



Dynamic action units slip in speech production errors [☆]

Louis Goldstein ^{a,b,1}, Marianne Pouplier ^{a,b,*,1},
Larissa Chen ^{a,b}, Elliot Saltzman ^{a,c}, Dani Byrd ^{a,d}

^a *Haskins Laboratories, 300 George Street, New Haven, CT 06511, USA*

^b *Yale University, Department of Linguistics, 370 Temple Street, New Haven, CT 06511, USA*

^c *Boston University, Department of Physical Therapy and Athletic Training, 635 Commonwealth Ave., Boston, MA 02115, USA*

^d *University of Southern California, Department of Linguistics, 3601 Watt Way, Los Angeles, CA 90089, USA*

Received 3 June 2005; revised 16 May 2006; accepted 19 May 2006

Abstract

In the past, the nature of the compositional units proposed for spoken language has largely diverged from the types of control units pursued in the domains of other skilled motor tasks. A classic source of evidence as to the units structuring speech has been patterns observed in speech errors – “slips of the tongue”. The present study reports, for the first time, on kinematic data from tongue and lip movements during speech errors elicited in the laboratory using a repetition task. Our data are consistent with the hypothesis that speech production results from the assembly of dynamically defined action units – gestures – in a linguistically structured environment. The experimental results support both the presence of gestural units and the dynamical properties of these units and their coordination. This study of speech articulation shows that it is possible to develop a principled account of spoken language within a more general theory of action.

© 2007 Elsevier B.V. All rights reserved.

[☆] This manuscript was accepted under the editorship of Jacques Mehler.

* Corresponding author. Present address: University of Edinburgh, Adam Ferguson Building, 40 George Square, Edinburgh EH8 9LL, UK. Tel.: +44 131 6506657.

E-mail address: pouplier@haskins.yale.edu (M. Pouplier).

¹ These authors have contributed equally to this paper and are in alphabetical order.

Keywords: Speech errors; Action units; Entrainment; Speech production

1. Introduction

While humans perform skilled acts of motor control in many domains, none have the communicative and information encoding properties of language. This special status of language has no doubt contributed to accounts of speech that make it appear to be quite different from other kinds of coordinated action. While phonological research is built on the insight that words are composed of combinatorial units of information such as features or segments, finding a physical basis for compositional units in spoken language, either in the articulatory actions humans perform when speaking or in the resultant acoustic signal, has proven elusive. There is no obvious discrete division of the speech acoustic signal into compositional units; there is no analogue to the spaces between the letters of written language or a phonetic transcription. However, it has been proposed that when articulatory kinematics are examined within a dynamical systems framework, it becomes possible to identify compositional action units as atomic units of speech production and to understand how general dynamical principles that apply to skilled action generally also shape the activity of speaking (Browman & Goldstein, 1986; Fowler, Rubin, Remez, & Turvey, 1980).

One reason for the difficulty in identifying units in speech production is the difficulty of observing the speech articulators – such as the tongue – in action. Some evidence for action units has been found in studies using experimentally induced mechanical perturbation of articulators during running speech. These have shown that certain vocal tract articulators cohere systematically in the production of particular information units. For example, the upper lip, lower lip, and jaw appear to work cooperatively (and compensatorily) to achieve lip closure (Kelso, Tuller, Vatikiotis-Bateson, & Fowler, 1984). Similar studies using phase-resetting techniques have demonstrated compensation in the temporal domain (Saltzman, Löfqvist, Kay, Kinsella-Shaw, & Rubin, 1998).

Based in part on such studies (as well traditional phonological investigations into how words in a language are systematically differentiated from one another and how they are modified when they are produced in different contexts), it has been proposed that utterances can be decomposed into sets of dynamically defined units of constriction action, called gestures (Browman & Goldstein, 1992, 1995). A gesture orchestrates the movements of several articulators (e.g., upper lip, lower lip and jaw) in order to achieve a linguistically significant goal (e.g., lip closure). Gestures are modeled as point attractors in a task space – where the attractors are, in their simplest form, characterized as critically damped mass-spring systems, and the task space dimensions are defined in terms of constrictions that can be created and released by the independently controllable constricting organs of the vocal tract (Saltzman & Munhall, 1989). The gestures comprising an utterance have activations that wax and wane over time as the corresponding task-space attractors come into and out of existence in the vocal tract. Simultaneously with being compositional units of action,

Download English Version:

<https://daneshyari.com/en/article/927216>

Download Persian Version:

<https://daneshyari.com/article/927216>

[Daneshyari.com](https://daneshyari.com)