Available online at www.sciencedirect.com

ScienceDirect

Speech Communication 57 (2014) 283–300

SPEECH COMMUNICATION

www.elsevier.com/locate/specom

Effects of perturbation and prosody on the coordination of speech and gesture

Heather Leavy Rusiewicz [a,*,1], Susan Shaiman [b], Jana M. Iverson [c], Neil Szuminsky [b]

[a] Duquesne University, Department of Speech-Language Pathology, 412 Fisher Hall, 600 Forbes Avenue, Pittsburgh, PA 15282, USA
[b] University of Pittsburgh, Department of Communication Science and Disorders, 4033 Forbes Tower, 3600 Forbes at Atwood Street, Pittsburgh, PA 15260, USA
[c] University of Pittsburgh, Department of Psychology, 3415 Sennott Square, 210 South Bouquet Street, Pittsburgh, PA 15260, USA

Available online 14 June 2013

Abstract

The temporal alignment of speech and gesture is widely acknowledged as primary evidence of the integration of spoken language and gesture systems. Yet there is a disconnect between the lack of experimental research on the variables that affect the temporal relationship of speech and gesture and the overwhelming acceptance that speech and gesture are temporally coordinated. Furthermore, the mechanism of the temporal coordination of speech and gesture is poorly represented. Recent experimental research suggests that gestures overlap prosodically prominent points in the speech stream, though the effects of other variables such as perturbation of speech are not yet studied in a controlled paradigm. The purpose of the present investigation was to further investigate the mechanism of this interaction according to a dynamic systems framework. Fifteen typical young adults completed a task that elicited the production of contrastive prosodic stress on different syllable positions with and without delayed auditory feedback while pointing to corresponding pictures. The coordination of deictic gestures and spoken language was examined as a function of perturbation, prosody, and position of the target syllable. Results indicated that the temporal parameters of gesture were affected by all three variables. The findings suggest that speech and gesture may be coordinated due to internal pulse-based temporal entrainment of the two motor systems.

Keywords: Gesture; Perturbation; Prosody; Entrainment; Speech-hand coordination; Capacitance

1. Introduction

The pervasive, but still primarily subjective, finding that speech and gesture temporally overlap is frequently cited as a theoretical pillar of proof regarding the integrative nature of the two systems (e.g., Goldin-Meadow, 1999; Iverson and Thelen, 1999; Krauss et al., 2000; McNeill, 1992). The common observation is that gesture and speech roughly occur at similar times during communication. Even though many individuals observe and state that speech and gesture are produced in tight temporal "synchrony", the mechanism responsible for this potential temporal alignment as well as factors that may affect the coordination of these two actions have not been elucidated. Although recent investigations have examined the effect of variables such as prosodic prominence upon the execution of gestures, there is still relatively sparse and mostly descriptive data on the impact of perturbation of the speech system on the gestural movement, and vice versa. This study seeks to explore the effect of perturbation of the speech stream in addition to the effect of prosodic prominence upon the temporal parameters of corresponding gestures in a controlled paradigm with the goal to inform theoretical postulations regarding the speech and gesture production and add to the growing literature base on the interaction of these two systems.

* Corresponding author. Tel.: +1 412 396 4205; fax: +1 412 396 4196.
E-mail address: rusiewih@duq.edu (H.L. Rusiewicz).

[1] This data were collected in partial fulfillment of doctoral studies at the University of Pittsburgh, Pittsburgh, Pennsylvania.

0167-6393/$ - see front matter © 2013 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.specom.2013.06.004

1.1. Theoretical accounts of speech–gesture temporal alignment

Despite the relatively recent surge of interest regarding the study of gesture, theoretical postulation regarding the interface of speech and gesture is limited and primarily rooted in a linguistic perspective (e.g., Krauss et al., 2000; McNeill, 1992; de Ruiter, 1998; de Ruiter, 2000). However, recent experimental findings that gestures align with prosodically stressed syllables (de Ruiter, 1998; Esteve-Gibert and Prieto, 2011; Leonard and Cummins, 2011; Rochet-Capellan et al., 2008; Rusiewicz et al., 2013) cannot be easily incorporated into a linear, linguistic-based theory of gesture and spoken language production. Nor can such theories incorporate the findings of Krahmer and Swerts (2007) who found that the production of gestures increased the acoustic prominence of syllables, even when the participants were instructed not to stress the target syllable. Likewise, a cognitive-linguistic account cannot explain the limited literature on speech-gesture perturbations which indicates that speech and gesture remain synchronized or alter the timing of movements when the timing of one of the movements is halted in some way, whether it be as a consequence of a speech error (de Ruiter, 1998), moments of stuttering (Mayberry and Jaques, 2000; Mayberry, 1998), delayed auditory feedback (McNeill, 1992), or perturbing the early part of pointing gestures (Levelt et al., 1985).

Recovery from perturbation can be examined to study the stability of a motor behavior, including speech production. Perturbation studies also provide abundant information regarding the role of sensory feedback for motor control. There is a wealth of data on the sensory information that is utilized by the motor system for upper limb movements. For instance the visual location of a target can be altered during reaching tasks to measure the compensation of movement (e.g., Paulignan et al., 1991; Prablanc and Martin, 1992). In addition, perturbation studies have demonstrated the importance of both visual feedback and proprioceptive sensory information for accurate pointing trajectories (e.g., Bard et al., 1999; Komilis et al., 1993. Typically, this literature is interpreted as evidence for the remarkable adaptability of the motor system as well as the dynamic coordination of structures. Even though examining perturbations can offer insight into the motoric coordination of speech and gesture, this relatively straight-forward paradigm has rarely been utilized as a tool to examine the temporal relationship of speech and gesture. In fact, a systematic investigation of the temporal coordination of speech and gesture following speech perturbation has yet to be conducted.

Though gestures certainly are a unique class of manual movements with the potential for substantial overlay of linguistic information, it is necessary to also consider the underlying motoric form of gestures and the similarity of the interactions of the motor processes for speech and manual movements, gesture or otherwise. For instance, movements such as finger tapping, align prosodically strong syllables and altering temporal parameters as a result of the manipulations in the speech stream (Franz et al., 1992; Kelso et al., 1981; Smith et al., 1986). Another interesting and relevant line of work was embarked upon by Gentilucci and colleagues (e.g., Gentilucci et al., 2001, 2004) which revealed the size of mouth opening and acoustic amplitude of syllables is influenced by the size of grasping movements. Taken together, these data imply that (1) the alteration of the temporal parameters of speech and manual movements due to perturbations as well as (2) the temporal coordination of prosodically prominent syllables and manual movements is at least in part explicable according to a motoric perspective.

1.1.1. Dynamic systems theory and entrainment

When contemplating the motoric coordination of speech and gesture, it is important to consider gesture first as a hand movement that can also be overlaid with language functions. As such, dynamic systems theory (DST) potentially offers the best launching pad for positing and testing the motoric coordination of speech and gesture. In recent decades, nonlinear science focusing on the coordinative structures and processes has emerged as a viable alternative to traditional top-down, executive controlled theories. The application of nonlinear dynamics stemmed originally from work in the area of physics, and later, the coordination of motor control within the human body (see [204] for a review). However, principles of nonlinear dynamics are now being applied to just about any behavior or event, from cognitive-linguistic processing (e.g., Thelen and Smith, 2002) to inanimate properties like biological and chemical processes (e.g., Lorenzo et al., 2007), but not often considered in relation to gestures. Indeed, the applications and specifications of DST even within the particular realms of speech (see van Lieshout, 2004 for a review), limb movement (see Schmidt and Lee, 1999 for review) and even the coordination of speech and manual movements like tapping (e.g., Kelso et al., 1981) are certainly vast and varied and beyond the scope of this manuscript. Hence, the present investigation is focused on studying the phenomenon of *entrainment* within a DST framework.

The idea of entrainment is certainly not new or specific to gestures, yet it is underrepresented in both the literature on spoken language production (though refer Franz et al., 1992; Kelso et al., 1981; Smith et al., 1986; van Lieshout, 2004) and gesture production. The mechanism of entrainment was first presented 70 years ago by von Holst (1973) in response to his observations of the mutual influence of fin movements of swimming fish. The general idea of entrainment is that the preferred temporal pattern of one oscillator will interface with the preferred temporal pattern of the second oscillator, resulting in either an identical rhythmic pattern or a compromise rhythmic pattern somewhere in between the two patterns relative to when they are produced in isolation.

Similarly, Thelen and Smith (2002) iterated "to the degree that two component systems have a history of time-locked activity, they will come to entrain each other