

Yoko Nakano*

Similarity effect in morphological generalization: Using the volitional form elicited production task of Japanese verbs with *suru* ending

<https://doi.org/10.1515/jjl-2025-2007>

Abstract: Two major views have been proposed regarding the mechanism underlying morphological generalization: rule-based and analogy-based mechanisms. The analogy-based mechanism is sensitive to the similarity of nonce words to existing words, whereas the rule-based mechanism is not. This study investigated the production of volitional forms of Japanese verbs whose non-past forms end with *suru* by using an elicited production experiment with low- and high-frequency real verbs and novel verbs with varying degrees of similarity to existing verbs. The findings are as follows: The conjugational form of the largest verb group was preferred, and an increase in the conjugational form of a smaller verb group was found only in the completely similar condition. These results can be explained by the rule-based mechanism.

Keywords: elicited production task; Japanese verbs; linguistic generalization; volitional forms

1 Introduction

Humans possess the ability to generalize rules, constraints, and patterns to novel linguistic data, which can also be observed in the production of morphologically complex inflectional verb forms (Pinker 1999; Veríssimo and Clahsen 2014). Two major views have been proposed regarding the mechanism underlying morphological generalization – rule-based (Albright and Hayes 2003; Hale 2001) and analogy-based (the exemplar model: Nosofsky 1989; the connectionist model: Rumelhart et al. 1986; Rumelhart and McClelland 1986) mechanisms, and there is also an eclectic view of the two – the dual mechanism model (Clahsen 1999; Pinker 1999; Pinker and Prince 1988).

The rule-based mechanism assumes that conjugational forms are produced by applying a rule or an operation to a verb, for instance, the past-tense form of English

***Corresponding author: Yoko Nakano**, The School of Human Welfare Studies, Kwansei Gakuin University, 1-155 Uegahara Ichiban-cho, Nishinomiya, Hyogo 662-8501, Japan, E-mail: y-k.nakano@kwansei.ac.jp

regular verbs is formed by the rule of adding *ed* to the verb root, and even to a novel verb (e.g., *ploamph-ploamphed*) (Pinker 1999). Conversely, analogy-based mechanisms have assumed that conjugational forms are produced based on the similarity and frequency of existing verbs. That is, English past-tense forms of verbs are produced from a cluster of frequently overlapping items (Rumelhart and McClelland 1986). For instance, the past participle *splung* of the novel English verb *spling* can be formed analogically with the existing English verbs *ring-rung* and *sing-sung* (Bybee and Moder 1983). Bybee (1995) suggested that items with higher type frequency are more productive in forming conjugational forms. The dual mechanism model assumes that the past-tense forms of English regular verbs are produced by the rule-based mechanism, namely, by applying the default rule to add *ed* to the verb stem, where past-tense forms are not stored in memory, and those of irregular verbs are produced by the analogy-based mechanism (Pinker 1999). One major method used to investigate this issue is the elicited production task (Berko 1958). In the written version of the task, a text is presented to participants. It contains a certain form of a nonce word and a blank space, in which participants are expected to fill with the target conjugational form. Because the elicited production task asks participants to conjugate nonce words, if the rule-based mechanism works, the task will make participants generalize the rules and operations they use for existing words to nonce words. If the analogy-based mechanism works, the task will make them analogize the conjugational form of a nonce word with existing words they know. The mechanism that works in Japanese is unclear. Therefore, this study investigated the applicability of these mechanisms in Japanese by checking whether the similarity of novel verbs to existing verbs influences the production of volitional verb forms.

Some studies have utilized the similarity between real and novel verbs as experimental stimuli in elicited production tasks. Verissimo and Clahsen (2014) constructed novel verbs that were similar to real Portuguese verbs. Portuguese contains three verb classes. The first conjugation class of verbs is the largest group (3396 verbs). The second conjugation (380 verbs) and third conjugation (348 verbs) classes of verbs are approximately the same size, and they are fewer than the first conjugation class verbs. Their corresponding infinitive forms end with *-ar*, *-er*, and *-ir*, respectively. The first-person singular indicative forms end with *-o*, regardless of class. In an elicited production task, they presented a two-sentence text. The first sentence contained a first-person singular indicative form of a novel verb, and the second sentence contained a blank space, which participants were expected to fill with the infinitive form. The endings of infinitive forms helped distinguish the classes of verbs in the participants' responses. The results indicated a similarity effect in the first conjugation class but not in the novel verbs of the second and third conjugation classes. They argued that their results were compatible with the dual-mechanism model.

Michon and Nakipoğlu (2020) also conducted an elicited production experiment on Turkish aorist (a kind of present-tense form that expresses truth and customs in Turkish) by using nonce verbs with varying degrees of similarity to existing verbs at a phonemic level. Different degrees of similarity were created by manipulating the overlaps in the CVC-phonemic structure of stems between nonce and real Turkish verbs, for example, two Cs overlap, the initial C overlaps, V overlaps, and second C overlaps.¹ In the experiment, the nonce root was presented in the first sentence auditorily, and the second sentence with a blank space, which participants were expected to fill with the aorist, was visually presented; participants were instructed to complete the sentence orally. They found that participants preferred the most regular patterns, which can be best explained by rule-based generalization and the effect of type frequency; namely, items with higher type frequency are more productive, as more items share the same phonological overlaps, which are more applicable to other items, including nonce words (Bybee 1995). Uygun et al. (2023) conducted an elicited production experiment on aorist with Turkish monolingual and heritage speakers and found the same pattern of results for monolingual speakers. With these previous studies in mind, we conducted an elicited production experiment on volitional verb forms with varying degrees of similarity in another agglutinative language, Japanese.

2 Background

2.1 Japanese verbs

Japanese verbs are classified into three groups (Groups I, II, and III). Group-I verbs are referred to as consonant-stem verbs because their stems end with a consonant. For instance, the non-past form of a Group-I verb *kaku* ‘write’ constitutes the stem *kak* and non-past affix *u*. Its volitional form is formed by attaching an affix *ō* to the stem, i.e., *kakō* ‘let’s write’. Group-II verbs are referred to as vowel-stem verbs because their stems always end with a vowel. For example, the non-past form of a Group-II verb *kiru* ‘wear’ constitutes the stem *ki* and non-past affix *ru*. Its volitional form constitutes the stem and volitional affix *yō*, i.e., *kiyō* ‘let’s wear.’ Group III has only two members, i.e., *kuru* ‘come’ and *suru* ‘do’ and they are categorized as irregular verbs. *Kuru* constitutes the stem *ku* and non-past affix *ru*. Furthermore, its past-tense form is *kita* ‘came’, which constitutes the stem *ki* and past-tense affix *ta*. Additionally, its volitional form is *koyō* ‘let’s come’, which constitutes the stem *ko* and volitional affix *yō*. *Suru* constitutes the stem *su* and non-past affix *ru*, and its past-

¹ Abbreviations used are as follows: C (consonant), and V (vowel).

tense form *sita* ‘did’ constitutes the stem *si* and affix *ta*. Furthermore, its volitional form *siyō* ‘let’s do’ constitutes the stem *si* and affix *yō*. Unlike Group-I verbs, Group-III verbs involve unpredictable vowel alternation, i.e., *u* is changed to *i*.

Note that the volitional forms of Group-I and Group-III verbs differ particularly in their word endings. The volitional form of Group-I verbs ends with *ō* and that of Group-III verbs typically ends with *yō*.² Some Group-I verbs end with *suru*, e.g., *sasuru* ‘rub’ to *sasurō* ‘let’s rub’. N-*suru* verbs typically change the non-past form: N-*suru* ‘do N’, to the volitional form: N-*siyō* ‘let’s do N,’ e.g., the non-past form: *benkyō-suru* ‘study’ to the volitional form: *benkyō-siyō* ‘let’s study.’ Even though some verbs end with the same *suru*, their volitional forms differ depending on the verb group to which they belong.

Pinker (1999) assumed that the rule of the largest and most regular verb group serves as the default rule for conjugational forms. In Japanese, Group I is the largest group, followed by Group II. According to Pinker (1999), this means that Group-I verbs are the most regular and serve as the default for producing conjugational forms. Group III has only two members; when one of them, *suru*, is attached to a noun (N), it functions as a light verb and forms an N-*suru* verb. For instance, when *suru* is attached to the noun *benkyō*, ‘study,’ it forms the verb *benkyō-suru*, ‘study.’ *Suru* can be attached to various nouns and is highly productive. Note that there are several verbs in Group I whose non-past forms end with *suru*, and Group-III verbs with the *suru* ending are numerous. Therefore, if the dual mechanism works, one possibility is that when the non-past form of a verb ending with *suru* is presented, the default rule will be based on Group-III verbs with the *suru* ending.

2.2 Japanese orthography and Group-I and N-*suru* verbs

Japanese words are written in kanji (logographic script) and in two types of kana (moraic script). Word roots are often written in kanji, affixes in hiragana, and loan words in katakana. Words are often written in a mixture of kanji and hiragana; however, the same words can also be written in kana only.

Morphological boundaries are not orthographically clear in Group-I verbs. For instance, the non-past form of the verb 擦る (*kosuru* ‘rub’) orthographically constitutes kanji 擦 (*kosu*) and hiragana る (*ru*). Morphologically, the stem is *kosur*, and the root corresponds to kanji 擦 and *u* is the non-past affix. The *r* at the stem end and *u* forms one mora and are written as one hiragana る. Therefore, the boundary between the stem end and non-past affix is orthographically unclear. Klafehn (2003) highlighted

2 Although the number of verbs is very small, shifts of some Group-III verbs to other verb groups have been reported (Tanomura 2001, 2009).

that ambiguous morpho-orthographic boundaries may make the application of rule-based mechanism to Japanese inflectional forms difficult because many Group-I verb stems end with a consonant and are never used in isolation.

The word roots of N-*suru* verbs can be compound nouns that constitute a noun (N) and a conjunctive form of a verb (V), and are referred to as NV-type compound nouns (Ishizuka 2013). A light verb construction can be formed by attaching *suru*, or a case marker (*o*: an accusative case marker, and *ga*: a nominative case marker) and *suru*. The word roots of N-*suru* verbs of Chinese origin are written in kanji, and *suru* in hiragana (する, underlined) is attached. Word roots of Japanese indigenous N-*suru* verbs are written only in hiragana (ぐずぐずする *guzuguzusuru*, ‘dawdle’), with a mixture of kanji and hiragana (足踏みする *asibumisuru*, ‘stomp’), and only in kanji (恋する *kouisuru*, ‘love’). Word-roots of N-*suru* verbs of origins other than Chinese and Japanese are typically written in katakana (スタートする *sutātosuru*, ‘start’). Different origins of the word roots of N-*suru* verbs may also affect the insertion of case markers between the N and *suru*, where the script type with which the verbs are written facilitates readers of the verbs identifying their origins. Because novel words do not have any corresponding kanji, they are bound to be presented in kana. This means that if we present real words with a mixture of kanji and hiragana, and novel words in hiragana, participants will be able to distinguish novel words from real words using different types of scripts. Therefore, to check the influence of script type, two lists of stimulus words were constructed. In one list, the stimulus verbs were written only in hiragana, and in the other, they were written with a mixture of kanji and hiragana (described in Section 3.2).

2.3 Previous studies on Japanese verbs

Some studies have investigated the applicability of mechanisms for producing Japanese inflectional verb forms; however, the results have been mixed. Vance (1991) conducted a forced choice completion task. A text containing a citation form of the made-up *godan* (Group-I) verbs, *hoku*, *homu*, *muru*, and *kapu*, in hiragana, and blank spaces was presented. The blank spaces were filled with volitional, past, negation, and conditional forms; the participants were asked to fill them with one of four choices. Participants indicated difficulty choosing the correct form, which can be produced analogically to existing Group-I verbs (for the nonce verb *hatu*, the volitional form *hattoo* and the past form *hatta*). Participants had more difficulty choosing the correct form of the phonotactically inadmissible verb *kapu*. According to Vance, these difficulties indicate that regular verb forms are stored in the lexicon and forms of made-up verbs are analogically produced.

Klafehn (2003) replicated Vance (1991) using native speakers (NSs) of Japanese and non-native speakers (NNSs) learning Japanese. The rule-based mechanism predicts that NSs outperform NNSs; however, NNSs chose correct forms more often than NSs. NSs did not necessarily generalize the conjugation pattern of Group-I verbs to nonce verbs; therefore, no effect of the default rule was observed. These results are more consistent with analogy-based mechanisms.

Batchelder (1999) conducted an aural/oral production task in which texts that contain a non-past form of a nonce verb (*ganosuru*) were aurally presented twice, and either its negative or past form was presented twice; participants were asked to utter the other form that they did not hear (the negative or past-tense form). The stimulus nonce verbs constituted the two-mora stem that end with either *o* or *i* (e.g., *gano*, *nimi*) and the ending of either *ru* or *suru* (e.g., *ganoru*, *ganosuru*, *nimiru* and *nimisuru*). Reaction times to utter the requested verb form and error rates were analyzed. According to Batchelder, if the rule-based mechanism works, producing new inflected forms of nonce verbs is easy, but if individual forms are in memory, producing new forms is difficult, and the ease or difficulty of producing inflected forms of nonce verbs is indicated by reaction times and error rates. The reaction times indicated a significant interaction between verb types (*i/o*, *ru/suru*), but no significant main effects of verb types. The expected forms were produced for the verbs *-oru* more than for the verbs *-osuru*, *-isuru*, and *iru*. No main effect or interaction was observed for error rate, but participants produced incorrect forms, which account for 12.3 %. This difficulty is not predicted by the rule-based mechanism, but can be better explained by an analogy-based mechanism (Batchelder and Ohta 2000).

Fushimi (2002) studied the morphological generalizations of Japanese verbs and utilized similarities between novel and real verbs. The non-past forms of Japanese verbs always end with *u*. Japanese is a moraic language, and one mora typically constitutes a consonant (C) and a vowel (V). The penultimate two syllables of a non-past form constitute $C_1V_1C_2$ and V_2 , and V_2 is always *u* (Table 1). Additionally, typically, the V_1 of Group-I verbs is *a*, *o*, or *u* and C_2 is a consonant other than *r*. The V_1 of Group-II verbs is *e* or *i* and C_2 is *r*. V_1 and C_2 are critical for judging similarities and dissimilarities between Group-I and Group-II verbs. However, some Group-I verbs have similar (S) or dissimilar (D) V_1 and C_2 to those of Group-II verbs. In Table 1, S and D in rows correspond to similar and dissimilar V_1 ; S and D in columns correspond to similar and dissimilar C_2 . For instance, SS-type verbs are Group-I verbs, but hold the identical V_1 and C_2 with those of Group-II verbs. The verb *hukeru* (耽る), ‘be absorbed in’ is a Group-I verb and the homonym of Group-II verb *hukeru* (老ける), ‘grow old’. DD-type verbs are not similar to Group-II verbs regarding V_1 and C_2 . The four categories of Group-I verbs have different degrees of similarity to Group-II verbs (Fushimi et al. 2006). However, they conjugate differently, as described in Section 2.1.

Table 1: Four categories of Group-I verbs by similarity of V₁ and C₂ to Group-II verbs in Japanese.^a

Similarities of V ₁ of Group-I Verbs to Group-II Verbs	Similarities of C ₂ of Group-I verbs to Group-II Verbs	
	Dissimilar: non- <i>r</i>	Similar: <i>r</i>
Dissimilar: <i>a</i> , <i>o</i> , and <i>u</i>	DD (<i>kaku</i>) ‘write’	DS (<i>karu</i>) ‘mow’
Similar: <i>i</i> and <i>e</i>	SD (<i>kiku</i>) ‘hear’	SS (<i>keru</i>) ‘kick’

^aFushimi (2002) and Fushimi et al. (2006) use the term *consistency* and *inconsistency* instead of *similarity* and *dissimilarity*. They assume the typical V₁ and C₂ of Group-I verbs. As this study is trying to describe one of the major differences between the morphological generalization mechanisms through similarity, we re-labeled the categories of Group-I verbs considering *similarity* and *dissimilarity* to Group-II verbs.

Fushimi (2002) presented non-past forms of real verbs with high and low frequencies and novel verbs in kana. They instructed participants to orally provide their corresponding negative forms and polite forms, which end with *mas(u)* (e.g., non-past forms: *kaku* ‘write’, the negative form: *kakanai* ‘do not write’, and *kakimasu* (the polite form of *kaku*)), and measured their naming latencies and error rates. They found that the magnitude of conjugation difficulty for low-frequency verbs varied in the order of DD real verbs, DS real verbs, and Group-II verbs; the order of naming difficulty reflected the similarity effect between the two types of verbs. The frequency effect in each of the four categories was the weakest in DD and DS and Group-II verbs indicated strong effects. For novel verbs, they found that the conjugation pattern of Group-II verbs was applied to novel verbs more frequently in SS than in SD and DS and in DS and SS than in DD, which also indicate a similarity effect. If conjugated forms are retrieved from memory, the time to retrieve them, i.e., naming latency, becomes shorter for high-frequency forms, and they have fewer similarity effects. For low-frequency forms, naming latency increases and is affected more by the similarity effect than for high-frequency forms. Furthermore, the more similar a particular verb is to typical Group-III verbs, the more likely it is to be conjugated in the Group-III pattern. Conversely, if the word is less similar to typical Group-III verbs, it is less likely for the word to be conjugated in the Group-III verb conjugation pattern. They also conducted an orally elicited production task for nonce verbs. As the similarity at V₁ and C₂ increased, Group-II forms increased, which, according to them, is an indication of an analogy-based generalization. However, the overall proportions of Group-I and -II forms were 75 % and 5 %, respectively. The production of Group-II forms was below chance level, which could be due to the block from analogizing further. Therefore, the results can also be interpreted as an indication of rule-based generalization. Fushimi (2005) also mentioned that these results could be explained by both rule- and analogy-based mechanisms.

Oseki et al. (2019) conducted bidirectional elicited-production tasks for past-tense and non-past forms of novel verbs with Japanese NSs as participants and input the same materials into a software program, referred to as a minimal generalization learner (MGL; Albright and Hayes 2003). The MGL learns patterns between the original and output forms during the learning phase and describes the rules from the patterns. In the output phase, when new words are provided to the MGL, rules are applied. Therefore, the MGL simulates rule-based linguistic generalizations (Albright and Hayes 2003).

The human and MGL data indicated a weak correlation in the non-past form to the past-tense form direction. By contrast, they indicated a strong correlation in the past-tense form to the non-past form direction. One possible interpretation of this result is that the analogy-based mechanism worked to produce past-tense forms from non-past forms, whereas the rule-based mechanism worked to produce non-past forms from past-tense forms. Another possible interpretation is that producing past-tense forms from non-past forms is more computationally complex than producing non-past forms from past-tense forms. Furthermore, phonological changes are involved in producing past-tense forms from non-past forms than in producing non-past forms from past-tense forms, which merely indicates different computational complexities or complexities of phonological changes involved in changing one form to a different form rather than the involvement of different mechanisms.

In summary, the aforementioned studies have investigated the mechanisms of morphological generalization in producing the past-tense forms of verbs in Groups I and II, and the results are mixed. No study has yet investigated the inflectional forms in Groups I and III in Japanese using graded similarities, which could be another way to evaluate views on mechanisms. Therefore, this study tested the mechanisms for producing volitional forms of Group-I and Group-III verbs with the *suru* ending through an elicited production task using varying degrees of similarity.³

2.4 Research questions

With the aforementioned background in mind, we made two research questions below.

- (i) When participants produce volitional forms of high- and low-frequency real Group-I verbs with varying degrees of similarity to Group-III verbs, do they rely on the similarity to Group-III verbs with the *suru* ending?

³ There can be a few different expressions of volitionality in Japanese verbs. Since this study deals with morphological generalization in verb forms, we refer to the verb form that constitutes the verb stem and the affix (y)ō as volitional forms of Japanese verbs.

- (ii) When participants produce volitional forms of novel Group-I verbs with varying degrees of similarity to Group-III verbs, do they rely on the similarity to Group-III verbs with the *suru* ending?

3 Volitional form elicited-production experiment

3.1 Participants

Forty Japanese native speakers (List 1: 6 males and 14 females, mean age: 40.65 years, List 2: 16 males and 4 females, mean age: 42.7 years) participated in the experiment.

3.2 Experimental stimuli

Stimuli consisted of low- and high-frequency real and novel words. As described in Section 2.2, orthography may help participants distinguish Group-I verbs from Group-III verbs and may affect their responses. We constructed two lists of stimulus words (Lists 1 and 2). In List 1, real and novel words were presented only in hiragana, while in List 2, the same real words were presented with a mixture of kanji and hiragana, and novel words were presented in hiragana. We indicated pronunciation of each kanji in hiragana in parentheses, e.g., 擦(こす)る.

Forty-eight high-frequency Group-I verbs, and forty-eight low-frequency Group I verbs, six high-frequency Group-III verbs with the *suru* ending, and six low-frequency Group-III verbs with the *suru* ending were chosen from the long unit word list data (2 or higher lemma frequency) of the Balanced Corpus of Contemporary Written Japanese (BCCWJ; National Institute of the Japanese Language 2015). A total of 80 novel words were constructed. Novel verbs consisted of either three or four morae. These were constructed by converting random numbers into hiragana using Excel. A total of 188 words were presented on each list. Table 2 shows the sample's real and novel verbs in the eight categories.

In Table 2, the consonants (C_1 and C_2) are those in the phonemic structure of the last two morae of a word ($C_1 V_1 C_2 V_2$). C_1 , V_1 , and C_2 are particularly important for creating different degrees of similarity of Group-I novel verbs to Group-III verbs, and V_2 of non-past verb forms is always *u*. S stands for “similar” and D for “dissimilar,” and the three capital letters, which are either S or D, indicate the (dis)similarity of C_1 , V_1 , and C_2 to those of Group-III verbs with the *suru* ending. The graded similarity of existing Group-I verbs to Group-III verbs with the *suru*-ending is indicated by the number of phonemes in the last two morae that overlap between Group-I and Group-III verbs. Any non-past verb form ends with the vowel *u*. Therefore, the minimum number of overlapping phonemes is one. The number of phonemes that overlapped between Group-I and Group-III verbs ranged from one to four (Table 2). In each cell of

Table 2: Sample real and novel Group-I verbs and their non-past and volitional forms.^a

Number of phoneme overlaps	C ₁ V ₁ C ₂ u	Real verbs		Novel verbs
		Non-past forms	Volitional forms	Non-past forms
1	DDDu	泳ぐ(およく) <i>oyogu</i> ‘swim’	およごう <i>oyogō</i> ‘let’s swim’	すおす <i>suosu</i>
		注ぐ(そそぐ) <i>sosogu</i> ‘wash’	そそごう <i>sosogō</i> ‘let’s wash’	うそく <i>usoku</i>
2	DSDu	破く(やぶく) <i>yabuku</i> ‘tear’	やぶごう <i>yabugō</i> ‘let’s tear’	ぎぬむ <i>ginumu</i>
	DDSu	縛る(しばる) <i>sibaru</i> ‘tie’	しばろう <i>sibarō</i> ‘let’s tie’	せとる <i>setoru</i>
	SDSu	去る(さる) <i>saru</i> ‘leave’	さろう <i>sarō</i> ‘let’s leave’	おむる <i>omuru</i>
	SSDu	濯ぐ(ゆすぐ) <i>yusugu</i> ‘rinse’	ゆすごう <i>susugō</i> ‘let’s rinse’	すそる <i>susoru</i>
3	DSSu	炙る(あぶる) <i>aburu</i> ‘grill’	あぶろう <i>aburō</i> ‘let’s grill’	いすぶ <i>isubu</i>
	SSSu	刷る(する) <i>suru</i> ‘print’	すろう <i>surō</i> ‘let’s print’	うする <i>usuru</i>

^aUnderlines indicate the phonemes in the last two morae that overlap with the phonemes of *suru*.

non-past forms, in the first tier, each sample verb is written in one kanji root with one hiragana, and only in hiragana in parentheses; in the second tier, Romanized letters and typical meanings are given.

Texts comprising two sentences were constructed. As shown in example (1), the first sentence contained the non-past form of a critical verb (in the Japanese parentheses 「」), which was followed by an affix that requires the non-past form of a verb, e.g., *koto* ‘the fact that’, *mae* ‘before’, and *yō* ‘in order to’. The second sentence has a blank space, which participants were instructed to fill with the volitional form of the verb, and the blank space was followed by an expression that often follows a verb volitional form, e.g., *to omotte imasu(yo)* ‘I think that.’⁴

- (1) A sample text for the novel verb こやる(*koyaru*)
The first line: お酒を「こやる」ことができます。
osake-o “*koyaru*” *koto-ga dekimasu*.⁵
saké-acc “*koyaru*” fact-nom can
‘(I) can *koyaru* saké.’

⁴ *yo* is an emphatic particle.

⁵ Abbreviations used are as follows: acc (Accusative), comp (Complementizer), and top (Topic).

The second line: 朝はコーヒーを ____ と思っています。
asa-wa koohii-o ____ to omotte imasu.
 morning-top coffee-acc ____ Comp I think.
 'I think I will ____ coffee in the morning.'

A total of 188 texts were analyzed. The stimuli were presented one by one on a separate screen via the Internet using PCibex Farm (Zehr and Schwarz 2018).⁶

3.3 Ethics statements

Ethical approval for this study was obtained from the Ethics Committee of Kwansei Gakuin University (No. 2022-54) and the study was conducted in accordance with the principles of the Declaration of Helsinki. All the participants received a written explanation of the purpose and risks of the experiment. Only those who provided informed consent participated in the study.

3.4 Procedure

The procedure was as follows: A brief survey collecting participants' personal information, including age, gender, and language background, was administered, and subsequently they proceeded to the main elicited production task of volitional forms. After the main experiment, a test was conducted in the Japanese language.

In the main experiment, participants were given a two-sentence text and instructed to fill in the blank with a volitional verb form in hiragana.

3.5 Predictions of results

The predicted results are as follows: If the rule-based mechanism was working, the rule of the largest group, that is, the conjugation rule of Group-I verbs, would be applied to the production of volitional forms; therefore, approximately the same production rates of Group-I volitional forms would be found regardless of conditions. However, if the analogy-based mechanism was to work, similarity effects would occur. Namely, the Group-III type of volitional verb forms increases as its similarity to existing verbs increases.

⁶ Although web-based experiments are not conducted face-to-face with participants, their effectiveness has been reported in several studies (Clahsen and Jessen 2021).

4 Results

4.1 Data cleaning

For data cleaning, two participants input their answers with a script other than hiragana and two participants input verb forms other than volitional forms. Data from the four participants did not undergo further analyses. Data from 36 participants were statistically analyzed. For real verbs, when the verb stem was completely different from the stimuli verb, for example, from the non-past form *sasuru* ‘rub’ to *ageyo* ‘let’s give’, its data was excluded from further analyses. Additionally, two words in classical Japanese, *sasagu* ‘dedicate’ and *kitau* ‘train’, were excluded from further analyses.

4.2 Analyses

Participants’ responses were converted to Romanized letters and categorized into three according to verb endings (Table 3). As described in Section 2.1, Group-I verbs typically conjugate from *u* to *ō*, and Group-III verbs with the *suru* ending, that is, N-*suru* verbs, typically conjugate from *ru* to *yō*. As the proportion of endings other than *ō* and *yō* is approximately 2 % or below, we analyzed only the produced forms that end with either *ō* or *yō*.

4.2.1 Analyses of real verbs

The mean proportions of high- and low-frequency real verbs in the four similarity categories by the number of overlapping phonemes and N-*suru* verbs were calculated for List 1 (with real verbs presented in hiragana) and List 2 (with real verbs presented in a mixture of kanji and kana), as shown in Tables 4 and 5, respectively.

Table 3: Overall means of produced forms with three types of ending ($n = 36$).

Ending	Real verbs (%)	Novel verbs (%)
-ō	92.50 (87.89 %)	71.06 (88.82 %)
-yō	10.83 (10.29 %)	7.33 (9.17 %)
Others	1.92 (1.82 %)	1.61 (2.01 %)
Sum	105.25 (100.00 %)	80.00 (100.00 %)

Note. - *ō*: volitional forms that end with *ō*; and -*yō*: volitional forms end with *yō*.

Table 4: Means and population proportions of volitional forms with $\bar{o}$ and $y\bar{o}$ for real verbs in four different similarities and N-*suru* verbs in List 1 (with real verbs only in hiragana) ($n = 17$).

Frequency type	Number of phoneme overlaps and N- <i>suru</i>	Mean (SD)		Proportion (SD)	
		$\bar{o}$	$y\bar{o}$	$\bar{o}$	$y\bar{o}$
High	1	5.94 (0.24)	0.00 (0.00)	1.000 (0.000)	0.000 (0.000)
	2	17.76 (0.44)	0.06 (0.24)	0.997 (0.057)	0.003 (0.057)
	3	17.12 (1.32)	0.00 (0.00)	1.000 (0.000)	0.000 (0.000)
	4	5.76 (0.44)	0.00 (0.00)	1.000 (0.000)	0.000 (0.000)
	N- <i>suru</i> verbs	0.18 (0.39)	5.59 (0.62)	0.031 (0.172)	0.969 (0.172)
Low	1	4.76 (0.66)	0.00 (0.00)	1.000 (0.000)	0.000 (0.000)
	2	16.24 (1.39)	0.06 (0.24)	0.996 (0.060)	0.004 (0.060)
	3	17.00 (1.58)	0.00 (0.00)	1.000 (0.000)	0.000 (0.000)
	4	5.82 (0.39)	0.00 (0.00)	1.000 (0.000)	0.000 (0.000)
	N- <i>suru</i> verbs	1.00 (1.27)	4.88 (1.27)	0.170 (0.376)	0.830 (0.376)

Table 5: Means and population proportions of volitional forms with $\bar{o}$ and $y\bar{o}$ for real verbs in four different similarities and N-*suru* verbs in List 2 (with real verbs in kanji and hiragana) ($n = 19$).

Frequency type	Number of phoneme overlaps and N- <i>suru</i>	Mean (SD)		Proportion (SD)	
		$\bar{o}$	$y\bar{o}$	$\bar{o}$	$y\bar{o}$
High	1	6.00 (0.00)	0.00 (0.00)	1.000 (0.000)	0.000 (0.000)
	2	17.95 (0.23)	0.00 (0.00)	1.000 (0.000)	0.000 (0.000)
	3	17.68 (0.48)	0.00 (0.00)	1.000 (0.000)	0.000 (0.000)
	4	6.00 (0.00)	0.00 (0.00)	1.000 (0.000)	0.000 (0.000)
	N- <i>suru</i> verbs	0.05 (0.23)	5.79 (0.42)	0.009 (0.094)	0.991 (0.094)
Low	1	4.74 (0.56)	0.00 (0.00)	1.000 (0.000)	0.000 (0.000)
	2	16.79 (0.54)	0.00 (0.00)	1.000 (0.000)	0.000 (0.000)
	3	17.53 (0.84)	0.00 (0.00)	1.000 (0.000)	0.000 (0.000)
	4	5.95 (0.23)	0.00 (0.00)	1.000 (0.000)	0.000 (0.000)
	N- <i>suru</i> verbs	0.63 (0.90)	5.26 (0.99)	0.107 (0.309)	0.893 (0.309)

Some participants produced Group-III volitional forms for Group-I verbs, and vice versa. In List 1 one participant produced Group-III volitional forms for Group-I real verbs twice, i.e., よそおよう *yosooyō* for the non-past form 装う *yosoou* ‘dress’, and しゅくしょう *syukusyō* for 祝す *shukusu* ‘congratulate’. For the high-frequency N-*suru* verbs, one participant produced the Group-I volitional form (ろんずろう *ronzurō* for 論ずる *ronzuru* ‘discuss’). For five N-*suru* verbs, ten participants conjugated non-past forms to the volitional form of Group-I verbs (献ずる *kenzuru* ‘offer’, to けんずろう *kenzurō*, 寄与する *kiyosuru* ‘contribute’ to きよそう *kiyosō*, 帰依する *kiesuru* ‘become a believer’

to きえすろう, *kiesurō*, 塗布する *tohusuru* ‘apply an ointment’ to とふすろう *tohusurō*, and 庇護する *higosuru* ‘protect’, 12 forms in total) and Group-I volitional form like forms (塗布する to とふろう *tohurō*, and 帰依する to きえさるう *kiesarō*, 11 forms in total). Furthermore, four participants, three of whom were included in the afore-mentioned ten participants, wrote the volitional forms of synonyms of the three verbs (ぬろう *nurō*, the volitional form of the verb 塗る *nuru* ‘apply’, which is synonymous to the verb 塗布する *tohusuru*, おそわろう *osowarō* ‘let’s learn’, the volitional form of 師事する *sijisuru* ‘become somebody’s pupil’, and つらぬこう *turanukō*, the volitional form of 貫く *turanuku*, the synonym of 固辞する *kojisuru* ‘stick to’, 7 forms in total).

For statistical analyses, general linear mixed models (GLMMs) for logistic regression were fitted into data by using *R* (R Core Team 2023) and the package *lme4* (Bates et al. 2015) with Script Types (kanji-hiragana mixture and hiragana), Frequency (high and low), and Verb Types (Group-I, and Group-III with the *suru* ending, i.e., N-*suru* verbs) as fixed factor, and the number of overlapping phonemes between Group-I and Group-III verbs with the *suru* ending (henceforth indicated, as the Number of Overlapping Phonemes) as covariate. Dependent variables were types of ending in volitional forms (*ō* and *yō*), which are binomial. The maximal model with random intercepts and slopes for both participants and items was adopted according to Barr et al. (2013). The optimizer *Bounded Optimization by Quadratic Optimization* (*bobyqa*, Powell 2009) was applied to the model. When the maximal model failed to converge, the next maximal model was adopted. The adopted model was “Response ~ Script Type * Frequency * Verb Type + (1 + Frequency + Verb Type | ss) + (1 + Script Type | item) + the Number of Overlapping Phonemes.”

The overall results (Table 6) indicated that the main effects of frequency and verb type were significant, and that the interaction between frequency and verb type was significant. The effect of the script type was not significant. For the frequency effect, volitional forms with the *ō* ending were produced more often for low-frequency Group-I verbs than for high-frequency Group-I verbs. Furthermore, they were also produced more often for Group-I verbs than for Group-III verbs with the *suru* ending. For the interaction between frequency and verb type, post-hoc analyses were conducted using the least square means (ls-means) method in the EMMEANS package (Lenth et al. 2023). In this method, the ls-means were calculated for each condition, and post-hoc analyses were performed using Holm’s sequential Bonferroni procedure. The results of post-hoc analyses indicated that in Group-I verbs, the production rates of volitional forms with the *ō* ending between high- and low-frequency verbs were marginally significant ($\beta = -3.44$, $SE = 1.8$, $df = \text{inf.}$, $z = -1.909$, $p = 0.0563$); no clear frequency effect was observed. In Group-III verbs with the *suru* ending, volitional forms with the *ō* ending were produced more often for low-frequency verbs than for high-frequency verbs ($\beta = 3.07$, $SE = 1.15$, $df = \text{inf.}$, $z = 2.675$, $p = 0.0075$). Therefore, Group-I real verbs did not indicate any frequency effect, whereas Group-III real verbs indicated the frequency effect.

Table 6: Results of GLMM for script types, high- and low-frequencies, and verb types.

	Estimate	SE	z-Value	pr(> z)
(Intercept)	−0.1576	1.5039	−0.105	0.91654
Script type	−0.99	0.7575	−1.307	0.19122
Frequency	−3.5266	1.3088	−2.695	0.00705**
Verb type	7.1349	1.1862	6.015	1.8e-09***
Number of overlapping phonemes	−0.4995	0.3335	−1.498	0.13422
Script type: number of overlapping phonemes	−1.2894	1.6725	−0.771	0.44076
Script type: verb type	2.8829	1.5917	1.811	0.07011.
Frequency: verb type	3.7279	1.4804	2.518	0.01179*
Script Type: frequency: verb type	2.3571	2.0614	1.143	0.25287

Signif. codes: ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘.’ 1.

4.2.2 Analyses of novel verbs

No effect of script type difference was found in real-word stimuli as mentioned in 4.2.1; hence, we excluded the script type from further analyses and collapsed the data of the two lists into one and recalculated the means and standard deviations according to the number of overlapping phonemes (Table 7).

For some novel verbs in the 1- to 3-phoneme overlapping conditions, 21 participants conjugated their non-past form to Group-III like volitional forms, e.g., 1 phoneme ぐよむ (*guyomu*) to ぐよう (*guyō*), 2 phoneme さうす (*sausu*) to さうしょう (*saushiyō*), and 3 phoneme おすう (*osuu*) to おしょう (*oshiyō*). Furthermore, six participants randomly wrote the volitional form of *suru*, しょう (*siyō*, ‘let’s do’) to any condition, and three of the six participants were the same participants who wrote volitional forms of synonymous higher-frequency verbs for real verbs as described in 4.2.1. For the 4-phoneme overlap condition, 33 participants conjugated non-past forms to Group-III like volitional

Table 7: Means and population proportions of volitional forms with *ō* and *yō* for novel verbs (*n* = 36).

Number of overlapping phonemes	Mean (SD)		Proportion (SD)	
	<i>ō</i>	<i>yō</i>	<i>ō</i>	<i>yō</i>
1	9.25 (1.23)	0.53 (0.99)	0.946 (0.23)	0.054 (0.23)
2	27.69 (4.20)	1.64 (3.24)	0.944 (0.23)	0.056 (0.23)
3	26.83 (4.84)	2.64 (4.16)	0.910 (0.29)	0.090 (0.29)
4	7.28 (1.76)	2.53 (1.71)	0.742 (0.44)	0.258 (0.44)

forms, e.g., げする (*gesuru*) to げしょう (*gesosiyō*) and いそする (*isosisuru*) to いしょう (*isosiyō*). Additionally, three participants conjugated any novel verbs including the 4-phoneme overlapping condition to Group-I like volitional forms, e.g., いそする (*isosuru*) to いそすろう (*isosurō*).

The overall analyses were conducted with GLMMs for logistic regression fitted to the data by using *R* and the package *lme4* with the number of overlapping phonemes between Group-I and Group-III verbs with the *suru* ending as a fixed factor. Dependent variables were ending types of volitional forms (*ō* and *yō*), which are binomial. The maximal model with random intercepts and slopes for both participants and items was adopted according to Barr et al. (2013). The optimizer *bobyqua* was used for the model. When the maximal model failed to converge, the next maximal model was adopted. The adopted model was “Response ~ The Number of Overlapping Phonemes + (1+ The Number of Overlapping Phonemes | ss) + (1 + The Number of Overlapping Phonemes | item).”

The results (Table 8) indicated a significant main effect of the number of overlapping phonemes of novel verbs with Group-III verbs with the *suru* ending, that is, *N-suru* verbs. We also conducted post-hoc multiple comparisons using the *ls-means* method. In the results (Table 9), only when the number of overlapping phonemes was 4, that is, when the critical novel verbs had the *suru* ending, volitional forms with the *yō* ending were produced significantly more often than when the number of overlapping

Table 8: Results of GLMM for the number of overlapping phonemes.

	Estimate	SE	z-Value	pr(> z)
(Intercept)	4.6024	0.5332	8.631	<2e-16***
Number of overlapping phonemes	−0.4531	0.1923	−2.356	0.0185*

Signif. codes: ‘***’ 0.001 ‘*’ 0.05 ‘.’ 0.05.

Table 9: Results of post-hoc analyses for the numbers of overlapping phonemes.

Contrast	Estimate	SE	df	z-Value	pr(> z)
1 vs. 2	−1.409	0.943	Inf	−1.495	0.4408
1 vs. 3	0.359	0.748	Inf	0.481	0.9634
1 vs. 4	2.697	0.985	Inf	2.739	0.0313*
2 vs. 3	1.769	0.779	Inf	2.271	0.1048
2 vs. 4	4.106	1.093	Inf	3.758	0.001***
3 vs. 4	2.338	0.821	Inf	2.846	0.023*

Signif. codes: ‘***’ 0.001 ‘*’ 0.05 ‘.’ 1. *Note.* The numbers in contrast indicate the number of overlapping phonemes of the last two morae in Group-I novel verbs and Group-III novel verbs with the *suru* ending.

phonemes ranged from 1 to 3. This means that only when the novel verb ended with *suru*, the Group-I type of volitional form was less produced. Furthermore, the Group-III type of volitional form was produced more than the other conditions, although the proportion of Group-III type was still below chance level (40 % or less).

5 Discussion

5.1 High and low-frequency real verbs

The results for the real verbs are as follows. First, for the real Group-I verbs, regardless of the similarity indicated by the number of overlapping phonemes in the last two morae between Group-I and Group-III verbs, frequency and script type, the conjugation pattern of Group-I verbs was dominantly preferred over that of Group-III verbs with the *suru* ending, regardless of similarity. The lack of a similarity effect suggests that the rule for Group I was applicable to any of the real Group-I verbs. Second, frequency effect was observed in Group-III real verbs but not in Group-I verbs. Finally, in Group-III verbs with the *suru* ending, more volitional forms with the *ō* ending were produced for low-frequency verbs than for high-frequency verbs. No frequency effect in Group-I verbs suggests that a rule is applied, regardless of frequency, whereas the frequency effect in Group-III verbs indicates that even though the verbs are real Group-III verbs, when their frequencies are low and participants are unfamiliar with them, a conjugational rule for Group I was conceivably applied to them. In summary, the results for real verbs can be best explained by the rule-based mechanism.

5.2 Novel verbs

One of the findings for novel verbs is that the conjugation pattern of Group-I verbs was dominantly preferred over that of Group-III verbs in the 1 to 3-phoneme overlapping conditions. This result indicated almost no similarity effect, which supports the rule-based mechanism. The use of the *ō* ending suggests that Group I is the largest of the three Japanese verb groups and its rule of attaching *ō* to the verb stem is the default rule to make volitional forms (Pinker 1999) in Japanese despite the fact that *suru* as a light verb is highly productive. Alternatively, Albright and Hayes (2003) assumed two or more default rules in a language. Considering the high productivity of the light verb *suru*, it is conceivable that both Group-I and N-*suru* verbs work as defaults, and inflectional forms are created by applying different rules to different verb groups in Japanese.

Second, as mentioned in 4.2.2, some participants produced forms with Group-III like forms with *yō* for novel verbs, but they also incorrectly produced Group-III forms

for real Group-verbs, as described in 4.2.1. This result indicates idiosyncrasy, which is also reported in Batchelder (1999). Third, the conjugation pattern of Group-III verbs increased only when the critical verbs ended with *suru* (*suru* to *siyo*) (Table 9), and the proportions did not exceed chance level (below 20 %). Michon and Nakipoğlu (2020) reported a similar pattern of results and they argued that although less preferred conjugational forms were produced for nonce verbs similar to irregular verbs, the proportion of the conjugational forms is below chance level because the default rule in Turkish is blocking the conjugational forms of irregular verbs. Because the pattern of results in this study is similar to that of Michon and Nakipoğlu (2020), it is possible that the default rule to add *ō* is blocking the use of *siyō*.

In summary, the results for real and novel verbs showed a preference for Group-I type of volitional forms over Group-III type of volitional forms, regardless of the similarities of Group-I real and novel verbs to Group-III verbs with the *suru* ending. Therefore, the rule to attach *ō* of Group-I verb served as the default rule to form volitional forms, and the results were more compatible with the rule-based mechanism than with the analogy-based mechanism. When the volitional forms of Group-III verbs are produced for Group-III verbs, one possibility is that, as a light verb is highly productive and its distribution is large in Japanese verbs (Nomura 1999), they are also produced by a rule.⁷

Vance (1991), Batchelder (1999), and Klafehn (2003) supported the analogy-based mechanism, whereas the results of this study are compatible with the rule-based mechanism. This difference can be attributed to the different tasks and materials used in the previous studies. Vance (1991) and Klafehn (2003) used four nonce verbs as stimuli: *hoku*, *homu*, *muru*, and *paku*. The first three words exist as literary and classical forms, that is, *hoku* corresponds to the modern verb *hokeru* ‘become senile’, *homu* to *homeru* ‘praise’ and *muru* to *mureru* ‘gather’. Some classical forms appear in literary works, biblical translations, and hymns. The two answer choices in the conditional, e.g., *homureba* and *homereba* were expected to be rated as incorrect, but they also exist. *Homureba* serves as the conditional of *homu* ‘praise’ and *homereba* is the conditional form of *homeru*. They may have semantically primed their corresponding modern verbs and conjugational forms. Vance (1991), Batchelder (1999), and Klafehn (2003) also dealt with other forms, such as past-tense and negative forms as well as volitional forms and tried to tap into the mechanism working for the whole conjugational paradigm. Although past-tense forms can be categorized as inflected forms, negative forms can be inflected forms and derived forms. *Nai* conjugates in the same way as adjectives, and some negative forms work as adjectives (e.g., *warikirenai*, the negative form of *warikireru* ‘can be divided’; Kishimoto 2015).

7 In the corpus used in Nomura (1999), 52,000 words are listed, and Japanese verbs number 21669 and N-*suru* verbs number 7649 words, which accounts for over 89 % of verbs.

Hagiwara et al. (1999) argued that the dual-mechanism model produces derived forms. Therefore, different mechanisms may produce different inflected and derived forms. As this study investigated only volitional forms, further studies should be conducted on other forms.

6 Conclusions

This study investigated the mechanisms that contribute to morphological generalization in the production of volitional forms of Japanese verbs. The results did not indicate any similarity effects. Therefore, the results of this study can be best explained by the rule-based mechanism.

Acknowledgments: I would like to thank all the participants in the experiment. Some parts of this paper were presented at *Morphology & Lexicon Forum 2023*, and I am grateful to the valuable comments received there. This study was supported by JSPS KAKENHI Grant Numbers 21K00492.

References

- Albright, Adam & Bruce Hayes. 2003. Rules vs. analogy in English past tenses: A computational/experimental study. *Cognition* 90(2). 119–161.
- Barr, Dale J., Roger Levy, Christoph Scheepers & Harry J. Tily. 2013. Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language* 68(3). 255–278.
- Batchelder, Eleanor Olds. 1999. Rule or rote? Native-speaker knowledge of Japanese verb inflection. In *Proceedings of the Second International Conference on Cognitive Science and the 16th Annual Meeting of the Japanese Cognitive Science Society Joint Conference (ICCS/JCSS99)*, Waseda University, Tokyo, Japan, July 27–30, 141–146.
- Batchelder, Eleanor Olds & Atsuko Ohta. 2000. Rule vs. rote in Japanese verb inflection. *LACUS Forum XXVI*, 55–66. Available at: <https://archive.org/details/lacus26-new/page/n55/mode/2up>.
- Bates, Douglas, Martin Mächler, Ben Bolker & Steve Walker. 2015. Fitting linear mixed-effects models using lme4. *Journal of Statistical Software* 67(1). 1–48.
- Berko, Jean. 1958. The child's learning of English morphology. *Word* 14(2-3). 150–177.
- Bybee, Joan L. 1995. Regular morphology and the lexicon. *Language and Cognitive Processes* 10(5). 425–455.
- Bybee, Joan L. & Carol Lynn Moder. 1983. Morphological classes as natural categories. *Language* 59(2). 251–270.
- Clahsen, Harald. 1999. Lexical entries and rules of language: A multidisciplinary study of German inflection. *Behavioral and Brain Sciences* 22(6). 991–1013.
- Clahsen, Harald & Anna Jessen. 2021. Morphological generalization in bilingual language production: Age of acquisition determines variability. *Language Acquisition* 28(4). 370–386.
- Fushimi, Takao. 2002. *Doshikatsuyo ni okeru katsuyogata ikkansei no koka* [The effect of consistency of verb conjugation patterns]. *The 5th Ninchi Shinkeigaku Kenkyūkai*, Nagoya University, Japan, 2–3 August. Available at: <http://cnps.umin.jp/pastcnp/PDFs2002/Fushimi.pdf>.

- Fushimi, Takao. 2005. Go no bunpō to sono shōgai-nijōsisutemukasetsu to tan'itsu sisutemu kasetu no tairitsu [Lexical grammar and its impairment—the dual system model v.s., the single system model]. In Sumiko Sasanuma (ed.), *Gengo komyunikēshon shōgai no atarasii siten to kainyū riron* [A new perspective and theory on impaired language communication]. *Koji no kino kenkyū* [Higher Brain Function Research], vol. 12(2), 157–182. Tokyo: Igakushoin.
- Fushimi, Takao, Karalyn Patterson, Matsuo Ijuin, Naoko Sakuma, Yoichi Kureta, Masayuki Tanaka, Tadashi Kondo, Shigeaki Amano & Itaru F. Tatsumi. 2006. Inflecting Japanese verbs: Two separate mechanisms or one graded system? [Manuscript submitted for publication.] In Takao Fushimi (Principal Investigator). (2003–2006). Dosi katsuyo mekanizumu no ninchisinrigakuteki, sinkeisinrigakuteki, keisankagakuteki kento [Cognitive psychological, neuropsychological and computational scientific studies on verb conjugation mechanism] (Project No. 15500176) [Grant]. Ministry of Education and Sciences Grants-in-Aid for Scientific Research Report, 53–77.
- Hagiwara, Hiroko, Yoko Sugioka, Takane Ito, Mitsuru Kawamura & Jun-ichi Shiota. 1999. Neurolinguistic evidence for rule-based nominal suffixation. *Language* 75(4). 739–763.
- Hale, John. 2001. A probabilistic Earley parser as a psycholinguistic model. In *Second meeting of the North American chapter of the association for computational linguistics*.
- Ishizuka, Naoko. 2013. Wago sahen dosi no yōhō to kakuhyōji [Usages and case marking of Japanese indigenous irregular verbs of sa-column]. *Tsukuba Nihongo Kenkyū* 17. 66–82.
- Kishimoto, Hideki. 2015. *Bunpo gensho kara toraeru nihongo* [Japanese as grammatical phenomena]. Tokyo: Kaitaku sha.
- Klaffehn, Terry. 2003. *Emergent properties of Japanese verbal inflection*. Manoa: University of Hawai'i Doctoral dissertation (No. 4350). Available at: <http://hdl.handle.net/10125/6864>.
- Lenth, Russell V., Ben Bolker, Paul Buerkner, Iago Giné-Vázquez, Maxime Herve, Maarten Jung, Jonathon Love, Fernando Miguez, Hannes Riebl & Henrik Singmann. 2023. *Emmeans: Estimated marginal means, aka least-squares means*. R package version 1.9.00.
- Michon, Elise & Mine Nakipoğlu. 2020. Rule-based vs. similarity-based generalization: An experimental study on the Turkish Aorist. *Mediterranean Morphology Meetings* 12. 54–63.
- National Institute for Japanese Language and Linguistics. 2015. Long unit word list data (2 or higher frequency) of the Balanced Corpus of Contemporary Written Japanese. Available at: <https://clrd.ninjal.ac.jp/bccwj/freq-list.html>.
- Nomura, Masaaki. 1999. Sahen dosi no kōzō [The structure of sa-column verbs]. In Morita Yoshiyuki kyōju koki kinen ronbunshū kankōkai (ed.), *Nihongo kenkyū to nihongo kyōiku* [Studies on the Japanese language and Japanese language education]. Japan: Meiji Shoin.
- Nosofsky, Robert M. 1989. Further tests of an exemplar-similarity approach to relating identification and categorization. *Perception & Psychophysics* 45(4). 279–290.
- Oseki, Yohei, Yasutada Sudo, Hiromu Sakai & Alec Marantz. 2019. Inverting and modeling morphological inflection. In *Proceedings of the 16th Workshop on Computational Research in Phonetics, Phonology, and Morphology (SIGMORPHON)*, 170–177.
- Pinker, Steven. 1999. *Words and rules: The ingredients of language*. New York: Basic Books.
- Pinker, Steven & Alan Prince. 1988. On language and connectionism: Analysis of a parallel distributed processing model of language acquisition. *Cognition* 28(1–2). 73–193.
- Powell, Michael J. D. 2009. The BOBYQA algorithm for bound constrained optimization without derivatives. *Cambridge NA Report NA2009/06*, Cambridge: University of Cambridge, 26.
- R Core Team. 2023. *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing. Available at: <https://www.R-project.org/>.
- Rumelhart, David E. & James L. McClelland. 1986. On learning the past tenses of English verbs. In David E. Rumelhart, James L. McClelland & PDP Research Group (eds.), *Parallel distributed processing*:

- Explorations in the microstructure of cognition. Volume 2: Psychological and biological models*, 216–271. MA, Cambridge: The MIT Press.
- Rumelhart, David E., James L. McClelland & PDP Research Group (eds.). 1986. *Parallel distributed processing, Volume 1: Explorations in the microstructure of cognition: Foundations*. MA, Cambridge: The MIT Press.
- Tanomura, Tadaharu. 2001. Sahen dōsi no katsuyō no yure nit suite: Densi siryō ni motozuku bunseki [An analysis of the morphological alternations of *sahen* verbs]. *Japanese Linguistics* 9. 9–32.
- Tanomura, Tadaharu. 2009. Sahen dōsi no yure nit suite · zoku: Daikibona densi siryo no riyō ni your bunseki no seimitsuka [Morphological changes of *sahen*-verbs revisited]. *Japanese Linguistics* 25. 91–103.
- Uygun, Serkan, Lara Schwarz & Harald Clahsen. 2023. Morphological generalization in heritage speakers: The Turkish aorist. *Second Language Research* 39(2). 519–538.
- Vance, Timothy J. 1991. A new experimental study of Japanese verb morphology. *Journal of Japanese Linguistics* 13. 145–166.
- Veríssimo, João & Harald Clahsen. 2014. Variables and similarity in linguistic generalization: Evidence from inflectional classes in Portuguese. *Journal of Memory and Language* 76. 61–79.
- Zehr, Jeremy & Florian Schwarz. 2018. PennController for internet based experiments (IBEX). Available at: <https://doi.org/10.17605/OSF.IO/MD832>.