Journal of Phonetics 57 (2016) 1–20

Contents lists available at ScienceDirect

Journal of Phonetics

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

Research Article

The effect of phonetic context on the dynamics of intrusions and reductions

Anneke W. Slis*, Pascal H.H.M. Van Lieshout

Department of Speech Language Pathology, Oral Dynamics Lab, University of Toronto, 160-500 University Avenue, Toronto, Ontario, Canada M5G 1V7

A R T I C L E I N F O

Article history:
Received 17 September 2014
Received in revised form
3 April 2016
Accepted 4 April 2016

Keywords:
Intrusion
Reduction
Speech error
Articulatory Phonology
Task Dynamics
Phonetic context
Gesture

A B S T R A C T

Recent studies have described speech errors as articulatory movements intruding during target constrictions as well as reduced movements of these target constrictions. These errors were hypothesized to originate from self-organizing mechanisms underlying context-free gestural coordination. The current study investigates whether such self-organizing mechanisms and resulting speech errors are influenced by coproduction constraints, which are introduced by phonetic context. Speakers repeated CVC–CVC word pairs, differing in onsets and sharing their rhymes. “Phonetic context” was manipulated by changing the rhyme across different word pairs using specific vowel- and consonant-combinations. Vertical movements of tongue tip, tongue dorsum, and lower lip were recorded with Electro-Magnetic Articulography. The results revealed that the tongue dorsum as an intruding articulator showed more intrusions in front vowel context than in low back vowel context. In addition, this articulator showed more intrusions than the tongue tip in /æ/ context, and more intrusions than the lower lip in /ɪ/ context. Reductions did not demonstrate this effect of vowel context. However, both tongue dorsum and tongue tip reduced more than the lower lip. The findings are explained in terms of the amount of spatial overlap of gestures and intruding articulatory movements, as defined in Articulatory Phonology.

1. Introduction

For more than a century, speech errors have been a fruitful phenomenon for scientists to study how speakers manage to plan and produce their speech. Observing natural speech error corpora and collecting data in more controlled experiments supported the view that errors resulted from a misordering of speech units (typically phonemes) at an abstract cognitive level. This means that errors provided a window into cognitive processes underlying speech planning (see e.g., Dell, 1986; Dell & Sullivan, 2004; Fromkin, 1971; Kent, 1996; Levelt, Roelofs, & Meyer, 1999). The current study examines errors from a different angle by investigating if articulatory constraints at the level of production, imposed by overlapping gestures and shared articulators, influence the occurrence of speech errors.

1.1. *Background*

In general, a speech error is an utterance produced by the speaker that differs from the originally intended one. A common type of error is a sound exchange such as *teep a cape* for intended *keep a tape* (Fromkin, 1971). In this example, the /k/ from *cape* substitutes the /t/ in *tape* and vice versa. This error type provided evidence that cognitive discrete units, such as phonemes or features, switch places *before* they translate into articulatory movements (Dell & Sullivan, 2004; Fromkin, 1971; Levelt, Roelofs, & Meyer, 1999). Most evidence supporting this account comes from studies that transcribe errors perceptually, a method highly susceptible to listener biases. In many cases, these biases caused listeners to identify errors as phonemic in nature (cf. Cutler, 1981; Kent, 1996; McMillan, Corley, & Lickley, 2009; Pouplier & Hardcastle, 2005). Consequently, the possibility that articulatory factors contribute to errors had not been addressed quantitatively in detail until the rise of studies employing measures that are more objective.

These studies indicate that speech errors frequently result in physical events that do not correspond to discrete speech segments (see e.g., Frisch & Wright, 2002; Goldrick & Blumstein, 2006; Goldstein, Pouplier, Chen, Saltzman & Byrd, 2007; McMillan & Corley,

* Corresponding author.
E-mail address: anneke.slis@alum.utoronto.ca (A.W. Slis).

0095-4470/$ - see front matter © 2016 Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.wocn.2016.04.001

2010; Mowrey & MacKay, 1990). A series of kinematic studies, of particular interest for the current study, revealed several error patterns that necessitated a different approach from the traditional view on how to explain the origin of speech errors (e.g., Goldstein et al., 2007; Pouplier, 2008). In these studies, speakers repeated word pairs like *cop top* many times. They frequently activated a set of articulators, which constricted the vocal tract in a manner appropriate for a specific speech segment (for example a tongue dorsum constriction for /k/), simultaneously with an articulation that was not supposed to be activated (for example a tongue tip constriction during /k/)[1]. The behavior of this non-target articulation was called an "intrusion". These intrusions ranged from small to full activation of the non-target articulator and were thus frequently gradual in nature. Likewise, although less frequent, a so-called reduction occurred in which speakers partially or fully reduced the movements of a target constriction. Based on distributional characteristics, these intrusions and reductions differed significantly from movement patterns considered part of normal articulatory variability (see Section 2.4.1).

Several factors affected the number of intrusions and reductions: more errors appeared at fast speaking rates (Goldstein et al., 2007) and CVC syllables with different onsets but similar vowels and coda consonants showed more errors than CV syllables, which only shared vowels (Pouplier, 2008). In addition, the number of intrusions and reductions increased over the course of a trial. Most importantly for the current study, Goldstein et al. (2007) revealed that more errors occurred in repetitions of the words *kip tip* than *cop top*, suggesting that certain combinations of articulatory movements were more prone to these intrusions than others were. They speculated that, compared to the tongue shape during /ɑ/, the tongue shape during /ɪ/ was more compatible with /k/ and /t/ productions, resulting in more intrusions, which suggests that articulatory constraints affected the error patterns.

Fig. 1 shows an example of tongue dorsum and tongue tip movements from the current study when a participant produces the word pair *cap tap* at a fast speaking rate. Circled peaks in tongue tip and tongue dorsum trajectories indicate intruding movements.

1.2. Theoretical framework

The gradual nature of errors, the finding that errors accumulated over time in repetitive speech, and the effect of vowel-context is problematic for models that explain errors as misplaced, but at the articulatory level correctly produced, phonemes or features (see also Goldrick and Blumstein (2006), Goldstein et al. (2007), Mowrey and MacKay (1990)). One of the models that can account for the gradual nature of errors, however, is the cascading activation model in which activation is flowing down from one level to the other (Goldrick & Blumstein, 2006; Goldrick, Ross Baker, Murphy, & Baese-Berk, 2011; McMillan & Corley, 2010; McMillan et al., 2009). In case of speech errors, this model suggests that the abstract target as well as the non-target phonological representation is selected, (partially) activated and, after cascading down to the articulatory level, executed. It must be noted that McMillan and Corley assume that articulatory variability is thus not caused by motoric factors but by "the tight coupling between speech plans and their execution" (McMillan & Corley, 2010, p. 244), thus reflecting the activation level of several abstract phonological representations. So-called spreading activation models (see e.g., Dell, 1986; Sevald & Dell, 1994) can possibly further account for the build-up of intrusions. In these models, nodes, representing units such as lexemes, words, syllables, and phonemes, are activated through a mechanism of spreading activation flowing top-down as well as bottom-up (Dell, 1986). After selecting a node, activation decays. By assuming that the activation of a phonological representation only gradually decays, some residual activation will be left during the next unit. If syllables and phonemes are reactivated repeatedly, residual activation builds up, resulting in more intrusions (see also Katsika, Shattuck-Hufnagel, Mooshammer, Tiede, and Goldstein (2014). Although it may be theoretically possible to explain the error patterns within frameworks of cascading- and spreading activation, an in-depth discussion of these models is not feasible because it would largely be speculative. We are not aware of any papers that have directly addressed this issue in this theoretical framework and such an attempt would be far beyond the scope of this paper. Most importantly, even if the build-up of errors can be explained within spreading- and cascading-activation models, it does not provide a framework for explaining possible errors originating at the articulatory level (see e.g, Pouplier & Goldstein, 2010), which is the main factor investigated in the current study.

An alternative theory that can potentially address the presence of intrusions and reductions, and that provides a framework for hypotheses related to phonological as well as phonetic constraints on speech error patterns, is Articulatory Phonology and the embedded model of Task Dynamics (Browman & Goldstein, 1989; Goldstein, Byrd, & Saltzman, 2006; Saltzman & Munhall, 1989).

1.2.1. Articulatory Phonology (AP) and Task Dynamics (TD)

In AP, the basic primitive unit is the gesture that is a goal-directed coordinated action of a set of articulators constricting the vocal tract. The gesture embodies speech events at a macroscopic (phonological) as well as at a microscopic (phonetic) level of speech (Browman & Goldstein, 1995). At the macroscopic level, gestures involve abstract vocal tract tasks, such as labial, alveolar, or dorsal closure. At the microscopic level, the actions of the individual gestures translate into coordinated movements of a set of articulators (Goldstein et al., 2006; Saltzman & Munhall, 1989). For example, to produce a /t/ with the tongue tip gesture (macroscopic), the individual articulators "jaw", "tongue tip", and "tongue body" are coordinated such that they realize the correct constriction task (microscopic).

The underlying linguistic requirements of individual gestures, such as degree and location of a constriction, are context-independent; context-dependency is introduced by combining gestures into larger units such as speech segments, syllables or words. As a result, gestures can overlap temporally and spatially such that they either share articulators or differ in types of

[1] The current paper uses the terms "target constrictions" and "target gestures" to indicate an intended gesture/ target constriction and "non-target articulation/ articulator" to indicate a non-intended articulatory movement respectively.