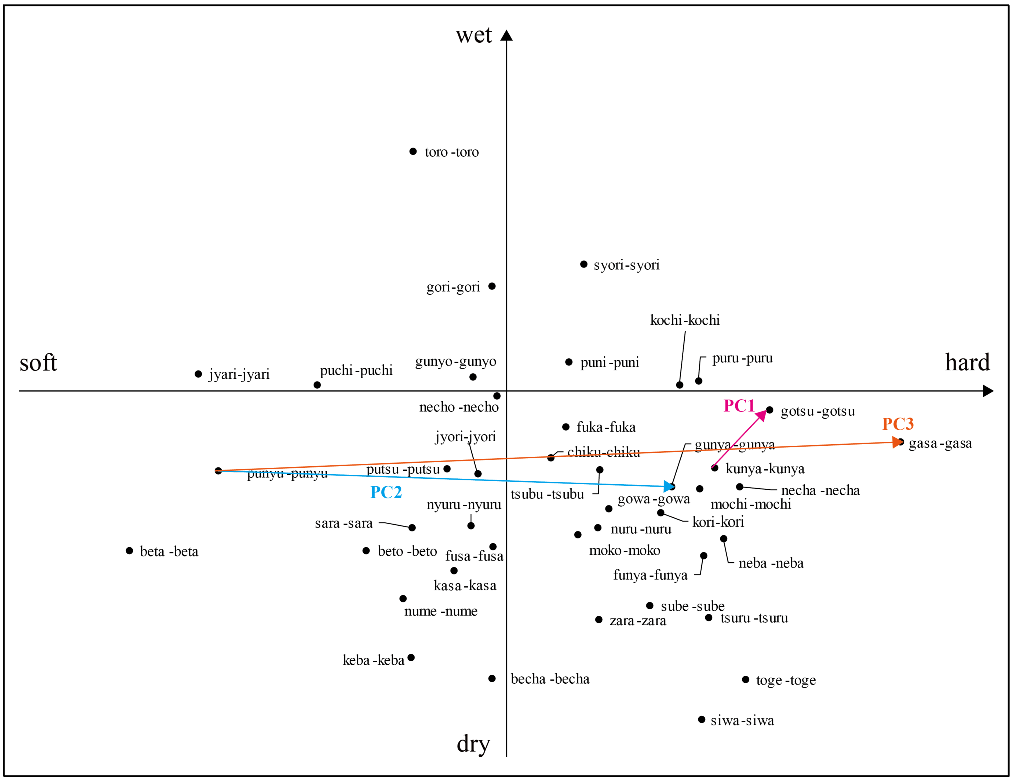


**FIGURE 6.** **Onomatopoeia group mapping of the hardness and wetness axes of the distributed representation obtained from the FastText model and calculated principal component axes. The magenta, cyan and orange axes represent the PC1 (”gowa-gowa“ and ”kunya-kunya“), PC2 (”gunya-gunya“ and ”punyu-punyu“) and PC3 (”gasa-gasa“ and ”punyu-punyu“) axes, respectively.**

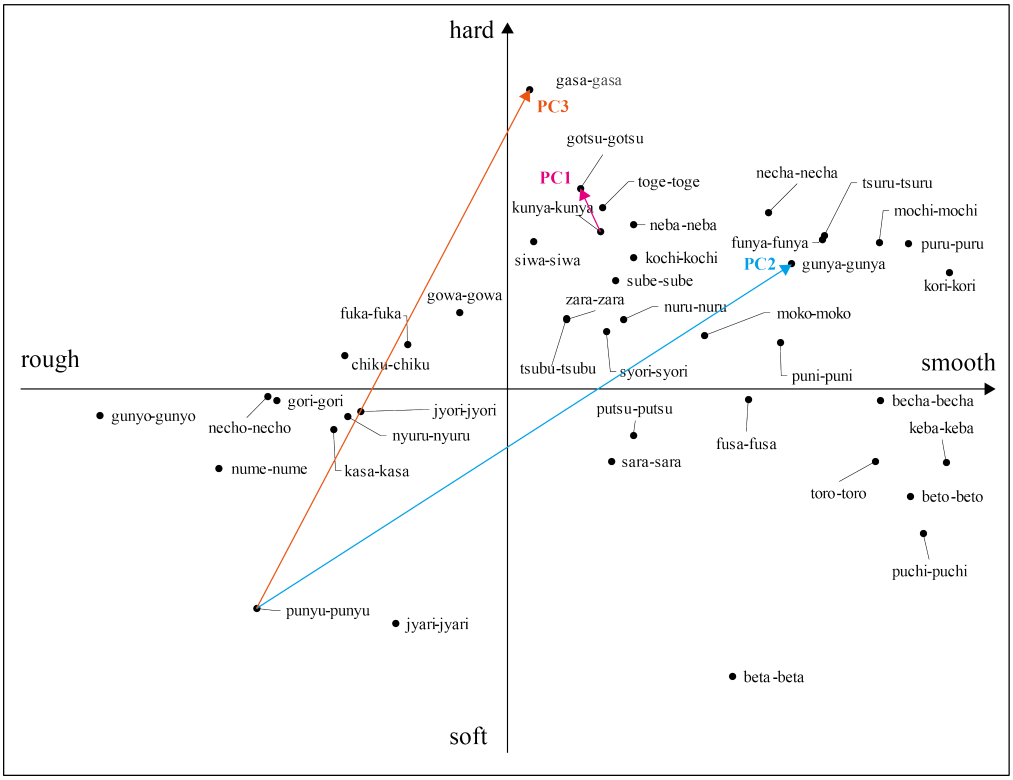


**FIGURE 7. Onomatopoeia group mapping of the roughness and hardness axes of the distributed representation obtained from the FastText model and calculated principal component axes. The magenta, cyan and orange axes represent the PC1 ("gowa-gowa" and "kunya-kunya"), PC2 ("gunya-gunya" and "punyu-punyu") and PC3 ("gasa-gasa" and "punyu-punyu") axes, respectively.**

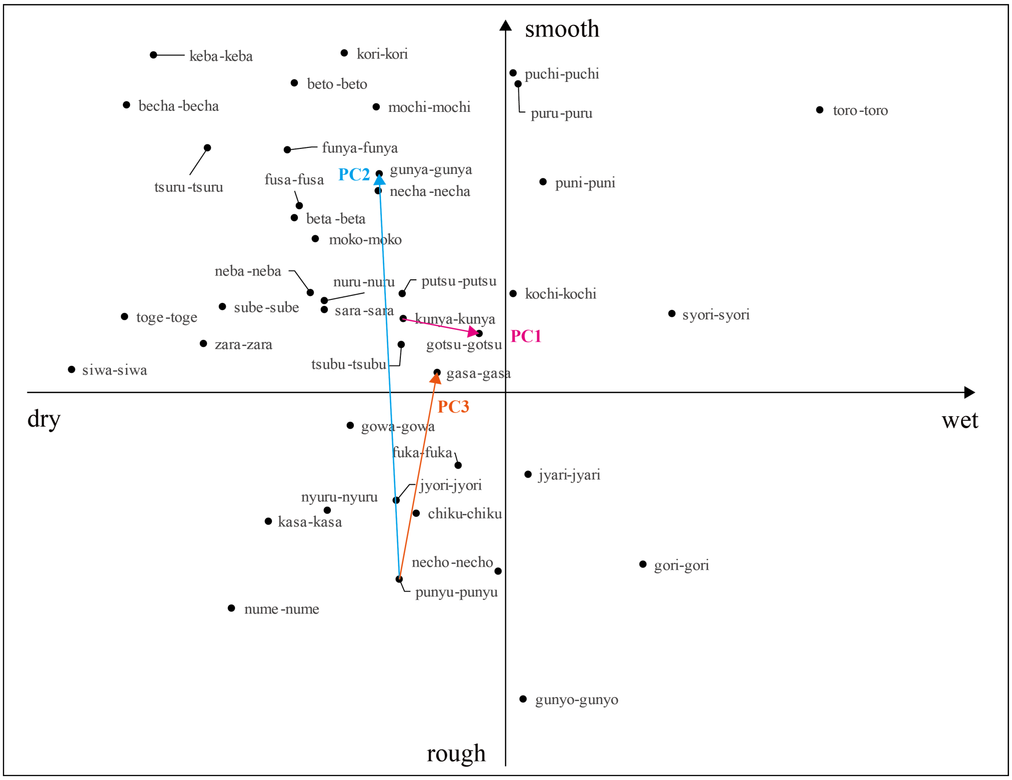


**FIGURE 8. Onomatopoeia group mapping of the wetness and roughness axes of the distributed representation obtained from the FastText model and calculated principal component axes. The magenta, cyan and orange axes represent the PC1 ("gowa-gowa" and "kunya-kunya"), PC2 ("gunya-gunya" and "punyu-punyu") and PC3 ("gasa-gasa" and "punyu-punyu") axes, respectively.**

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2021.3130277, IEEE Access



iment A (i.e., the three axes with textural properties).

### C. APPLICATION TO MULTIPLE LANGUAGES AND LINGUISTIC EXPRESSIONS OTHER THAN ONOMATOPOEIA

The present NLP-based method for tactile evaluation using word embedding can also be applied to languages other than Japanese. In the proposed method, we used onomatopoeia as the main tactile embedding, but it should be possible to extract tactile information by embedding adjectives and representative nouns as well. Our results suggest that although BERT's performance is inferior to FastText in directly visualizing word embedding, fine-tuning can greatly improve BERT's performance. Furthermore, by translating word embeddings that have learned the meaning of tactile sensation using a generic model such as BERT, it should be possible to visualize the differences in tactile sensation between multiple languages.

## V. CONCLUSION

In this paper, we proposed a new and effective approach in tactile research that uses natural language processing of archival text corpuses. This approach allows us to access a large amount of useful data, i.e., verbalized tactile words, without time-consuming perception experiments. We highlighted onomatopoeias, which are frequently used to represent tactile information in Japanese. They were successfully mapped with respect to physical properties, which agreed with the previous report based on an enquete-based survey. This verified the effectiveness of the NLP-based approach.

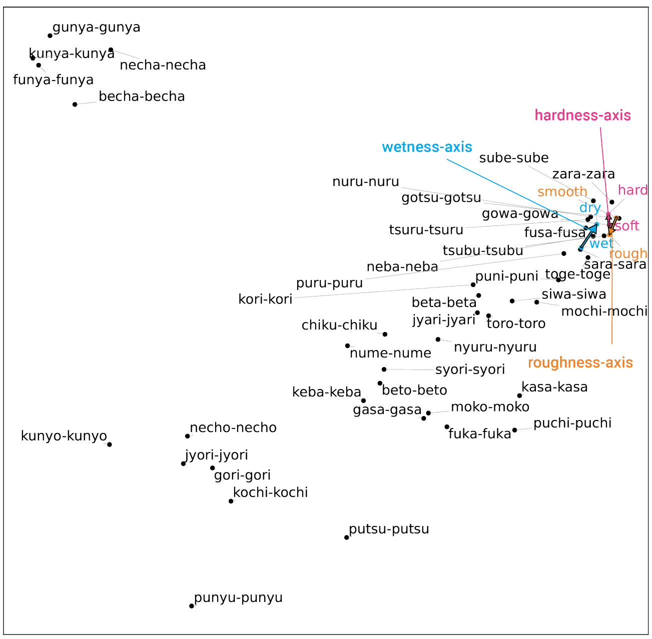


**FIGURE 9.** **Onomatopoeia group mapping of the PC1 and PC2 axes of the distributed representation obtained from the FastText model and calculated principal component axes. The magenta, the cyan, and the orange axes represent the hardness, wetness, and roughness axes, respectively.**

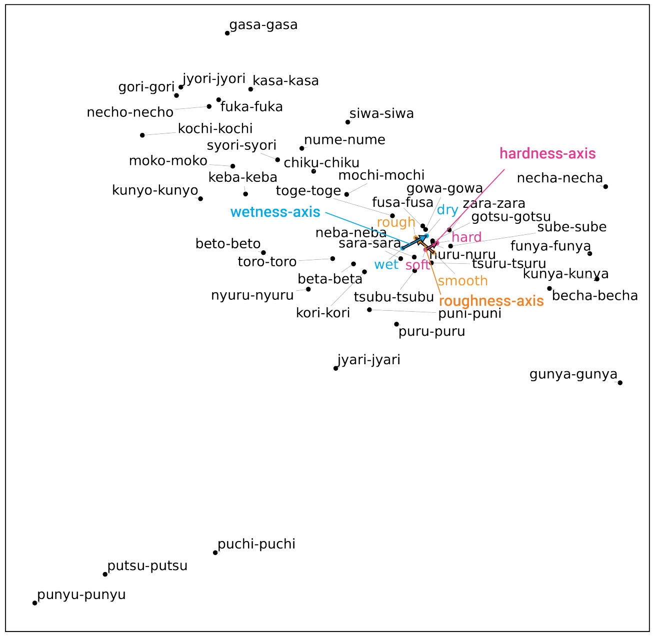


**FIGURE 10.** **Onomatopoeias group mapping the PC2 and PC3 axes of the distributed representation obtained from the FastText model and calculated principal component axes. The magenta, cyan, and orange axes represent the hardness, wetness, and roughness axes, respectively.**


**FIGURE 11.** **Onomatopoeias group mapping the PC3 and PC1 axes of the distributed representation obtained from the FastText model and calculated principal component axes. The magenta, cyan, and orange axes represent the hardness, wetness, and roughness axes, respectively.**

In addition, principal component analysis of the distributed representation suggested that onomatopoeias are used to represent affective tactile information. PCA successfully extracted affective tactile dimensions, including "kunya-gotsu", "punyu-gunya", and "punyu-gasa".

## REFERENCES

[1] S. Okamoto, H. Nagano, and Y. Yamada, "Psychophysical dimensions of tactile perception of textures," IEEE Transactions on Haptics, vol. 6, no. 1, pp. 81–93, 2013.
[2] S. Tachi, "Telexistence: Enabling humans to be virtually ubiquitous," IEEE Computer Graphics and Applications, vol. 36, no. 1, pp. 8–14, 2016.
[3] H. Shirado and T. Maeno, "Modeling of human texture perception for tactile displays and sensors," in Proceedings of the 1st Joint Eurohaptics Conference and Symposium on Haptic Interfaces for Virtual Environment and Teleoperator Systems, 2005, pp. 629–630.
[4] M. Kawazoe, Y. Kosemura, and N. Miki, "Encoding and presentation of surface textures using a mechanotactile display," Sensors and Actuators A: Physical, vol. 261, pp. 30–39, 2017.
[5] S. Hirai and N. Miki, "A thermal tactile sensation display with controllable thermal conductivity," Micromachines, vol. 10, no. 6, p. 359, May 2019.
[6] R. Etzi, C. Spence, M. Zampini, and A. Gallace, "When sandpaper is 'kiki' and satin is 'bouba': an exploration of the associations between words, emotional states, and the tactile attributes of everyday materials," vol. 29, pp. 133–155. [Online]. Available: https://brill.com/view/journals/msr/29/1-3/article-p133_7.xml
[7] M. Miyazaki, S. Hidaka, M. Imai, K. Kantartzis, H. Okada, and S. Kita, "The facilitatory role of sound symbolism in infant word learning," in the Annual Meeting of the Cognitive Science Society,, 2013, pp. 3080–3085.
[8] M. Arata, M. Imai, J. Okuda, H. Okada, and T. Matsuda, "Gesture in language: How sound symbolic words are processed in the brain," in the Annual Meeting of the Cognitive Science Society,, 2010, pp. 1374–1379.
[9] M. Sakamoto, T. Tahara, and J. Watanabe, "A system to visualize individual variation in tactile perception using onomatopoeia map," vol. 21, no. 2, pp. 213–216.
[10] M. Sakamoto, "System to quantify the impression of sounds expressed by onomatopoeias," Acoustical Science and Technology, vol. 41, no. 1, pp. 229–232, 2020.
[11] V. S. Ramachandran and D. Brang, "Tactile-emotion synesthesia," vol. 14, no. 5, pp. 390–399.
[12] T. Mikolov, K. Chen, G. Corrado, and J. Dean, "Efficient estimation of word representations in vector space," in Proceeding of the 1st International Conference on Learning Representations, Workshops Track, 2013, p. 12 pages.
[13] P. Bojanowski, E. Grave, A. Joulin, and T. Mikolov, "Enriching word vectors with subword information," Transactions of the Association for Computational Linguistics, vol. 5, pp. 135–146, 2017.
[14] J. Devlin, M.-W. Chang, K. Lee, and K. Toutanova, "BERT: Pre-training of deep bidirectional transformers for language understanding," in Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, vol. 1, jun 2019, pp. 4171–4186.
[15] A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A. N. Gomez, L. Kaiser, and I. Polosukhin, "Attention is all you need," in Proceedings of the 31st International Conference on Neural Information Processing Systems, 2017, p. 6000–6010.
[16] T. B. Brown, B. Mann, N. Ryder, M. Subbiah, J. Kaplan, P. Dhariwal, A. Neelakantan, P. Shyam, G. Sastry, A. Askell, S. Agarwal, A. Herbert-Voss, G. Krueger, T. Henighan, R. Child, A. Ramesh, D. M. Ziegler, J. Wu, C. Winter, C. Hesse, M. Chen, E. Sigler, M. Litwin, S. Gray, B. Chess, J. Clark, C. Berner, S. McCandlish, A. Radford, I. Sutskever, and D. Amodei, "Language models are few-shot learners," CoRR, vol. abs/2005.14165, 2020. [Online]. Available: https://arxiv.org/abs/2005.14165
[17] A. Ramesh, M. Pavlov, G. Goh, S. Gray, C. Voss, A. Radford, M. Chen, and I. Sutskever, "Zero-shot text-to-image generation," CoRR, vol. abs/2102.12092, 2021. [Online]. Available: https://arxiv.org/abs/2102.12092
[18] Z. Yang, Z. Dai, Y. Yang, J. G. Carbonell, R. Salakhutdinov, and Q. V. Le, "Xlnet: Generalized autoregressive pretraining for language understanding," CoRR, vol. abs/1906.08237, 2019. [Online]. Available: http://arxiv.org/abs/1906.08237
[19] Y. Liu, M. Ott, N. Goyal, J. Du, M. Joshi, D. Chen, O. Levy, M. Lewis, L. Zettlemoyer, and V. Stoyanov, "Roberta: A robustly optimized BERT pretraining approach," CoRR, vol. abs/1907.11692, 2019. [Online]. Available: http://arxiv.org/abs/1907.11692
[20] Z. Lan, M. Chen, S. Goodman, K. Gimpel, P. Sharma, and R. Soricut, "Albert: A lite bert for self-supervised learning of language representations," in International Conference on Learning Representations, 2020. [Online]. Available: https://openreview.net/forum?id=H1eA7AEtvS
[21] K. Clark, M. Luong, Q. V. Le, and C. D. Manning, "ELECTRA: pre-training text encoders as discriminators rather than generators," CoRR, vol. abs/2003.10555, 2020. [Online]. Available: https://arxiv.org/abs/2003.10555
[22] Y. Uchida, K. Araki, and J. Yoneyama, "Classification of emotional onomatopoeias based on questionnaire surveys," in 2012 International Conference on Asian Language Processing, pp. 1–4.

TATSUHO NAGATOMO received the B.S. and M.S. degrees from Keio University, Japan, in 2017 and 2019, respectively. His research interests include tactile sensors, haptic displays, natural language processing, machine learning, and augmented reality.

TAKEFUMI HIRAKI (S'16-M'19) received the B.S., M.S., and Ph.D. degrees from the University of Tokyo, Japan, in 2014, 2016, and 2019, respectively. He was a Visiting Researcher with Microsoft Research, China, in 2019 and a JSPS Research Fellow at the Graduate School of Engineering Science, Osaka University, from 2019 to 2021. He is currently an Assistant Professor with the Faculty of Library, Information and Media Science, University of Tsukuba. His research interests include augmented reality, haptic displays, soft robotics, and human–computer interaction. He is a member of ACM.

HIROKI ISHIZUKA received the Ph.D. degree in integrated design engineering from Keio University in 2016. He is an Assistant Professor with the Graduate School of Engineering Science of Osaka University. His research interests include haptics and soft robotics.

NORIHISA MIKI received the Ph.D. degree in mechano-informatics from University of Tokyo in 2001. He developed the world's smallest drone using MEMS technology during his Ph.D. studies. Then, he worked on the MIT microengine project as a postdoc (2001-2003), and later as a research engineer (2003-2004). He joined the Department of Mechanical Engineering at Keio University in 2004 as an Assistant Professor and became a Full Professor in 2017. His research interests started with development of MEMS-based biomedical and human interface devices. Currently, he also explores the fields of medical engineering, neuroscience, and media arts using his innovative devices. He was a researcher of JST PRESTO (Information Environment and Humans) from 2010 to 2016 and Kanagawa Institute of Industrial Science and Technology (formerly, Kanagawa Academy of Science and Technology) from 2010 to present. He is a member of IEEE and JSME micro-nano science and technology division. He is a general chair of the 8th and 9th Symposium on Micro-Nano Science and Technology in 2017 and 2018 sponsored by JSME. He co-founded a healthcare startup LTaste Inc. in 2017.