: +Running head: Within-word speech errors

Within-word serial order control: Adjacent mora exchange and serial position effects in repeated single-word production.

Masataka Nakayama[1,2]

Satoru Saito[1]

1. Graduate School of Education Kyoto University, Yoshida-honmachi, Sakyo-ku, Kyoto 606-8501, Japan

2. Japan Society for the Promotion of Sciences, 8 Ichibancho, Chiyoda-ku, Tokyo, 102-8472, Japan

Corresponding Authors:

Masataka Nakayama or Satoru Saito

Graduate School of Education Kyoto University, Yoshida-honmachi, Sakyo-ku, Kyoto 606-8501, Japan.

Telephone: +81 75 753 3067

Fax: +81 75 753 3066

E-mail: nakayama.masataka.36x@st.kyoto-u.ac.jp

    or saito.satoru.2z@kyoto-u.ac.jp

Abstract

An essential function of language processing is serial order control. Computational models of serial ordering and empirical data suggest that plan representations for ordered output of sound are governed by principles related to similarity. Among these principles, the temporal distance and edge principles at a within-word level have not been empirically demonstrated separately from other principles. Specifically, the temporal distance principle assumes that phonemes that are in the same word and thus temporally close are represented similarly. This principle would manifest as phoneme movement errors within the same word. However, such errors are rarely observed in English, likely reflecting stronger effects of syllabic constraints (i.e., phonemes in different positions within the syllable are distinctly represented). The edge principle assumes that the edges of a sequence are represented distinctly from other elements/positions. This principle has been repeatedly observed as a serial position effect in the context of phonological short-term memory. However, it has not been demonstrated in single-word production. This study provides direct evidence for the two abovementioned principles by using a speech-error induction technique to show the exchange of adjacent morae and serial position effects in Japanese four-mora words. Participants repeatedly produced a target word or nonword, immediately after hearing an aurally presented distractor word. The phonologically similar distractor words, which were created by exchanging adjacent morae in the target, induced adjacent-mora-exchange errors, demonstrating the within-word temporal distance principle. There was also a serial position effect in error rates, such that errors were mostly induced at the middle positions within a word. The results provide empirical evidence for the temporal distance and edge principles in within-word serial order control.

Within-word serial order control: Adjacent mora exchange and serial position effects in repeated single-word production

1. Introduction

The compositional nature of language allows humans to express and comprehend a nearly infinite number of ideas via a finite repertoire of elements. Although the units used to represent elements may differ within a language (e.g., sentence, word, phoneme, phonemic feature) and between languages (e.g., syllable, mora), the flexible use of element combinations allows us to deal with an enormous number of concepts and meanings. For example, in Japanese, "tatsumaki" (*tornado*) and "tamatsuki" (*billiard*) are different concepts, but they are represented by identical sound units. Similarly, "a half-formed wish" and "a half-warmed fish" express different meanings but employ identical sound units. However, "a half-formed wish" and "a half-warmed fish" *must* be differentiated for accurate communication (Jefferies, Grogan, Mapelli, & Isella, 2012). As these examples indicate, an essential characteristic of language, especially spoken language, is its sequential nature and compositionality, which raise the problem of serial order control.

1.1. Models of serial order control

To tackle the problem of serial order control, a number of computational models have been proposed in the domains of speech production and serial order memory. They include localist connectionist models (Burgess & Hitch, 1999; Dell, Burger, & Svec, 1997; Dell, 1986;

Hartley & Houghton, 1996; Houghton, 1990), parallel distributed recurrent connectionist models

(Botvinick & Plaut, 2006; Dell, Juliano, & Govindjee, 1993; Elman, 1990; Gupta & Tisdale,

2009; Plaut & Kello, 1999; Ueno, Saito, Rogers, & Lambon Ralph, 2011), and other types of

mathematical models (Brown, Preece, & Hulme, 2000; Henson, 1998; Page & Norris, 1998;

Vousden, Brown, & Harley, 2000). Although the details of these models vary according to the

research topic, we focus on their functional similarities and common principles.

A fundamental problem for serial order control is how to deal with plans in which

elements and order information are represented in advance (Lashley, 1951). In speech production,

an intended abstract concept should be decoded into time-varying phonological representations

(to produce a sequence of sounds), and this requires an intermediate phonological plan

representation in which all phoneme and order information is compressed (e.g., Plaut & Kello,

1999). In a similar vein, to reproduce a sequence of sounds from time-varying auditory input (i.e.,

in a task based on phonological short-term memory; pSTM), the entire sequence must be

maintained simultaneously in the form of a plan representation and decoded into time-varying

phonological representations (e.g., Gupta & Tisdale, 2009). Although their input/encoding

processes may differ, speech production and pSTM are assumed to share a similar mechanism

for representing and decoding abstract phonological plans (e.g., Saito & Baddeley, 2004). A

general principle governing plan representation and subsequent behavior (i.e., production and

reproduction of single words and sentences/lists) is the similarity principle (Acheson &

MacDonald, 2009a). In the following section, we review evidence for levels of similarity from

empirical data and models of serial order. The data mostly relate to errors in speech

(re)production, which provide information about serial order control mechanisms (e.g., Fromkin,

1971; Garrett, 1975; Henson, Norris, Page, & Baddeley, 1996). In this context, the movement of

elements, in particular exchanges of elements, reveals how similarly these elements and positions are represented. We then consider what support for the similarity principle is missing from the empirical data and describe our experimental approach.

1.1.1. Phonological similarity principle

One source of similarity is phonological, as plan representations should contain information about phonological elements. Phonologically similar phonemes or items tend to be misordered, typically by exchanging one with another, in the context of both speech production and pSTM (e.g., Acheson & MacDonald, 2009b; Page, Madge, Cumming, & Norris, 2007). Similarly, single-word production is vulnerable to distraction by phonologically similar words (Saito & Baddeley, 2004). Almost all models of serial order simulate this phonological similarity effect, though they implement it differently (i.e., feedback from phoneme to lexical representations: Dell, 1986, 1988; misselection of phonologically similar and thus confusing items: Brown et al., 2000; Burgess & Hitch, 1999; Henson, 1998; Page & Norris, 1998; Vousden et al., 2000; distributed coding of plan representations: Botvinick & Plaut, 2006; Dell et al., 1993).

1.1.2. Temporal distance principle

Another important source of similarity is derived from the temporal aspect of language. Plan representations should contain not only element information, but also order information, or information about the position of each element about to be output, and this should be mapped

onto time. Thus, some similarity inevitably reflects the temporal aspect of language (i.e., temporal distance and edgeness). The temporal distance between to-be-output positions determines similarity of the elements and/or of their associated position representations, such that temporally near elements/positions are more similarly represented. Consistent with the temporal distance principle, the transpositions exhibit a gradient whereby elements in adjacent/nearer positions are more likely to be transposed/exchanged in the context of immediate serial recall (Henson et al., 1996), and phonemes in adjacent/nearer syllables/words are more likely to be exchanged in the context of spontaneous speech production (Vousden et al., 2000). This principle is a consequence of the way models represent order information. In models that represent order information by context–element associations (i.e., where element representations are associated with context representations), context representations directly reflect temporal distance using oscillators (Brown et al., 2000; Burgess & Hitch, 1999; Henson & Burgess, 1997; Henson, 1998; Vousden et al., 2000). Elements that are associated with similar contexts (i.e., temporally near elements) tend to move toward or switch positions with each other.

Other models represent order information by an activation gradient. These models employ spread or preparatory activation with a primacy gradient of element representations (Dell, 1986; Houghton, 1990; Page & Norris, 1998) or connections from plan to element (or frame) representations that lead to graded activation (Dell et al., 1997). Thus, elements that are temporally close and receive similar activation values tend to move toward or switch positions with each other.

A further class of serial order models includes parallel distributed-processing recurrent networks. Recurrent networks represent elements and order information conjunctively within a hidden layer of three-layer networks with recurrent connections (Botvinick & Plaut, 2006; Dell

et al., 1993; Elman, 1990; Gupta & Tisdale, 2009; Plaut & Kello, 1999; Ueno et al., 2011).

Distributed representation and learning algorithms (i.e., back propagation; Rumelhart, Hinton, &

Williams, 1986) enable the model to capture the similarity structure of a given environment. As

Botvinick and Plaut (2006) reported, close elements are represented similarly and tend to move

toward or switch positions with one another. Despite their different implementations, all models

reviewed above are governed by the temporal distance principle.

1.1.3. Edge principle

Another source of similarity is the edge effect. All words and sentences begin and end.

The beginnings and ends of sequences are represented distinctly from other positions. A clear

manifestation of this edge effect is the serial position effect, a well-established phenomenon in

the domain of serial order memory. Items at the initial and final list positions (i.e., the list edges)

are recalled better than are items in other positions, with initial positions having the greatest

advantage. This effect has also been reported in the repetition of single nonwords (e.g., Gupta,

2005; Gupta, Lipinski, Abbs, & Lin, 2005; Gupta & Tisdale, 2009). The temporal distinctiveness

theory proposed by Glenberg and Swanson (1986) suggests that the recency effect occurs due to

higher distinctiveness toward the end of a memory list, whereas the primacy effect is based on

different mechanisms. Other models of serial order have replicated the serial position effect by

directly representing edges as distinct from other positions/elements (Henson, 1998; Henson &

Burgess, 1997) or by learning to represent edges distinctively because they always follow or are

followed by a null context (Botvinick & Plaut, 2006; Gupta & Tisdale, 2009).

1.1.4. Syllabic constraint

Syllabic constraint is also a widely accepted organizing principle. Different syllabic positions are distinctly represented (and same syllabic positions in different syllables are represented rather similarly). This constraint is indicated by the fact that phonemes do not transpose randomly, but rather transpose to the same within-syllable positions (e.g., onset phonemes transpose to onset, but not coda, positions; Vousden et al., 2000). Although syllabic representations may be regarded as phonological (e.g., which phoneme can occupy which syllable position; Dell, 1986, 1988), positional (Vousden et al., 2000), or both (Dell et al., 1993, 1997; Hartley and Houghton, 1996), the critical point is that such constraints work at least partly independently of similarity, reflecting the other aspects of language described above (we return to this point in *4.6. Toward a universal account of syllabic constraint*).

1.2. Missing empirical data

Many models assume that distinct representations capture within- and between-word temporal information and that they are governed by the same principles reviewed above (Dell, 1986; Dell et al., 1997; Gupta & MacWhinney, 1997; Vousden et al., 2000; see also Botvinick, 2007; Burgess & Hitch, 1999, 2006; Henson, 1998, for use of multi-level representations). However, the assumed temporal distance and edge principles at the within-word level have yet to be examined empirically. In fact, within-word adjacent element exchange and serial position effects have not been fully elucidated in an experimental setting. As described above, movement errors exhibit a transposition gradient when distance is calculated in terms of syllable/word units,

which only indicates the temporal distance principle at the syllable/word level. Nonetheless, phonemes do not move to a unit in the same word or an adjacent unit. In other words, phonemes transpose to distant positions rather than to the nearest position when distance is calculated in terms of phoneme units. For example, Vousden et al. (2000) reported that only 12.2% of naturally observed speech errors in English involve movement errors within a word. Syllabic constrains are so strong in English that they may prevent the occurrence of within-word movement errors and obscure the within-word temporal distance principle. Thus, the assumed but hidden within-word temporal distance principle needs to be directly demonstrated.

In addition to the lack of within-word adjacent-element exchange, it has not yet been demonstrated that the serial position effect occurs in single-word production in a manner consistent with edge distinctiveness. Though Gupta (2005; Gupta et al., 2005) has shown within-item serial position effects, these studies were conducted in the context of nonword repetition and detailed analyses of the error patterns were not reported. Concluding that word edges are distinctly represented requires the demonstration that phonemes at the edge positions of a word are less susceptible to exchange with adjacent phonemes.

1.3. The current study

To dissociate the effects of temporal distance and edge, we need a language that is structured to minimize the influence of syllabic constraints on within-word adjacent-element exchanges. Japanese is an ideal language[1] for this purpose, as the mora[2] sub-syllabic unit plays a

---

[1] Japanese may not be the only ideal language. Within-word speech errors and primacy effects are observed in naturally occurring speech errors in Spanish (García-Albea, del Viso, & Igoa,

more important role than the syllable in Japanese speech production. Importantly, a mora

behaves as a moving unit that is at least partly independent of syllabic constraints (Kubozono,

1989). A core issue regarding the rarity of within-word phoneme misordering in English is that

adjacent phonological units are not exchanged, and this gradient has not been shown to be

dissociated from syllabic constraints. Thus, the demonstration of within-word adjacent-element

exchanges, irrespective of unit, are theoretically important (cf. proximate unit principle;

O'Seaghdha & Chen, 2009; O'Seaghdha, Chen, & Chen, 2010). Japanese is an ideal language

for the present investigation in this respect.[3] Indeed, a number of within-word phoneme

exchanges are common in Japanese. More than 80% of exchange errors (including all types of

moving units; phoneme, mora, and syllable) in naturally spoken Japanese occur within a word

(Terao, 2002). Although the observation of naturally occurring speech errors can provide clues

about how the language system works, experimental techniques that induce speech errors allow

active tests of these hypotheses in a more rigorous manner and allow detection of otherwise

hidden patterns. In this study, a speech-error induction technique developed by Saito and

Baddeley (2004) was used. In their experiments, Japanese participants were required to

repeatedly utter a single Japanese word target. An auditory distractor was suddenly presented to

---

1989). Nonetheless, besides language selection, our experimental approach has clear advantages
for rigorously demonstrating these phenomena.

[2] A mora includes a vocalic nucleus (V), a nucleus onset (CV or CCV), a nasal consonant (N) in
the syllabic coda position, and a geminate consonant (Q; Otake, Hatano, Cutler, & Mehler, 1993).
For example, the two-syllable word *Kyoto* has three morae (CCV-V-CV: Kyo-o-to), as does the
word *ninja* (CV-N-CV: ni-n-ja).

[3] Many features other than those listed in the main text are suitable for testing the temporal
distance and edge principles. Compared to other languages such as English, Japanese has less
strict phonotactic constraints (e.g., almost all morae can occupy any position in a word), a faster
speech rate or shorter syllable/mora duration (Kohno, 1998), a simpler and more mutually
similar syllabic structure (i.e., the mora; see *4.6. Toward a universal account of syllabic
constraint*), and a smaller phoneme and syllable inventory (Tamaoka & Makioka, 2004, 2009;
see Terao, 2002, for a discussion).

induce speech errors (see Fig. 1). Distractors that were phonologically similar to the target (e.g.,

shi-o-zu-ke,[4] "pickles," against *Shi-zu-o-ka*, name of a prefecture) successfully induced

phonological speech errors (e.g., *shi-o-zu-ka*, "mound of salt"). To induce within-word mora-

exchange errors, we used a specific type of phonologically similar distractor. All similar

distractors were words in which the exchange of adjacent morae resulted in a target word

(Experiments 1 and 2) or a target nonword (Experiment 2). For example, when the target was

"ta-*ma-tsu*-ki" (billiard), the similar distractor was "ta-*tsu-ma*-ki" (tornado). We expected that

phonologically similar distractors would induce a consistent number of speech errors that would

be primarily characterized by mora exchanges.

This technique (and the fact that multi-syllabic/moraic words are common in Japanese)

also allows us to demonstrate a within-word serial position effect. To this end, words (and

nonwords) with a four-mora structure were used and the position of to-be-exchanged morae was

manipulated (i.e., 1-2, 2-3, and 3-4 morae exchanges). For example, *ku-ro*-da-i (black porgy) to

ro-ku-da-i (sixth) is a 1-2 morae exchange; ta-*ma-tsu*-ki (billiard) to ta-*tsu-ma*-ki (tornado) is a 2-

3 morae exchange; and ya-ki-*me-shi* (fried rice) to ya-ki-*shi-me* (ash-glazed pottery) is a 3-4

morae exchange. This manipulation allowed us to examine whether similarity of position

representations led to mora exchange by strongly biasing the direction of movement. For

example, the second mora should move to the first mora position with a 1-2 exchange distractor,

and to the third mora position with a 2-3 exchange distractor. Comparing the error rates between

these position pairs allowed us to examine which pairs are represented more similarly. Using this

manipulation and control (see *2.1. Methods*), overall comparisons between similar and dissimilar

distractor conditions served as a rigorous replication of the phonological similarity effect

---

[4] Hyphen indicates a mora boundary.

demonstrated by Saito and Baddeley (2004), in which the only controlled variables were those related to phonological similarity. In the three phonemically similar conditions, the serial position of the to-be-exchanged morae, and not phonemic similarity, determined similarity, as all phonemes were shared between the target and distractor, allowing for a relatively pure test of serial position effects. We predicted that distractors created by exchanging the middle part of the target (i.e., 2-3 exchange distractors) would induce speech errors more frequently than distractors involving edge exchange (i.e., 1-2 and 3-4 exchange distractors). Both the serial position effect and edge distinctiveness typically confer an advantage for the initial position (e.g., Gupta, 2005) and the initial position is (learned to be) represented more distinctly from final positions, at least in Botvinick and Plaut's (2006) model. Therefore, we also predicted relatively more errors for 3-4 than 1-2 exchange distractors.

Experiment 1 aimed to induce within-word phoneme exchanges and to demonstrate the serial position effect with word stimuli. Experiment 2 aimed to test whether lexical/semantic representations contribute to error prevention using this technique.

2. Experiment 1

2.1. Methods

2.1.1. Participants

Twenty-four graduate and undergraduate students from Kyoto University participated in this experiment in exchange for a book coupon worth 500 yen. One additional participant was

not included in the analysis because of a programming error. The mean age of participants was 22.4 years (range: 20–32 years). All participants were native Japanese speakers.

2.1.2. Materials

A total of 402 four-morae words were selected from a Japanese corpus (Amano & Kondo, 2000). The morae were CV, V, or N (nasal consonant). The nasal consonant was included as a candidate mora, as this mora appears frequently in the Japanese corpus (Tamaoka & Makioka, 2004). The selected words were used as either targets, similar distractors, dissimilar distractors, or filler targets. Participants were asked to utter the target words repeatedly. The distractor words were aurally presented between participant utterances (see *2.1.3. Design and Procedure*; Fig. 1). Similar distractors were words that resulted in the target when adjacent morae were exchanged. For example, when ta-*ma-tsu*-ki (billiard) was a target, the similar distracter was ta-*tsu-ma*-ki (tornado), and vice versa. The three types of similar distractors differed in the positions of to-be-exchanged morae (i.e., 1-2, 2-3, and 3-4 morae exchanges). All distracters were real words and, by definition, phonotactically legal. Logged word frequency[5] (Amano & Kondo, 2000) and the logged sum of the three bi-mora frequencies (Tamaoka & Makioka, 2004) of the targets were matched between conditions (Table A.1) using Match (van Casteren & Davis, 2007). Dissimilar distractor words were phonologically dissimilar to the target word (e.g., *ha-ka-do-ru* for *ta-ma-tsu-ki*) and shared no more than 50% of the phonemes with the target word. Logged word frequency and the logged sum of the three bi-mora frequencies of similar and dissimilar distractors (e.g., *ta-tsu-ma-ki* and *ha-ka-do-ru* for the target *ta-ma-tsu-ki*) were matched (Table

---

[5] We used the frequency of the most frequent written form, which was highly correlated with the frequency of ignored variations of written forms ($r = .99$).

A.1). Similar distractors (e.g., *ta-tsu-ma-ki*) and targets (e.g., *ta-ma-tsu-ki*) were counterbalanced across participants, and dissimilar distractors were paired with similar distractors (e.g., *ta-tsu-ma-ki* and *ha-ka-do-ru,* and *ta-ma-tsu-ki* and *gu-ra-tsu-ku* were paired). In this way, 96 quartets of a target words (or similar distractors), a similar distractor (or target word), and two dissimilar distractors were pooled for each of the three to-be-exchanged morae-position conditions (see Appendix). For each participant, one target and one distractor were selected and used from each quartet. Target and distractor selection from the quartet was counterbalanced. Eighteen additional words were used as filler targets, which also had potentially similar distractors (see *2.1.3. Design and Procedure*). All auditory distractors used in Experiment 1 and Experiment 2 were created with Japanese text-to-speech software (VoiceText editor SAYAKA; HOYA).

2.1.3. Design and Procedure

Figure 1 presents a schematic illustration of the procedure. Each trial began with a visual presentation of the target word in the most frequent written form (*hiragana, katakana*, *kanji*, or a mixture of these[6]) that remained in the center of the computer display (Diamondcrysta RDTI92WLM) until participants pressed the space bar. A *katakana* form of the word was also presented below the most frequent written form to ensure that participants read the word correctly. Participants were instructed to say the target word in time with the appearance of a visual signal (*), which was presented for 250 ms 12–14 times at a rate of 1 signal per second. Thus, participants uttered the same target word 12–14 times in every trial. For well-timed

---

[6] Written Japanese consists of three types of characters. *Hiragana* and *katakana* are phonograms, and each letter corresponds to one mora in most cases (although there are some exceptions). *Kanji* has both semantic and phonological values and thus there is a quasi-systematic (albeit sometimes arbitrary) mapping between meanings and linguistic sounds.

utterances, each visual signal was preceded by a signal tone, which was presented 500 ms before the visual signal for 250 ms at 440 Hz. To induce speech errors, the signal tone was replaced by an auditory distractor word during the third, fourth, fifth, ninth, tenth, or eleventh presentation. The word was presented within the 500-ms interval immediately preceding the visual signal. The signal tones and aurally presented words were delivered via headphones (Audio-Technica ATH-f2SG). The experiment was controlled by Hot Soup Processer (http://hsp.tv/) on a Dell Dimension 9150. Participant responses were noted by the experimenter and recorded with a video camera (Panasonic HDC-HS200) and IC recorder (OLYMPUS Voice Trek V-85). All participants completed all six conditions in the 2 (distractor similarity: similar/dissimilar) × 3 (exchanged morae position: 1-2/2-3/3-4) design. There were 16 trials for each condition, and an additional 18 filler trials, in which no distractor was presented, for a total of 114 trials. Participants could rest between trials. Trial order was pseudo-randomized using Mix (van Casteren & Davis, 2006) such that the same condition (e.g., the dissimilar condition) appeared in no more than three consecutive trials. Filler trials were separated by no fewer than three trials. Before the experimental trials, participants signed a consent form, received instructions, and performed seven practice trials.

Visual presentation of the target (e.g., "ta-*tsu-ma*-ki")
in the most frequent form and *katakana*.

Visual signal

竜巻
タツマキ   0.5s   1s

Tone   1s

The distracter was presented replacing
3rd, 4th, 5th, 9th, 10th or 11th tone.
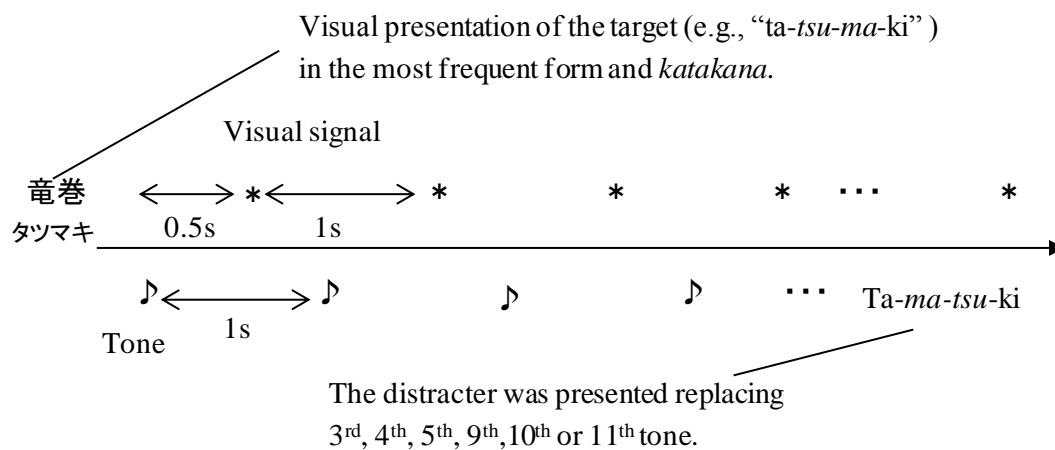
Ta-*ma-tsu*-ki

Fig. 1. A schematic illustration of the procedure.

2.2. Results and Discussion

2.2.1. Preliminary analysis

The experimenter noted errors online, and these were subsequently confirmed by referring to the video and/or voice records. An independent rater coded errors based on the video and/or voice recordings. All experimenters and raters were native Japanese speakers. Errors were categorized as incorrect or rhythm errors. Incorrect errors were defined as phonological errors and included phoneme misordering and substitution. Rhythm errors included delayed responses, intervals where no response was made,[7] non-fluent speech, and false starts (i.e., a correct word was produced after the first utterance of one or more correct phonemes). Co-occurrences of incorrect and rhythm errors were categorized as incorrect errors because they could be analyzed as phonological errors. Although all errors observed in any utterance in each trial were noted, with the exception of trials in which the first utterance was an erroneous response, the following analyses focused only on utterances that occurred immediately after distractor presentation,[8] as in Saito and Baddeley (2004). In the following section, we report only incorrect error data, because the experiments were designed to measure speech errors, not reaction times. Thus, it was difficult to reliably evaluate rhythm errors. Notably, combining incorrect and rhythm errors in a single analysis did not change critical aspects of the results (i.e., the serial position effect). Inter-

---

[7] It may be inappropriate to include "no response" as a rhythm error. However, it is difficult to distinguish "no response" from "delayed response" in this paradigm, as the target is repeatedly produced. Importantly, it is reasonable to exclude "no response" errors from phonological errors, which are our focus.

[8] As our aim was to investigate the principles that govern plan representations for ordered production of phonological elements, we focused only on the utterances where planning could have been influenced by the distractor (i.e., utterances just after distractor presentation). Hence, incorrect responses (i.e., phonological errors) were most frequently produced as the first erroneous utterance in this position (52) in a trial, followed by the next position (16).

rater agreement on incorrect judgments was $\kappa = .84$, and data coded by the experimenter were analyzed.

2.2.2. Error categorization

Table 1 presents a mora-based qualitative categorization of errors. Movement errors (i.e., correct morae appearing in the wrong positions) were the most common (movement = 36, non-movement = 17; $p < .05$, exact test), and exchanges were the most frequent movement errors (exchanges = 29, non-exchanges = 7; $p < .05$, exact test). This experiment successfully induced within-word mora movement errors.

Table 1. Qualitative categorization of errors

		Exp 1	Exp 2	
			word	nonword
movement	total	36	18	18
	exchange	23	15	13
	exchange with deletion	2	0	1
	exchange with addition	4	0	0
	anticipation	1	1	0
	anticipation with deletion	0	0	0
	anticipation with addition	0	0	0
	persiveration	1	0	1
	persiveration with deletion	0	0	0
	persiveration with addition	2	0	2
	others with deletion	3	2	1
substitution	total	12	2	12
	substitution	8	2	11
	substitution with deletion	3	0	1
	substitution with addition	1	0	0
mixed (movement + substitution)		0	3	5
pure deletion (imcomplete)		5	0	3
correct		2242	1126	1103
# of trial		2295	1149	1141

Notes: All error categories are based on mora sized units. Errors were broadly categorized into movement, substation, and mixed errors. *Movement* errors in which correct mora(e) appeared in incorrect positions consisted of exchange, anticipation, perseveration, and other movement. *Exchange*: two morae within a word were exchanged. *Anticipation*: one mora appeared both at an earlier position and at the correct position. *Perseveration*: one mora appeared both at the correct position and at a later position. (Although these three errors are theoretically not mutually exclusive, as there are four mora positions in a target word, they were mutually exclusive in this dataset). *Other movement*: except for the error types described above, all morae were correct, but at least one mora appeared in the wrong position. *Substitution* errors were defined as erroneous utterances in which the wrong mora(e) appeared. *Mixed* errors were defined as at least one correct mora in the wrong position and at least one incorrect mora present. Independent of these criteria, word length of erroneous utterances was considered. *Addition* and *deletion* indicate utterances longer or shorter than four morae, respectively. Pure deletion is an utterance in which all uttered morae were located in the correct positions, but at least one mora was omitted (i.e., an incomplete utterance).

2.2.3. Error rate

Error rates (Fig. 2) were analyzed using logistic mixed-effects regressions (Jaeger, 2008)

in the statistical software package R (R Core Team, 2013). Fixed effects (i.e., phonological

similarity and to-be-exchanged positions) were coded to be centered and orthogonal to each other,

and to generate positive effects on error rate. For "phonological similarity," the similar and

dissimilar conditions were coded as 1 and -1, respectively. For to-be-exchanged positions, we

used two variables ("edge" and "primacy gradient") to capture serial position effects more

precisely. Note that serial position effects are nested within the similarity effect. The dissimilar

distractors were not constructed by exchanging morae in the target word, but by matching

frequency values with similar distractors. In theory, dissimilar distractors should not induce a

serial position effect, and inspection of the data supports this view (Fig. 2.). Thus, for these

variables all dissimilar conditions were coded as 0. For similar distractors, for the "edge"

variable, 1-2 and 3-4 exchanges were coded as -1, and 2-3 exchanges were coded as 2. For

"primacy gradient," 1-2 exchanges were coded as -1, 2-3 exchanges were coded as 0, and 3-4

exchanges were coded as 1. In addition to the fixed effects, all possible random intercepts and

slopes for participants and items were included in the analysis (Barr, Levy, Scheepers, & Tily,

2013). Fixed effects were tested using likelihood ratio tests comparing the full model with

models missing each fixed effect. The likelihood tests indicated significant effects of

phonological similarity and edge (Table 2).
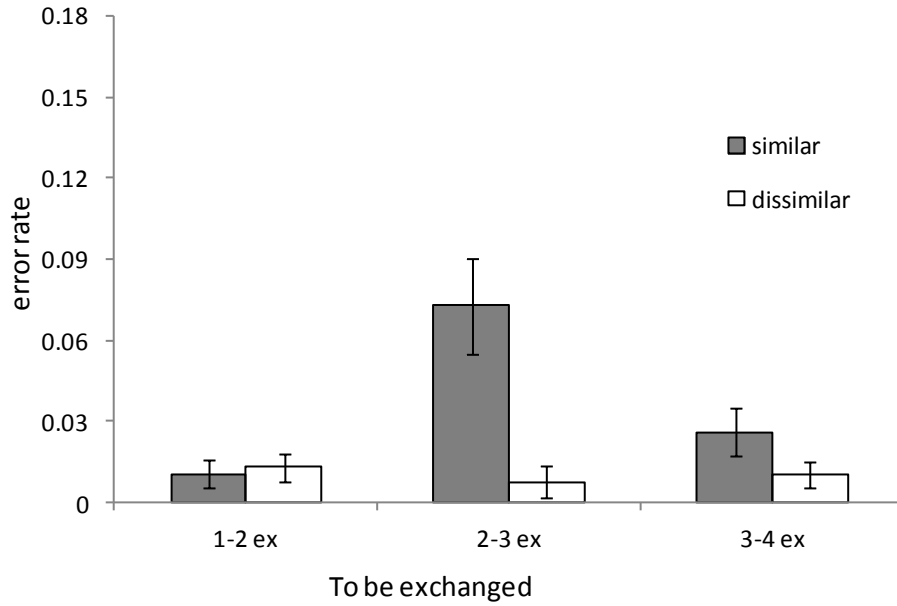
Fig. 2. Incorrect error rate as a function of phonological similarity and exchanged mora positions in Experiment 1.

Note. Gray and white bars indicate mean error rates in the phonologically similar and dissimilar conditions, respectively. Error bars indicate standard errors of the means based on participant analysis.

Table 2. Results of logistic mixed-effects regressions in Experiment 1.

Random effects

Groups		Variance	SD	Correlation		
Participant	(Intercept)	3.12	1.77			
	phonological similarity	0.77	0.88	-0.96		
	edge	0.21	0.46	-0.61	0.81	
	primacy gradient	0.67	0.82	-0.28	0.43	0.62
Item	(Intercept)	23.57	4.86			
	phonological similarity	23.59	4.86	-0.95		

Fixed effects

	Estimate	SE	z value	Pr($>$|z|)	χ^2	Pr($>\chi^2$)
(Intercept)	-9.49	1.35	-7.03	<.001		
phonological similarity	4.34	1.29	3.36	<.001	14.97	<.001
edge	0.70	0.21	3.34	<.001	7.13	<.01
primacy gradient	0.68	0.49	1.40	>.1	1.11	>.1

2.2.4. Movement error rate

These errors could be regarded as intrusions from distractors rather than within-target movement errors because all similar distractors were mora-exchanged sequences of the respective targets. However, our view that these errors were movement errors was supported by more detailed analysis of similar distractor conditions where we determined the origin of each mora in an utterance. Relevant to the present discussion, the movement errors at output position mora 2 were induced more by 2-3 than by 1-2 distractors (18/384 vs. 1/381, $p < .05$, exact test). This was also true when the source mora was restricted to mora 3 in 2-3 exchanges and mora 1 in 1-2 exchanges (17/384 vs. 1/381, $p < .05$, exact test). In this case, the source morae positions in the distractors were the same (i.e., mora 2), but were different in the targets (i.e., mora 1 or mora 3). If the errors were intrusions from the distractors, we would have found similar error rates for 2-3 and 1-2 distractors at output position mora 2. However, this was not the case. The errors observed at output position mora 3 showed a similar tendency (i.e., more errors in 2-3 distractors [19/384 for all movement errors and 17/384 for those from mora 2] than in 3-4 distractors [9/383 both for all movement errors and for those from mora 4]), but this difference was not statistically significant ($p = .081$ and $p = .12$, respectively, exact test). At least in mora 2, error rates were affected by target morae positions rather than by distractor morae positions, indicating that these errors reflected morae movements within target words.[9] In summary, Experiment 1 clearly demonstrated that adjacent morae exchanges were also affected by their serial position, indicating the operation of both temporal distance and edge principles in single-word production.

---

[9] One may argue that these errors reflect lexical intrusions from distractors. However, a supplementary experiment (Nakayama et al., 2012) also indicated that nonword distractors elicited morae exchange and the serial position effect (thus, the error utterances were nonwords rather than words), and that lexical intrusions do not fully explain these errors.

3. Experiment 2

Experiment 2 examined the influence of lexical/semantic representations on the error-induction paradigm used in Experiment 1 and replicated the previous experiment. The present technique apparently tapped the phonological planning component of speech production and reproduction (Saito & Baddeley, 2004). The distractor was presented at a late stage of production (i.e., 500 ms before the utterance was filled by the distractor), which is assumed to correspond to phonological planning. Manipulating phonological similarity led to phonological errors. However, the literature on speech production and pSTM suggests that the mechanisms underpinning speech-error induction need to be examined more carefully. Because several models and empirical data suggest that lexical/semantic representations interact with phonological representations and contribute to phoneme ordering (e.g., Dell, 1986, 1990; Gollan & Goldrick, 2012; Jefferies, Frankish, & Lambon Ralph, 2006; Martin & Saffran, 1997; Patterson, Graham, & Hodges, 1994), it is necessary to test whether lexical/semantic representations affect the occurrence of phonological errors in the error-induction technique used in Experiment 1. In this experiment, nonword targets, which by definition are assumed to have no lexical/semantic representations, were included, and error rates for nonword and word targets were compared. If this technique taps into an interaction between phonological and lexical/semantic representations as well as phonological planning, nonword targets should elicit more speech errors, especially phoneme order errors, than word targets should.

3.1. Methods

3.1.1. Participants

Twenty-four graduate and undergraduate students from Kyoto University participated in

this experiment in exchange for a book coupon worth 500 yen. An additional participant was not

included in the analysis because of a programming error. The mean age of participants was 23.5

years (range: 19–29 years). All participants were native Japanese speakers and not participated in

Experiment 1.

3.1.2. Materials

We selected 240 four-mora words from the Japanese corpus (Amano & Kondo, 2000) and

created 48 four-mora nonwords. Forty-eight of these words were target words that were paired

with 48 similar distractor words, and 48 were similar distractor words for 48 target nonwords. A

total of 192 words served as dissimilar distractors. As in the previous experiment, exchanging the

adjacent mora pair in a similar distractor resulted in the corresponding target word/nonword.[10]

Dissimilar distractor words were phonologically dissimilar to the target. The dissimilar

distractors shared no more than 50% of the phonemes with the target (Table A.3).

Table A.3 presents the psycholinguistic variables in each condition. Bi-mora frequency

was matched between target words and nonwords. Word frequency and the bi-mora frequency of

distracter words were matched across conditions. All participants were exposed to all of the

target words/nonwords (i.e., total of 96 trials). Participants were exposed to either similar or

---

[10] All nonwords were readable when written in *katakana*, and they were phonotactically legal in
that a nasal consonant did not occupy the onset mora position.

dissimilar distractor words for each target, and distractor selection (similar or dissimilar) was counterbalanced.

3.1.3 Design and Procedure

Experiment 2 included a manipulation of target lexicality, and was a 2 (distracter similarity: similar/dissimilar) × 3 (exchanged mora position: 1-2/2-3/3-4) × 2 (target lexicality: word/nonword) design. Participants were tested in all conditions, with eight trials per condition. The procedure was almost identical to Experiment 1, with the following exceptions: (a) filler targets were not included; b) only audio (not video) recording was conducted; and (c) participants read the target word at the beginning of each trial, and an experimenter confirmed that they did so correctly. This reading task was introduced because nonwords may be vulnerable to reading errors, especially when they are created from real words by exchanging adjacent morae (see also Perea, Nakatani, & Leeuwen, 2011; Perea & Pérez, 2008).

3.2. Results and Discussion

3.2.1. Preliminary analysis

Coding was the same as in Experiment 1, except that an independent rater coded data from five of the participants. Inter-rater agreement on incorrect judgments was $\kappa = .78$, and data coded by the experimenter were analyzed.

3.2.2. Error categorization

Table 1 summarizes the qualitative categorization of errors. As in Experiment 1, movement errors were the most common (movement = 36, non-movement excluding mixed errors = 17; $p < .05$, exact test), and exchanges were the most common movement errors (exchanges = 29, non-exchanges 7; $p < .05$, exact test). The prevalence of movement errors, including exchanges, was comparable for words and nonwords in the present experiment and words in Experiment 1 (18, 18, and 36, respectively).[11] There was a lexicality effect for substitutions (words = 2, nonwords = 12; $p < .05$, exact test; see also *3.2.3. Error rate*). Further examination revealed that 9/12 mora-substitution errors in nonwords were segment-movement errors (e.g., "u-k*o*-g*a*-shi" for the target "u-*ka-go*-shi" and the similar distractor "u-*go-ka*-shi").

3.2.3. Error rate

Error rates (Fig. 3) were analyzed with logistic mixed-effects regressions using R. For fixed effects, phonological similarity, edge, primacy gradient, and lexicality (words and nonwords were coded as -1 and 1, respectively) were included in the model. We report a model without lexicality-related interactions because a model including these interactions failed to converge correctly, and because inspection of the data (Fig. 3) indicated that interactions were correlated with the main effect of lexicality (i.e., any interactions would be produced by a lexicality effect in similar 2-3 and 3-4 exchange conditions, see also note 13). For random effects, all possible intercepts and slopes for participants and items were included. As in Experiment 1, a likelihood ratio test was used for significance testing.

---

[11] There were almost twice as many trials in Experiment 1 than there were word and nonword trials in Experiment 2.

This analysis failed to show a significant effect of lexicality ($\chi^2 = 0.68$, $p > .1$).

Inspection of the data indicates large variance (>8.58) for random-effects intercepts and slopes

for similarity, and we included the effect of bi-mora frequency in the analysis to reduce the

random variance. Although this was a post-hoc analysis, an examination of the stimuli indicated

that the lexicality effect may have been influenced by differences in bi-mora frequency between

words and nonwords. As described earlier, the bi-mora frequency for (non)words was calculated

by summing the three bi-morae frequencies (i.e., bi-mora 1-2, bi-mora 2-3, and bi-mora 3-4).

Then, the summed value was matched between words and nonwords ($t(94) = 0.05$, $p = .961$).

However, given that each bi-mora frequency was a probabilistic variable, it may have been more

accurate to calculate the bi-mora frequency of (non)words by multiplying the three bi-mora

frequencies and logging the result. Notably, these values differed between words and nonwords

($t(94) = 2.01$, $p = .047$). Thus, we included this calculation of bi-mora frequency in the analysis.

The final model is described in Table 3. Phonological similarity and edge effects were

replicated. There was also a marginally significant effect of primacy gradient, which may

indicate a slight advantage for initial versus final positions. Although the analysis also showed a

significant effect of bi-mora frequency,[12] the lexicality effect remained non-significant. It may

not be surprising that error rate was sensitive to bi-mora frequency, as accumulating evidence

suggests that sublexical/phonotactic frequencies contribute to serial ordering (Gathercole,

Frankish, Pikkering, & Peaker, 1999; Gupta & Tisdale, 2009; Tamaoka & Makioka, 2009; Tanida,

Ueno, Saito, & Lambon Ralph, 2010; Vitevitch & Luce, 2004), and this confirms that the

technique used here taps into phonological planning. In contrast, the absence of a lexicality effect

---

[12] Re-analysis of the error rate in Experiment 1 including the new calculations of bi-mora frequency showed no effect of bi-mora frequency. Similarity and edge effects remained significant.

might be surprising considering accumulating evidence for the contribution of lexical/semantic

representations to phoneme ordering. However, the present results should be interpreted with

caution because the experimental design might not be optimal for capturing lexicality effects.

Most of the errors were produced in the similar 2-3 and 3-4 exchange conditions in which there

were numerically fewer errors for word targets (Fig. 3). Such higher order interactions might be

difficult to detect.[13] The present technique is relatively new and further experimentation is

needed on the lexicality effect in which, for example, only similar 2-3 and 3-4 exchange

conditions are included with more stimuli.

---

[13] An analysis in which the lexicality effect was only expressed by an interaction and separate
analyses for phonologically similar conditions also showed no lexicality effects.
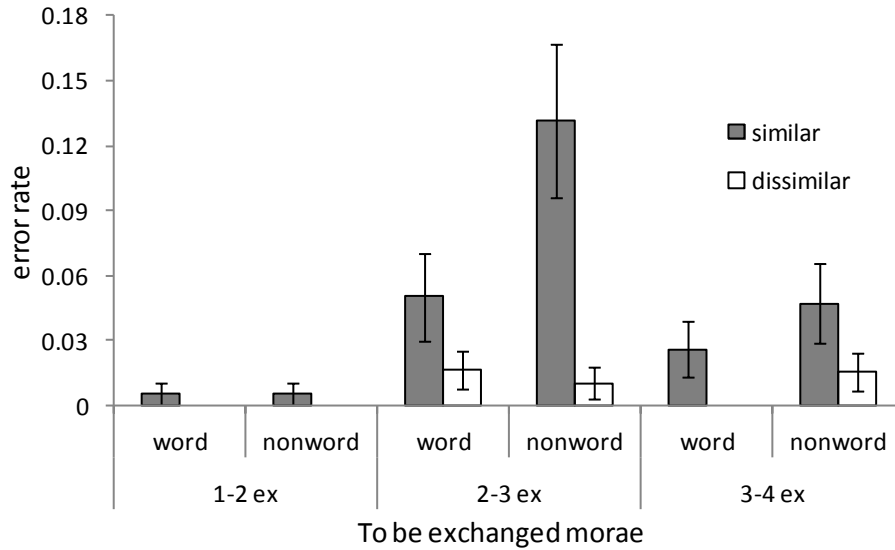
Fig. 3. Incorrect error rate as a function of phonological similarity, exchanged mora positions, and target lexicality in Experiment 2.

Note. Gray and white bars indicate mean raw error rate in the phonologically similar and dissimilar conditions, respectively. Error bars indicate standard errors of the means based on participant analysis.

Table 3. Results of logistic mixed-effects regression in Experiment 2.

Random effects

Groups		Variance	SD	Correlation				
Participant	(Intercept)	0.97	0.98					
	Phonological similarity	0.82	0.90	-0.24				
	Edge	0.00	0.03	-0.46	-0.76			
	Primacy gradient	0.04	0.21	0.15	0.93	-0.95		
	Lexicality	0.06	0.24	-0.04	0.98	-0.87	0.982	
	Bi-mora frequency	0.01	0.12	-0.92	0.60	0.07	0.259	0.436
Item	(Intercept)	9.71	3.12					
	phonological similarity	9.31	3.05	-0.93				

Fixed effects

	Estimate	SE	z value	Pr(>|z|)	χ^2	Pr(>χ^2)
(Intercept)	-7.69	1.11	-6.89	<.001		
Phonological similarity	2.84	1.11	2.56	<.05	5.57	<.05
Edge	0.69	0.20	3.44	<.001	9.64	<.01
Primacy gradient	0.98	0.51	1.93	<.1	2.87	<.1
Lexicality	0.06	0.25	0.22	>.1	0.04	>.1
Bi-mora frequency	-0.44	0.19	-2.39	<.05	3.92	<.05

3.2.2. Movement error rate

As in the previous experiment, the movement error rate varied as a function of the original target position. Movement error output at mora 2 was induced more by 2-3 (21/381 for errors from mora 3, which was the only source mora position) than by 1-2 exchanged distractors (1/382 for errors from mora 1 and 2/382 for those from any other mora positions; *ps* < .05, exact test). Movement errors output at mora 3 were induced more by 2-3 (20/381 for errors from mora 2 and 25/381 for those from any other mora positions) than by 3-4 exchanged distractors (7/382 for errors from mora 4, which was the only source mora position; *ps* < .05, exact test). These results confirm that errors were morae movement errors within target words.

In summary, Experiment 2 confirmed that this paradigm taps into the phonological planning process that is responsible for serial order control and further confirmed the governing principles of plan representations (i.e., temporal distance and edge principles), replicating adjacent mora exchange and serial position effects within a word.

4. General discussion

4.1. Summary of results

In two experiments, phonologically similar distractors induced exchange of adjacent morae and other phoneme movements in a repeated single-word production task in Japanese. Additionally, there was a serial position effect in within-word mora misordering, such that distractors whose second and third morae (i.e., the middle portion of the word) were exchanged in the target word induced more errors than 1-2 and 3-4 exchanged distractors. This pattern was

confirmed with nonword targets in Experiment 2, in which target bi-mora frequency also affected error rates, such that participants made more errors for (non)words with low bi-mora frequency. In the following sections, we discuss the implications of these results.

4.2. Phonological similarity

The overall phonological similarity effect confirms the principle of phonological similarity. Saito and Baddeley (2004) have already used this experimental technique to demonstrate phonological similarity effects. However, they only tested a limited number of stimuli (i.e., eight target words), which were selected from error-prone words in a language game. Thus, the target and distractor words they used to examine phonological similarity effects were not controlled for other variables. This limitation was overcome in the present study, which used more stimuli that were selected according to objective criteria, and which were carefully controlled for various psycholinguistic variables.

4.3. Temporal distance principle

Induced adjacent mora exchanges within a word are a clear demonstration of the within-word temporal distance principle. Models of serial order assume that the principle of temporal distance operates in both within- and between-word serial ordering. However, within-word adjacent-element (e.g., phoneme) exchanges have not been previously demonstrated empirically, resulting in a lack of direct evidence for this principle. The present results provide this missing evidence.

4.4. Edge principle

These are the first results to reveal a serial position effect in the context of single-word production, such that word edges were resistant to transposition. This indicates that word edges are represented distinctly from other positions, indicating a common edge similarity principle in speech production and pSTM. Notably, this effect was shown to be dissociated from the phonological/phonemic similarity principle, as phonemic similarity was controlled between the three exchange conditions (i.e., all phonemes were shared between targets and distractors).

4.5. Linguistic knowledge

The presence of a bi-mora frequency effect indicates that (phonological) linguistic knowledge contributes to serial ordering. Considering the ample evidence supporting the contribution of linguistic knowledge, this appears to be a general principle that determines the similarity/distinctness of positions/elements. More frequent sequences or their constituents are represented with a finer grain and are less susceptible to movement and other errors. This principle can provide an integrated account for some aspects of syllabic constraints, as discussed in the next section.

4.6. Toward a universal account of syllabic constraints

The successful induction of within-word adjacent-mora exchanges implies that syllabic constraints do not affect serial ordering in Japanese in the same way that they do in English. In English, syllabic constraints such as onset-rhyme organization (e.g., coda is relatively predictable based on vowels, but vowels are relatively unpredictable based on onset) decrease within-syllable similarity among positions/segments (and may increase between-syllable similarity). In Japanese, syllabic constraints primarily cluster around CVs (e.g., a consonant is usually followed by a vowel) as a sub-syllabic mora unit. This type of constraint increases the similarity between adjacent morae, especially when the mora structures (e.g., CV) are identical or similar (in fact, the majority of morae share a CV form and they are quite similar structurally). The syllabic constraint in English prevents exchanges of adjacent elements (i.e., phonemes or sub-syllabic units) and may encourage exchange of elements in the same within-syllable positions in different syllables/words (i.e., between-syllable or -word exchanges). In contrast, the syllabic constraints in Japanese encourage exchange of adjacent morae. These different effects of syllabic constraints can at least partly emerge from the different statistical structures in different languages (Chen, Dell, & Chen, 2004; Dell et al., 1993; Kessler & Treiman, 1997; Lee & Goldrick, 2008). In other words, some aspects of syllabic constraints are language-specific, but the mechanisms generating such constraints and their implementation are based on universal cognitive computational processing.

Language-specific statistical structure may not be the only source of syllabic constraints. If syllable representations are regarded as position representations (Vousden et al., 2000; see also Henson & Burgess, 1997), syllabic constraints may also be captured by the edge principle, which could be applied to any language, and is thus a universal principle. The distinctness of onset and

offset syllable positions should prevent phoneme misordering within a syllable and/or

monosyllabic words (Acheson & MacDonald, 2009a).

4.7. Integration and interaction between levels of similarity

Although the purpose of this study was to dissociate the temporal distance and edge

principles from other levels of similarity (i.e., whether and how temporal distance influences

serial ordering when the influence of syllabic constraints are minimized, and whether and how

serial position affects ordering when phonemic similarity is equivalent), the establishment of

these distinguishing principles indicates that these levels of similarity are integrated and interact

(e.g., Botvinick & Watanabe, 2007; Henson, 1998; Sevald & Dell, 1994). How within- and

between-word temporal information is integrated still requires investigation. Within-word

position representations may differentially influence within- and between-word phoneme

ordering if within-word positional representations are re-used for phoneme ordering in different

words, as many models assume (e.g., Dell, 1986; Vousden et al., 2000; see also Henson, 1998).

Consistent with this possibility, grouping elements in a list recall task decreased within-group

movement errors at the cost of increased between-group movement errors in the same within-

group positions (Henson, 1996; Experiment 2). Another theoretically interesting point is how

serial order control mechanisms, which are governed by temporal distance and edge principles,

operate in the acquisition of linguistic knowledge such as syllabic constraints, bi-mora frequency,

and lexical and semantic knowledge (Botvinick & Plaut, 2006; Burgess & Hitch, 2006; Dell et

al., 1993; Gupta & Tisdale, 2009; Plaut & Kello, 1999; Ueno et al., 2011). For example, serial

position modulates the effect of (phonological) linguistic knowledge (Hitch, Chiara Fastame, &

Flude, 2005), which requires further modeling (e.g., Botvinick & Plaut, 2006; Burgess & Hitch, 2006). Conversely, investigating the principles of serial order control mechanisms such as the edge principle could provide some clues as to how syllabic constraints such as onset-rhyme organization emerge over the course of vocabulary evolution (Gupta & Dell, 1999).

5. Conclusion

Serial ordering is influenced by similarity in certain aspects of language. Insofar as language is a temporal sequence that has a beginning and an end, the temporal distance and edge principles are general similarity principles that operate irrespective of ordering level (e.g., within- or between-word ordering) or language-specific element units (e.g., segment, mora, or syllable), even if they are hidden in some languages and reflect other similarity principles such as (language-specific) syllabic constraints. With a speech-error induction technique, we uncovered empirical evidence for these principles in within-word serial order control.

Acknowledgment

Appendix

Table A.1 Psycholinguistic variables of the stimuli in Experiment 1

			1-2 exchange	2-3 exchange	3-4 exchange
Target/Similar distracter	Word frequency (log10)	*Mean*	1.13	1.34	1.24
		SD	0.83	0.99	0.75
		Min	0.00	0.00	0.00
		Max	3.31	3.09	3.17
	Bi-mora frequency (log10)	*Mean*	5.80	5.82	6.03
		SD	0.33	0.43	0.44
		Min	4.90	4.99	5.10
		Max	6.60	6.61	6.73
Dissimilar distracter	Word frequency (log10)	*Mean*	1.14	1.34	1.28
		SD	0.82	0.99	0.77
		Min	0.00	0.00	0.00
		Max	3.31	3.09	3.20
	Bi-mora frequency (log10)	*Mean*	5.79	5.80	5.99
		SD	0.33	0.43	0.43
		Min	4.90	4.97	5.08
		Max	6.59	6.60	6.65
% Phoneme shared between target and dissimilar distracter		*Mean*	27.1	25.9	23.3
		SD	12.0	12.9	10.8
		Min	7.1	0.0	7.1
		Max	50.0	50.0	50.0

Table A.2 Stimuli for Experiment 1

Exchanged morae position	Target A	Target B	Dissimilar distracter A	Dissimilar distracter B
1-2 ex	i-da-te-N	da-i-te-N	ka-N-o-u	bu-N-bo-u
1-2 ex	i-ka-zo-ku	ka-i-zo-ku	ho-u-ta-i	ba-i-ni-N
1-2 ex	i-ku-bu-N	ku-i-bu-N	bo-u-ta-ka	i-chi-ga-N
1-2 ex	i-su-se-ki	su-i-se-ki	he-N-se-ki	ka-shi-chi-N
1-2 ex	i-su-to-ri	su-i-to-ri	a-i-ga-mi	ko-N-pa-ku
1-2 ex	i-ta-me-shi	ta-i-me-shi	ni-be-na-i	ka-shi-a-ge
1-2 ex	ka-na-ga-ki	na-ka-ga-ki	fu-ku-a-tsu	ra-i-ge-ki
1-2 ex	ka-sa-ga-mi	sa-ka-ga-mi	ta-e-nu-ki	de-shi-go-to
1-2 ex	ka-ta-a-shi	ta-ka-a-shi	mi-chi-ga-u	me-ga-shi-ra
1-2 ex	ka-wa-mu-ki	wa-ka-mu-ki	ya-do-ro-ku	ka-ke-o-chi
1-2 ex	ka-wa-se-ru	wa-ka-se-ru	he-N-bo-u	de-a-ge-ru
1-2 ex	ke-ta-ba-shi	ta-ke-ba-shi	ha-ya-shi-za	te-a-mi-ki
1-2 ex	ki-nu-i-to	nu-ki-i-to	e-do-ko-ro	so-bo-ku-sa
1-2 ex	ki-shi-me-N	shi-ki-me-N	a-i-e-su	ya-ma-sa-ka
1-2 ex	ki-shi-ra-su	shi-ki-ra-su	ko-tsu-ga-i	hi-ra-o-shi
1-2 ex	ki-ta-gu-chi	ta-ki-gu-chi	ka-sa-ne-gi	te-ga-ta-sa
1-2 ex	ki-wa-da-chi	wa-ki-da-chi	a-ra-ku-chi	ta-na-mu-ki
1-2 ex	ko-to-a-ge	to-ko-a-ge	u-chi-tsu-gi	ke-i-so-N
1-2 ex	ku-ro-da-i	ro-ku-da-i	se-N-no-u	da-ku-su-ru
1-2 ex	ku-ro-ma-ru	ro-ku-ma-ru	ha-tsu-a-ki	shi-tsu-ba-N
1-2 ex	ma-ro-N-ha	ro-ma-N-ha	de-su-gi-ru	za-ga-shi-ra
1-2 ex	o-shi-i-ri	shi-o-i-ri	do-ku-ga-N	he-N-a-i
1-2 ex	o-shi-ki-ri	shi-o-ki-ri	o-i-mi-zu	bu-ra-i-do
1-2 ex	o-shi-u-ri	shi-o-u-ri	ra-ku-ba-i	ga-N-a-tsu
1-2 ex	shi-me-a-ge	me-shi-a-ge	ka-ta-bo-ne	ku-ni-mo-chi
1-2 ex	su-mi-ga-ta	mi-su-ga-ta	o-ki-ga-sa	u-me-wa-ri
1-2 ex	ta-ma-ge-ru	ma-ta-ge-ru	ki-no-to-i	yo-mi-to-ki
1-2 ex	tsu-mi-mo-no	mi-tsu-mo-no	zo-ku-se-ki	ku-mi-u-chi
1-2 ex	u-ka-be-ru	ka-u-be-ru	de-mo-do-ru	ze-i-ta-ku
1-2 ex	u-o-he-N	o-u-he-N	mo-chi-ka-ke	i-e-da-ni
1-2 ex	zi-ma-wa-ri	ma-zi-wa-ri	tsu-ke-da-shi	ka-ri-a-to
1-2 ex	zi-tsu-do-u	tsu-zi-do-u	ko-ku-go-ka	ke-i-fu-ku
2-3 ex	bu-ra-N-ku	bu-N-ra-ku	o-bi-e-ru	a-wa-re-da
2-3 ex	e-ba-ra-su	e-ra-ba-su	te-u-su-da	u-o-to-ka
2-3 ex	fu-zi-ma-me	fu-ma-zi-me	ho-ga-ra-ka	hi-ra-nu-i
2-3 ex	ha-ka-na-ge	ha-na-ka-ge	yu-u-na-mi	ne-mu-ra-su

2-3 ex	ha-ma-ya-ki	ha-ya-ma-ki	o-mo-te-do	a-shi-nu-ki
2-3 ex	ha-ta-na-ka	ha-na-ta-ka	tsu-ku-ri-e	mo-no-i-re
2-3 ex	hi-ki-to-ri	hi-to-ki-ri	da-ku-te-N	na-N-po-u
2-3 ex	ho-ke-N-i	ho-N-ke-i	so-N-ta-i	sa-N-su-i
2-3 ex	i-o-to-su	i-to-o-su	te-ra-se-ru	shi-ma-a-zi
2-3 ex	i-shi-mu-ro	i-mu-shi-ro	za-N-bu-to	fu-N-su-i
2-3 ex	ka-bi-N-da	ka-N-bi-da	sa-N-shi-ki	su-i-ga-ra
2-3 ex	ka-i-ku-N	ka-ku-i-N	do-u-ke-tsu	se-i-ko-tsu
2-3 ex	ka-ki-ta-shi	ka-ta-ki-shi	bo-u-a-ku	se-ki-a-ku
2-3 ex	ka-na-ma-ri	ka-ma-na-ri	ya-su-mi-bi	zi-tsu-bu-shi
2-3 ex	ka-ta-mi-chi	ka-mi-ta-chi	sa-to-ka-ta	i-to-na-mi
2-3 ex	ka-ta-wa-re	ka-wa-ta-re	so-ku-ha-tsu	de-i-ga-N
2-3 ex	ki-ka-N-ku	ki-N-ka-ku	wa-ta-i-re	mi-ka-du-ki
2-3 ex	ko-i-sa-N	ko-sa-i-N	te-ki-shi-N	so-u-zi-N
2-3 ex	ko-u-ku-N	ko-ku-u-N	ta-N-so-u	fu-ri-ko-u
2-3 ex	ku-u-ra-N	ku-ra-u-N	o-ya-mo-to	ki-wa-ma-ru
2-3 ex	mi-ga-ru-i	mi-ru-ga-i	ge-N-bi-ki	ha-tsu-za-N
2-3 ex	mi-o-to-su	mi-to-o-su	ni-ba-N-me	to-ke-ko-mu
2-3 ex	mo-tsu-ya-ku	mo-ya-tsu-ku	fu-ta-ga-ru	i-shi-wa-ri
2-3 ex	o-i-u-tsu	o-u-i-tsu	ha-na-ga-sa	go-ma-N-to
2-3 ex	sa-i-ku-N	sa-ku-i-N	ho-u-ka-go	ke-N-pi-tsu
2-3 ex	shi-ta-ba-ki	shi-ba-ta-ki	nu-ki-yo-mi	ka-mi-u-ta
2-3 ex	su-o-do-ri	su-do-o-ri	fu-ri-ma-ku	a-ma-zo-ra
2-3 ex	ta-i-ku-tsu	ta-ku-i-tsu	ko-wa-se-ru	se-N-i-ki
2-3 ex	ta-tsu-ma-ki	ta-ma-tsu-ki	gu-ra-tsu-ku	ha-ka-do-ru
2-3 ex	yo-i-ko-to	yo-ko-i-to	fu-tsu-zi-N	ko-N-sa-i
2-3 ex	yo-ma-wa-ri	yo-wa-ma-ri	u-na-ba-ra	ka-ke-mo-chi
2-3 ex	zi-shi-ba-ri	zi-ba-shi-ri	mi-zu-gu-ki	a-do-o-N
3-4 ex	a-i-u-tsu	a-i-tsu-u	to-shi-o-i	ke-mi-su-ru
3-4 ex	a-ki-ku-sa	a-ki-sa-ku	mo-da-N-sa	ri-ku-go-u
3-4 ex	chi-N-u-tsu	chi-N-tsu-u	ho-i-ku-ki	bo-ru-shi-chi
3-4 ex	do-u-to-mo	do-u-mo-to	do-u-ha-i	ho-N-za-i
3-4 ex	fu-yu-ku-sa	fu-yu-sa-ku	ni-i-ku-sa	yo-shi-na-ni
3-4 ex	hi-ma-zi-N	hi-ma-N-zi	he-N-zi-ru	ga-ku-o-N
3-4 ex	hi-to-sa-shi	hi-to-shi-sa	i-N-su-u	mu-no-u-sa
3-4 ex	ho-N-ka-wa	ho-N-wa-ka	re-N-se-tsu	i-ba-re-ru
3-4 ex	ka-bu-da-N	ka-bu-N-da	ka-ku-re-ba	hi-to-tsu-me
3-4 ex	ka-i-ki-ri	ka-i-ri-ki	shi-se-i-zi	tsu-i-ta-te
3-4 ex	ka-i-ki-sa	ka-i-sa-ki	a-ka-nu-ke	me-N-ko-u
3-4 ex	ka-mi-na-bi	ka-mi-bi-na	so-de-ta-ke	he-pa-ri-N
3-4 ex	ka-N-i-da	ka-N-da-i	na-N-te-i	na-i-ku-u

3-4 ex	ki-ri-zu-mi	ki-ri-mi-zu	wa-ke-do-ri	tsu-chi-bu-ta
3-4 ex	ko-ro-shi-ba	ko-ro-ba-shi	do-ku-so-N	a-ya-se-ru
3-4 ex	ko-shi-ka-ta	ko-shi-ta-ka	e-bi-te-N	sa-ku-zo-u
3-4 ex	ko-u-ka-da	ko-u-da-ka	te-i-ta-sa	na-N-se-i
3-4 ex	me-N-i-ta	me-N-ta-i	ka-N-ge-tsu	ba-ku-me-i
3-4 ex	mu-shi-sa-N	mu-shi-N-sa	ke-i-fu-u	ni-chi-no-u
3-4 ex	o-ku-mi-ya	o-ku-ya-mi	ki-ro-su-u	ha-ga-ta-me
3-4 ex	o-ya-da-ma	o-ya-ma-da	na-su-gi-ru	hi-ri-ki-sa
3-4 ex	se-N-i-tsu	se-N-tsu-i	sa-ka-ra-i	sa-N-zo-u
3-4 ex	shi-bu-i-to	shi-bu-to-i	i-wa-sa-ki	shi-ga-ra-mu
3-4 ex	su-i-gu-N	su-i-N-gu	ko-yu-u-da	i-ya-se-ru
3-4 ex	te-N-i-ta	te-N-ta-i	ko-u-sa-tsu	zi-i-shi-ki
3-4 ex	to-u-ki-shi	to-u-shi-ki	ka-i-do-ri	te-ki-se-N
3-4 ex	ya-ki-shi-me	ya-ki-me-shi	ku-tsu-zu-re	mi-zu-u-ri
3-4 ex	ya-ma-ba-N	ya-ma-N-ba	so-N-ta-ku	o-gu-ra-i
3-4 ex	ya-ma-i-shi	ya-ma-shi-i	go-u-ma-N	mu-ko-N-da
3-4 ex	yo-i-mi-ya	yo-i-ya-mi	i-ra-tsu-me	ku-ma-ri-N
3-4 ex	yo-u-i-sa	yo-u-sa-i	ka-i-mu-da	ho-tsu-re-ru
3-4 ex	yu-u-i-da	yu-u-da-i	se-i-ba-tsu	so-u-he-i

Note: Target B served for target A as similar distracter and vice versa. Dissimilar distracter A was used for target A and B for B.

Table A.3 Psycholinguistic variables of the stimuli in Experiment 2

			1-2 exchange		2-3 exchange		3-4 exchange	
			word	nonword	word	nonword	word	nonword
Target	Word frequency (log10)	*Mean*	2.66	-	2.81	-	2.36	-
		SD	0.71	-	0.79	-	0.71	-
		Min	1.68	-	1.93	-	1.28	-
		Max	4.22	-	4.65	-	3.59	-
	Bi-mora frequency (log10)	*Mean*	5.97	5.96	5.92	5.93	6.10	6.08
		SD	0.38	0.35	0.39	0.39	0.36	0.36
		Min	5.49	5.50	5.24	5.24	5.61	5.52
		Max	6.60	6.52	6.54	6.59	6.64	6.65
Similar distracter	Word frequency (log10)	*Mean*	1.79	1.79	1.94	1.94	1.54	1.54
		SD	0.47	0.46	0.53	0.54	0.40	0.40
		Min	1.04	1.04	1.20	1.20	1.00	1.00
		Max	2.55	2.54	2.93	2.96	2.46	2.46
	Bi-mora frequency (log10)	*Mean*	5.82	5.84	5.86	5.88	6.11	6.09
		SD	0.37	0.38	0.44	0.43	0.38	0.39
		Min	5.06	5.02	4.99	5.04	5.35	5.33
		Max	6.46	6.45	6.51	6.52	6.63	6.61
Dissimilar distracter	Word frequency (log10)	*Mean*	1.79	1.79	1.94	1.94	1.55	1.54
		SD	0.47	0.46	0.54	0.53	0.39	0.40
		Min	1.04	1.04	1.20	1.20	1.00	1.00
		Max	2.54	2.53	2.93	2.95	2.46	2.46
	Bi-mora frequency (log10)	*Mean*	5.82	5.85	5.86	5.88	6.11	6.09
		SD	0.38	0.38	0.44	0.44	0.38	0.39
		Min	5.07	5.03	4.97	5.02	5.36	5.36
		Max	6.52	6.44	6.50	6.52	6.63	6.64
% Phoneme shared between target and dissimilar distracter		*Mean*	25.8	26.1	28.7	26.2	25.8	21.5
		SD	6.0	9.4	10.2	10.3	8.3	11.2
		Min	16.7	8.3	15.4	7.7	8.3	0.0
		Max	36.4	40.0	50.0	40.0	40.0	50.0

Table A.4 Stimuli for Experiment 2

Target lexicality	Exchanged morae position	Target	Similar distracter	Dissimilar distracter
word	1-2 ex	fu-ko-u-da	ko-fu-u-da	ka-tsu-ra-ku
word	1-2 ex	ka-i-zo-ku	i-ka-zo-ku	ba-i-ni-N
word	1-2 ex	ka-ta-ga-ki	ta-ka-ga-ki	u-so-tsu-ki
word	1-2 ex	ka-wa-mu-ki	wa-ka-mu-ki	ya-do-ro-ku
word	1-2 ex	ki-ta-ga-wa	ta-ki-ga-wa	u-tsu-ro-i
word	1-2 ex	ki-ta-gu-chi	ta-ki-gu-chi	a-to-ga-ma
word	1-2 ex	ku-ri-a-ge	ri-ku-a-ge	hi-ki-ni-ku
word	1-2 ex	ro-ku-da-i	ku-ro-da-i	sa-i-de-N
word	1-2 ex	ro-ma-N-ha	ma-ro-N-ha	za-ga-shi-ra
word	1-2 ex	shi-a-ga-ru	a-shi-ga-ru	shi-fu-zo-u
word	1-2 ex	shi-ki-me-N	ki-shi-me-N	ya-ma-sa-ka
word	1-2 ex	ta-ka-me-da	ka-ta-me-da	ga-ku-se-ki
word	1-2 ex	wa-ka-se-ru	ka-wa-se-ru	de-a-ge-ru
word	1-2 ex	wa-ka-ta-ke	ka-wa-ta-ke	ya-ku-za-da
word	1-2 ex	wa-su-re-ru	su-wa-re-ru	ra-i-to-u
word	1-2 ex	zi-tsu-do-u	tsu-zi-do-u	ke-i-ha-ku
word	2-3 ex	bu-N-ra-ku	bu-ra-N-ku	de-ki-ba-e
word	2-3 ex	hi-ki-to-ri	hi-to-ki-ri	da-ku-te-N
word	2-3 ex	ka-bi-N-da	ka-N-bi-da	shi-ga-na-i
word	2-3 ex	ka-wa-ki-ri	ka-ki-wa-ri	ho-ga-ra-ka
word	2-3 ex	ki-ka-N-ku	ki-N-ka-ku	ke-i-ne-N
word	2-3 ex	ko-u-ku-N	ko-ku-u-N	yo-u-be-N
word	2-3 ex	ku-ra-u-N	ku-u-ra-N	ki-wa-ma-ru
word	2-3 ex	mi-to-o-shi	mi-o-to-shi	su-i-na-N
word	2-3 ex	mi-to-o-su	mi-o-to-su	tsu-mi-ko-mu
word	2-3 ex	sa-ku-i-N	sa-i-ku-N	ta-ku-so-u
word	2-3 ex	su-do-o-ri	su-o-do-ri	a-ma-zo-ra
word	2-3 ex	ta-i-ku-tsu	ta-ku-i-tsu	ko-wa-se-ru
word	2-3 ex	ta-tsu-ma-ki	ta-ma-tsu-ki	gu-ra-tsu-ku
word	2-3 ex	tsu-yo-me-ru	tsu-me-yo-ru	i-to-na-mi
word	2-3 ex	yo-ko-i-to	yo-i-ko-to	ka-ra-ku-ni
word	2-3 ex	yo-ma-wa-ri	yo-wa-ma-ri	u-na-ba-ra
word	3-4 ex	a-ra-ka-ta	a-ra-ta-ka	ba-ka-mo-no
word	3-4 ex	chi-N-tsu-u	chi-N-u-tsu	bo-ru-shi-chi
word	3-4 ex	fu-tsu-go-u	fu-tsu-u-go	mi-so-ku-ra
word	3-4 ex	ha-na-bi-ra	ha-na-ra-bi	no-u-sa-i
word	3-4 ex	hi-ma-N-zi	hi-ma-zi-N	ga-ku-o-N
word	3-4 ex	ho-N-ka-wa	ho-N-wa-ka	tsu-u-zi-N

word	3-4 ex	ka-bu-da-N	ka-bu-N-da	sa-i-su-N
word	3-4 ex	ka-i-ki-ri	ka-i-ri-ki	ma-i-ru-do
word	3-4 ex	shi-ro-mi-zu	shi-ro-zu-mi	te-mi-zi-ka
word	3-4 ex	su-i-N-gu	su-i-gu-N	i-ya-se-ru
word	3-4 ex	ta-i-da-N	ta-i-N-da	sa-i-to-u
word	3-4 ex	ta-ma-shi-i	ta-ma-i-shi	ge-i-da-N
word	3-4 ex	te-N-ta-i	te-N-i-ta	ni-ku-me-ru
word	3-4 ex	to-ku-i-da	to-ku-da-i	to-u-ku-tsu
word	3-4 ex	yo-u-sa-i	yo-u-i-sa	ma-o-ta-i
word	3-4 ex	yu-u-i-da	yu-u-da-i	se-i-ba-tsu
nonword	1-2 ex	bu-mi-ru-i	mi-bu-ru-i	chi-ri-me-N
nonword	1-2 ex	chi-u-da-shi	u-chi-da-shi	e-pu-ro-N
nonword	1-2 ex	da-i-ke-ru	i-da-ke-ru	ho-u-chi-ku
nonword	1-2 ex	ka-ta-mi-ya	ta-ka-mi-ya	fu-ra-N-ki
nonword	1-2 ex	ki-i-shi-ni	i-ki-shi-ni	so-re-i-yu
nonword	1-2 ex	ku-a-go-u	a-ku-go-u	to-ri-ni-ku
nonword	1-2 ex	ku-da-su-ru	da-ku-su-ru	ri-ke-N-ya
nonword	1-2 ex	ku-tsu-se-ru	tsu-ku-se-ru	i-ta-N-shi
nonword	1-2 ex	ma-a-i-ro	a-ma-i-ro	ha-na-shi-ru
nonword	1-2 ex	o-shi-ga-ma	shi-o-ga-ma	so-ko-bi-e
nonword	1-2 ex	o-to-ra-su	to-o-ra-su	ba-a-sa-ma
nonword	1-2 ex	ra-sa-ke-ru	sa-ra-ke-ru	u-ka-tsu-da
nonword	1-2 ex	ra-u-na-i	u-ra-na-i	sa-i-shi-ki
nonword	1-2 ex	shi-mu-ke-N	mu-shi-ke-N	yo-u-ki-sa
nonword	1-2 ex	ta-i-wa-ru	i-ta-wa-ru	mu-na-mo-to
nonword	1-2 ex	te-o-ma-e	o-te-ma-e	ha-tsu-bo-shi
nonword	2-3 ex	a-ha-ka-ta	a-ka-ha-ta	u-ki-yo-e
nonword	2-3 ex	a-sa-to-ki	a-to-sa-ki	fu-ta-su-zi
nonword	2-3 ex	da-to-i-ku	da-i-to-ku	ya-ku-se-N
nonword	2-3 ex	ga-sa-ku-i	ga-ku-sa-i	sa-i-zi-tsu
nonword	2-3 ex	ge-fu-i-u	ge-i-fu-u	ri-ku-u-N
nonword	2-3 ex	ha-ka-i-ra	ha-i-ka-ra	se-i-he-ki
nonword	2-3 ex	ha-ta-ba-ki	ha-ba-ta-ki	ni-i-bo-N
nonword	2-3 ex	hi-ki-ri-da	hi-ri-ki-da	ha-i-hi-N
nonword	2-3 ex	me-ta-da-su	me-da-ta-su	u-ma-no-ri
nonword	2-3 ex	na-yo-ka-ku	na-ka-yo-ku	su-i-gi-N
nonword	2-3 ex	ne-bo-tsu-u	ne-tsu-bo-u	u-ro-tsu-ku
nonword	2-3 ex	no-ho-u-N	no-u-ho-N	yo-u-ga-sa
nonword	2-3 ex	sa-da-ka-i	sa-ka-da-i	bo-da-i-zi
nonword	2-3 ex	shi-a-ro-N	shi-ro-a-N	tsu-ku-ri-da
nonword	2-3 ex	su-ta-i-ku	su-i-ta-ku	i-chi-o-N
nonword	2-3 ex	u-ka-go-shi	u-go-ka-shi	mi-se-mo-no
nonword	3-4 ex	a-ma-to-o	a-ma-o-to	ni-za-ka-na

nonword	3-4 ex	fu-ku-N-se	fu-ku-se-N	do-u-za-i
nonword	3-4 ex	go-ku-u-do	go-ku-do-u	ke-i-zo-u
nonword	3-4 ex	i-to-ri-ki	i-to-ki-ri	hi-to-mu-ki
nonword	3-4 ex	ka-ta-chi-ku	ka-ta-ku-chi	ki-yo-bu-N
nonword	3-4 ex	ko-u-tsu-ke	ko-u-ke-tsu	pi-ka-i-chi
nonword	3-4 ex	mo-u-N-ki	mo-u-ki-N	na-ga-i-mo
nonword	3-4 ex	o-ki-shi-i	o-ki-i-shi	te-ki-me-N
nonword	3-4 ex	ra-ku-tsu-ge	ra-ku-ge-tsu	mo-u-be-N
nonword	3-4 ex	ri-me-shi-N	ri-me-N-shi	na-ki-hi-to
nonword	3-4 ex	se-i-ha-gi	se-i-gi-ha	to-u-ko-N
nonword	3-4 ex	se-i-u-fu	se-i-fu-u	ho-ko-u-ki
nonword	3-4 ex	so-u-tsu-ku	so-u-ku-tsu	he-i-sa-ku
nonword	3-4 ex	ta-ke-tsu-mi	ta-ke-mi-tsu	u-tsu-ra-su
nonword	3-4 ex	te-ka-N-ge	te-ka-ge-N	to-shi-tsu-ki
nonword	3-4 ex	yu-u-N-ke	yu-u-ke-N	mo-u-shi-N

References

Acheson, D. J., & MacDonald, M. C. (2009a). Verbal working memory and language production: Common approaches to the serial ordering of verbal information. *Psychological Bulletin*, *135*(1), 50–68. doi:10.1037/a0014411

Acheson, D. J., & MacDonald, M. C. (2009b). Twisting tongues and memories : Explorations of the relationship between language production and verbal working memory. *Journal of Memory and Language*, *60*(3), 329–350. doi:10.1016/j.jml.2008.12.002

Amano, S., & Kondo, T. (2000). *Japanese NTT database series: Lexical properties of Japanese, word-frequency (II).* Tokyo: Sanseido [in Japanese].

Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language*, *68*(3), 255–278. doi:10.1016/j.jml.2012.11.001

Botvinick, M. M. (2007). Multilevel structure in behaviour and in the brain: a model of Fuster's hierarchy. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences*, *362*(1485), 1615–26. doi:10.1098/rstb.2007.2056

Botvinick, M. M., & Plaut, D. C. (2006). Short-term memory for serial order: a recurrent neural network model. *Psychological Review*, *113*(2), 201–233. doi:10.1037/0033-295X.113.2.201

Botvinick, M. M., & Watanabe, T. (2007). From numerosity to ordinal rank: a gain-field model of serial order representation in cortical working memory. *The Journal of Neuroscience*, *27*(32), 8636–8642. doi:10.1523/JNEUROSCI.2110-07.2007

Brown, G. D. a., Preece, T., & Hulme, C. (2000). Oscillator-based memory for serial order. *Psychological Review*, *107*(1), 127–181. doi:10.1037//0033-295X.107.1.127

Burgess, N., & Hitch, G. J. (1999). Memory for serial order: A network model of the phonological loop and its timing. *Psychological Review*, *106*(3), 551–581.

Burgess, N., & Hitch, G. J. (2006). A revised model of short-term memory and long-term learning of verbal sequences. *Journal of Memory and Language*, *55*(4), 627–652. doi:10.1016/j.jml.2006.08.005

Chen, T.-M., Dell, G. S., & Chen, J.-Y. (2004). A cross-linguistic study of phonological units: Syllables emerge from the statistics of Mandarin Chinese, but not from the statistics of English. *Proceedings of the 26th annual conference of the Cognitive Science Society*.

Dell, G. S. (1986). A spreading-activation theory of retrieval in sentence production. *Psychological Review*, *93*(3), 283–321.

Dell, G. S. (1988). The retrieval of phonological forms in production: Tests of predictions from a connectionist model. *Journal of Memory and Language*, *27*(2), 124–142.

Dell, G. S. (1990). Effects of frequency and vocabulary type on phonological speech errors. *Language and cognitive processes*, *5*(4), 313–349.

Dell, G. S., Burger, L. K., & Svec, W. R. (1997). Language production and serial order: a functional analysis and a model. *Psychological Review*, *104*(1), 123–47.

Dell, G. S., Juliano, C., & Govindjee, A. (1993). Structure and content in language production: A theory of frame constraints in phonological speech errors. *Cognitive Science*, *17*(2), 149–195. doi:10.1016/0364-0213(93)90010-6

Elman, J. L. (1990). Finding structure in time. *Cognitive Science*, *14*(2), 179–211.

Fromkin, V. (1971). The non-anomalous nature of anomalous utterances. *Language*, *47*(1), 27–52.

García-Albea, J. E., del Viso, S., & Igoa, J. M. (1989). Movement errors and levels of processing in sentence production. *Journal of Psycholinguistic Research*, *18*(I), 145–161.

Garrett, M. F. (1975). The Analysis of Sentence Production. In G. H. Bower (Ed.), *The psychology of learning and motivation* (pp. 133–175). San Diego, CA: Academic Press.

Gathercole, S. E., Frankish, C. R., Pikkering, S. J., & Peaker, S. (1999). Phonotactic Influences on Short-Term Memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *25*(1), 84–95.

Gollan, T., & Goldrick, M. (2012). Does bilingualism twist your tongue? *Cognition*, *125*, 491–497.

Gupta, P. (2005). Primacy and recency in nonword repetition. *Memory*, *13*(3-4), 318–324. doi:10.1080/09658210344000350

Gupta, P., & Dell, G. S. (1999). The emergence of language from serial order and procedural memory. In B. MacWhinney (Ed.), *The Emergence of Language* (pp. 447–481). Mahwah, New Jersey: Lawrence Erlbaum Associates.

Gupta, P., Lipinski, J., Abbs, B., & Lin, P.-H. (2005). Serial position effects in nonword repetition. *Journal of Memory and Language*, *53*(1), 141–162. doi:10.1016/j.jml.2004.12.002

Gupta, P., & MacWhinney, B. (1997). Vocabulary acquisition and verbal short-term memory: computational and neural bases. *Brain and Language*, *59*(2), 267–333. doi:10.1006/brln.1997.1819

Gupta, P., & Tisdale, J. (2009). Does phonological short-term memory causally determine vocabulary learning? Toward a computational resolution of the debate. *Journal of Memory and Language*, *61*(4), 481–502. doi:10.1016/j.jml.2009.08.001

Hartley, T., & Houghton, G. (1996). A linguistically constrained model of short-term memory for nonwords. *Journal of Memory and Language*, *35*, 1–31.

Henson, R. N. A. (1996). *Short-term Memory for Serial Order*. *Unpublished doctrial dissertation*. University of Cambridge.

Henson, R. N. A. (1998). Short-term memory for serial order: the Start-End Model. *Cognitive Psychology*, *36*(2), 73–137. doi:10.1006/cogp.1998.0685

Henson, R. N. A., & Burgess, N. (1997). Representations of serial order. In J. Bullinaria, D. Glasspool, & G. Houghton (Eds.), *Proceedings of fourth neural computation and psychology workshop.* (pp. 283–300). London: Springer.

Henson, R. N. A., Norris, D. G., Page, M. P. A., & Baddeley, A. D. (1996). Unchained memory: Error patterns rule out chaining models of immediate serial recall. *The Quarterly Journal of Experimental Psychology:*, (1), 80–115.

Hitch, G. J., Chiara Fastame, M., & Flude, B. (2005). How is the serial order of a verbal sequence coded? Some comparisons between models. *Memory*, *13*(3-4), 247–258. doi:10.1080/09658210344000314

Houghton, G. (1990). The problem of serial order: A neural network model of sequence learning and recall. In R. Dale, C. Mellish, & M. Zocks (Eds.), *Cerrent Research in Natural Lnguage Generation* (pp. 287–319). London: Academic Press.

Jaeger, T. F. (2008). Categorical Data Analysis: Away from ANOVAs (transformation or not) and towards Logit Mixed Models. *Journal of memory and language*, *59*(4), 434–446. doi:10.1016/j.jml.2007.11.007

Jefferies, E., Frankish, C. R., & Lambon Ralph, M. a. (2006). Lexical and semantic binding in verbal short-term memory. *Journal of Memory and Language*, *54*(1), 81–98. doi:10.1016/j.jml.2005.08.001

Jefferies, E., Grogan, J., Mapelli, C., & Isella, V. (2012). Paced reading in semantic dementia: Word knowledge contributes to phoneme binding in rapid speech production. *Neuropsychologia*, *50*(5), 723–732. doi:10.1016/j.neuropsychologia.2012.01.006

Kessler, B., & Treiman, R. (1997). Syllable structure and the distribution of phonemes in English syllables. *Journal of Memory and Language*, *311*(37), 295–311.

Kohno, M. (1998). Mora, syllable and Rhythm - A Psycholimguistic Sutdy [Mora, onsetsu, rizumu no shinrigengogakutekikohsatsu]. *Journal of the Phonetic Society of Japan*, *2*(1), 16–24.

Kubozono, H. (1989). The mora and syllable structure in Japanese: Evidence from speech errors. *Language and Speech*, *32*(3), 249–278. doi:10.1177/002383098903200304

Lashley, K. (1951). The problem of serial order in behavior. In L. A. Jeffress (Ed.), *Cerebral Mechanisms in Behaviour: the Hixon Symposium* (pp. 112–136). New York: Wiley.

Lee, Y., & Goldrick, M. (2008). The emergence of sub-syllabic representations. *Journal of Memory and Language*, *59*(2), 155–168. doi:10.1016/j.jml.2008.03.002

Martin, N., & Saffran, E. M. (1997). Language and auditory-verbal short-term memory impairments: Evidence for common underlying processes. *Cognitive Neuropsychology*, *14*(5), 641–682.

Nakayama, M., Kajimura, S., Sugimoto, M., Kuraya, K., Inoue, M., Ishibashi, R., & Saito, S. (2012). Phoneme exchange in, serial-position effect on, and lexical/semantic contributions to single-word production: An investigation using speech-error induction techniques in Japanese. In *Poster presented at the 34th Annual Conference of the Cognitive Science Society.* Sapporo, Japan.

O'Seaghdha, P. G., & Chen, J.-Y. (2009). Toward a Language-General Account of Word Production : The Proximate Units Principle. In N. A. Taatgen & H. van Rijn (Eds.), *Proceedings of the 31st annual conference of the Cognitive Science Society* (pp. 68–73). Austin, TX: Cognitive Science Society.

O'Seaghdha, P. G., Chen, J.-Y., & Chen, T.-M. (2010). Proximate units in word production: phonological encoding begins with syllables in Mandarin Chinese but with segments in English. *Cognition*, *115*(2), 282–302. doi:10.1016/j.cognition.2010.01.001

Otake, T., Hatano, G., Cutler, A., & Mehler, J. (1993). Mora or Syllable? Speech Segmentation in Japanese. *Memory & Cognition*, *32*, 358–378.

Page, M. P. A., Madge, A., Cumming, N., & Norris, D. G. (2007). Speech errors and the phonological similarity effect in short-term memory: Evidence suggesting a common locus. *Journal of Memory and Language*, *56*(1), 49–64. doi:10.1016/j.jml.2006.09.002

Page, M. P. A., & Norris, D. (1998). The primacy model: a new model of immediate serial recall. *Psychological Review*, *105*(4), 761–781.

Patterson, K., Graham, N., & Hodges, J. (1994). The impact of semantic memory loss on phonological representations. *Journal of Cognitive Neuroscience*, 57–63.

Perea, M., Nakatani, C., & Leeuwen, C. Van. (2011). Transposition effects in reading Japanese Kana: Are they orthographic in nature? *Memory & cognition*, *39*, 700–707. doi:10.3758/s13421-010-0052-1

Perea, M., & Pérez, E. (2009). Beyond alphabetic orthographies : The role of form and phonology in transposition effects in Katakana Beyond alphabetic orthographies : The role of form and phonology in transposition effects in Katakana. *Language and Cognitive Processes*, *24*(1), 67–88. doi:10.1080/01690960802053924

Plaut, D. C., & Kello, C. T. (1999). The emergence of phonology from the interplay of speech comprehension and production: A distributed connectionist approach. In B MacWhinney (Ed.), *The Emergence of Language* (pp. 381–415). Hillsdale, NJ: Erlbaum.

R Core Team. (2013). R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing.

Rumelhart, D. E., Hinton, G. E., & Williams, R. J. (1986). Learning internal representations by error propagation. In D. E. Rumelhart, J. L. McClelland, & the P. research Group (Eds.), *Parallel distributed processing: Explorations in the microstructure of cognition. Volume 1: Foundations* (pp. 318–362). Cambridge, MA: MIT Press.

Saito, S., & Baddeley, A. D. (2004). Irrelevant sound disrupts speech production: exploring the relationship between short-term memory and experimentally induced slips of the tongue. *Quarterly Journal of Experimental Psychology: Human Experimental Psychology*, *57*(7), 1309–40. doi:10.1080/02724980343000783

Sevald, C. A., & Dell, G. S. (1994). The sequential cuing effect in speech production. *Cognition*, *53*, 91–127.

Tamaoka, K., & Makioka, S. (2004). Frequency of occurrence for units of phonemes, morae, and syllables appearing in a lexical corpus of a Japanese newspaper. *Behavior Research Methods*, *36*(3), 531–547.

Tamaoka, K., & Makioka, S. (2009). Japanese Mental Syllabary and Effects of Mora, Syllable, Bi-mora and Word Frequencies on Japanese Speech Production. *Language and Speech*, *52*(1), 79–112. doi:10.1177/0023830908099884

Tanida, Y., Ueno, T., Saito, S., & Lambon Ralph, M. A. (2010). Effects of accent typicality and phonotactic frequency on nonword immediate serial recall performance in Japanese. *Eleventh Annual Conference of the International Speech Communication Association*.

Terao, Y. (2002). *Iimachigai wa dooshite okoru? [How do speech errors occur?]*. Tokyo: Iwanami Shoten [in Japanese].

Ueno, T., Saito, S., Rogers, T. T., & Lambon Ralph, M. A. (2011). Lichtheim 2: Synthesizing Aphasia and the Neural Basis of Language in a Neurocomputational Model of the Dual

Dorsal-Ventral Language Pathways. *Neuron*, *72*(2), 385–396. doi:10.1016/j.neuron.2011.09.013

Van Casteren, M., & Davis, M. H. (2006). Mix, a program for pseudorandomization. *Behavior Research Methods*, *38*(4), 584–589.

Van Casteren, M., & Davis, M. H. (2007). Match : A program to assist in matching the conditions of factorial experiments. *Behavior Research Methods*, *39*(4), 973–978.

Vitevitch, M. S., & Luce, P. A. (2004). A web-based interface to calculate phonotactic probability for words and nonwords in English. *Behavior Research Methods*, *36*(3), 481–7.

Vousden, J. I., Brown, G. D. A., & Harley, T. A. (2000). Serial control of phonology in speech production: a hierarchical model. *Cognitive Psychology*, *41*(2), 101–75. doi:10.1006/cogp.2000.0739