Cognition 131 (2014) 415–430

Contents lists available at ScienceDirect

Cognition

journal homepage: www.elsevier.com/locate/COGNIT

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

Masataka Nakayama [a,b,*], Satoru Saito [a,*]

[a] Graduate School of Education, Kyoto University, Yoshida-honmachi, Sakyo-ku, Kyoto 606-8501, Japan
[b] Japan Society for the Promotion of Sciences, 8 Ichibancho, Chiyoda-ku, Tokyo 102-8472, Japan

ARTICLE INFO

Article history:
Received 12 November 2012
Revised 17 February 2014
Accepted 4 March 2014
Available online 2 April 2014

Keywords:
Speech error
Within-word exchange
Serial position effect
Speech production
Phonological short-term memory
Japanese

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.

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

---

* Corresponding authors. Address: Graduate School of Education, Kyoto University, Yoshida-honmachi, Sakyo-ku, Kyoto 606-8501, Japan (M. Nakayama). Tel.: +81 75 753 3067; fax: +81 75 753 3066.

E-mail addresses: nakayama.masataka.36x@st.kyoto-u.ac.jp (M. Nakayama), saito.satoru.2z@kyoto-u.ac.jp (S. Saito).

http://dx.doi.org/10.1016/j.cognition.2014.03.002

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, 1986; Dell, Burger, & Svec, 1997; 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, 1998; Henson & Burgess, 1997; 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 &