Journal of Phonetics 55 (2016) 53–77

Contents lists available at ScienceDirect

Journal of Phonetics

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

Research Article

Selection and coordination: The articulatory basis for the emergence of phonological structure

Sam Tilsen

Cornell University, 203 Morrill Hall, Ithaca, NY 14853, USA

ARTICLE INFO

Article history:
Received 27 May 2015
Received in revised form
27 October 2015
Accepted 30 November 2015
Available online 17 December 2015

Keywords:
Speech production
Motor control
Motor feedback
Speech development
Language acquisition
Articulatory phonology
Phonological typology

ABSTRACT

Phonological theories commonly analyze speech utterances as composed of hierarchically organized units, such as features/gestures, segments, moras, and syllables, but it is not well understood why this hierarchical organization is observed. Moreover, current phonological theories and speech production models fail to explain cross-linguistic and developmental variation in the organization of units. This paper presents the *selection-coordination theory* of speech production, which attempts to unify our understanding of developmental and cross-linguistic variation in phonological structure. The theory holds that hierarchical organization emerges from a recurring trend in speech development whereby children acquire coordinative regimes of control over articulatory gestures that were previously competitively selected. In this framework, segments, moras, and syllables are understood as differently-sized instantiations of the same type of motor planning unit, and cross-linguistic and developmental phonological patterns are derived from distinguishing competitive and coordinative regimes of articulatory control. Evidence for the theory is drawn from research in motor control, speech development, and phonological and phonetic patterns in speech.

1. Introduction

This paper describes the selection-coordination theory of speech production, and explains how the theory helps us understand cross-linguistic and developmental phonological patterns. The gist of the theory is the interaction of two motor control mechanisms—selection and coordination—results in two prototypical articulatory control regimes: competitive control and coordinative control. Phonological structure involving segments, moras, and syllables emerges through developmental transitions from competitive to coordinative control, mediated by the internalization of sensory feedback.

The problems which motivate the development of selection-coordination theory can be summarized as three ways in which current approaches are insufficiently explanatory: (1) *Hierarchical structure* of utterances is not explained by formal representations or existing production models; these models presuppose structure rather than motivating it from cognitive mechanisms. (2) *Cross-linguistic variation* in phonological structure is described but not explained by current approaches; mechanisms for understanding how or why such variation emerges are absent. (3) *Developmental patterns* in speech are not explained, nor related to other forms of variation; existing approaches are either too rigidly structured or lack the necessary mechanisms to generate empirically observed phenomena.

Here we show that selection-coordination theory addresses these problems and provides a unified account of articulatory control, developmental patterns, and phonological structure. The theory elaborates on ideas which have been implemented computationally and to a limited extent tested experimentally in previous work (Tilsen, 2013, 2014a); here the ideas are developed in detail and applied to a broad range of phenomena, although the scope of the theory is such that treatment of all relevant issues is not possible. The manuscript is organized as follows:

Section 1 discusses how current frameworks are either too rigidly structured (in the case of hierarchical production models) or lack the necessary mechanisms for understanding hierarchical structure (in the case of Articulatory Phonology).

E-mail address: tilsen@cornell.edu

0095-4470/$ - see front matter © 2015 Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.wocn.2015.11.005

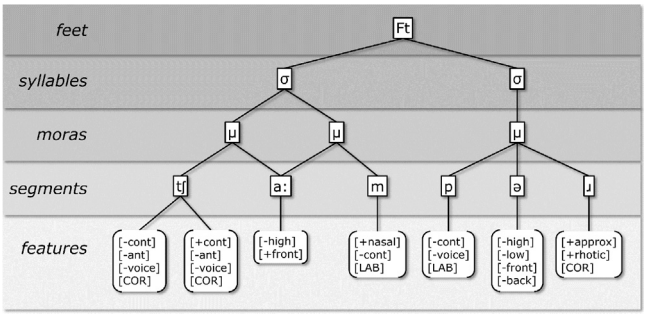

Fig. 1. Hierarchical representation of the word *chomper* illustrating various structural non-isomorphisms: multiple sets of features associated with one segment, multiple segments associated with one mora, and multiple moras associated with one segment.

Section 2 presents the theory in detail, describing the mechanisms of selection and coordination, their interaction, and the role of sensory feedback in developmental transitions from competitive to coordinative control.

Section 3 shows how the theory helps us understand phonetic and phonological patterns in development.

Sections 4 and 5, focusing on segments and moras/syllables respectively, provide a new understanding of phonological structure and apply selection-coordination theory to shed light on phonological and phonetic variation.

1.1. Hierarchical structure in phonological representations

Most phonological theories analyze speech utterances into hierarchically organized structures that consist of features or gestures, segments, moras, syllables, feet, and phrases. Problematically, current theories presuppose that these units are essential components of the representation, failing to satisfactorily motivate the existence of levels in the hierarchy. A hierarchically structured representation of the word “chomper” is shown in Fig. 1. The representation consists of units which are organized into spatially distinct levels or tiers. Representations of this sort rely on several conceptual metaphors (cf. Tilsen, 2009b): speech sounds and groups of speech sounds are objects, time is a space in which objects are located, and functional associations between sounds and groups of speech sounds are connections between objects. Consequently formal phonological theories have aimed to elucidate principles and constraints on the presence, absence, ordering, and connections between units.

Formal phonological representations dissociate the content, ordering, and representations of segments, thereby allowing for various non-isomorphisms as shown in Fig. 1. For the affricate [tʃ] a single segment is associated with multiple sets of features (Lombardi, 1990; Sagey, 1986), for the long vowel [aː] one set of features is associated with multiple moras, and all three moras are associated with multiple segments (Hyman, 1984, 1985; McCarthy & Prince, 1986).

Hierarchical phonological representations and non-isomorphisms between levels provide useful symbolic/geometric metaphors for *describing* phonological patterns but do not imply very much about cognitive mechanisms that could *explain* how such patterns arise. Phonological representations make only very limited predictions regarding the physical realizations of utterances; such predictions are generally viewed as the province of speech production models. Below we distinguish two general classes of such models.

1.2. Hierarchically structured production models

Most models of speech production are hierarchically structured and assume that segments and syllables play key roles in organizing speech motor planning. Importantly, none of these models incorporates a moraic level of motor organization, and they have no mechanisms for restructuring the network of units over time.

One example is the spreading activation model of Dell (1986), which views morphemes, syllables, syllable sub-constituents, segments, and features as nodes in a hierarchically structured network through which activation spreads. The selection of items on one level of representation is preceded by the selection of items on the next highest level. Thus the representation of a speech plan is constructed in a cascade, where the unit selected on a given level is the one with the highest amount of activation. The model provides an account of sequencing errors, and related models (Goldrick & Blumstein, 2006; Goldrick, Baker, Murphy, & Baese-Berk, 2011) have incorporated more detail regarding how spreading activation mechanisms interact with articulatory processes.

The gradient-ordering, directions into velocities of articulators model (GODIVA: Bohland, Bullock, & Guenther, 2010), an extension of the DIVA model (Guenther, Ghosh, & Tourville, 2006), views speech planning as a process in which segments are competitively selected to fill positions in syllable frames. Articulations are produced sequentially in association with segments via a mapping from sensory targets to motor commands, in conjunction with feedforward and feedback controllers.

The production model of Levelt (1993) presents a similar view of speech planning as the filling of timing slots by segmental material. Abstract plans are specified in the lexicon as a sequence of segments. A phonological encoding module groups segments into syllable-sized units and an articulatory module implements gestures with timing patterns determined by their syllabification. In recent versions of the model, moraic structure has been cited as “an emerging property of the syllabification process” (1999: 21), but without elaboration of how this occurs.