cortex 69 (2015) 24–39

Available online at www.sciencedirect.com

ScienceDirect

Journal homepage: www.elsevier.com/locate/cortex

Research report

How much is a word? Predicting ease of articulation planning from apraxic speech error patterns

Wolfram Ziegler* and Ingrid Aichert

EKN – Clinical Neuropsychology Research Group, Clinic for Neuropsychology, City Hospital, Munich, Germany

ARTICLE INFO

Article history:
Received 4 September 2014
Reviewed 12 January 2015
Revised 27 February 2015
Accepted 1 April 2015
Action editor Roberto Cubelli
Published online 15 April 2015

Keywords:
Speech apraxia
Articulation
Motor planning
Model
Syllable
Prosody

ABSTRACT

Background: According to intuitive concepts, 'ease of articulation' is influenced by factors like word length or the presence of consonant clusters in an utterance. Imaging studies of speech motor control use these factors to systematically tax the speech motor system. Evidence from apraxia of speech, a disorder supposed to result from speech motor planning impairment after lesions to speech motor centers in the left hemisphere, supports the relevance of these and other factors in disordered speech planning and the genesis of apraxic speech errors. Yet, there is no unified account of the structural properties rendering a word easy or difficult to pronounce.
Aim: To model the motor planning demands of word articulation by a nonlinear regression model trained to predict the likelihood of accurate word production in apraxia of speech.
Method: We used a tree-structure model in which vocal tract gestures are embedded in hierarchically nested prosodic domains to derive a recursive set of terms for the computation of the likelihood of accurate word production. The model was trained with accuracy data from a set of 136 words averaged over 66 samples from apraxic speakers. In a second step, the model coefficients were used to predict a test dataset of accuracy values for 96 new words, averaged over 120 samples produced by a different group of apraxic speakers.
Results: Accurate modeling of the first dataset was achieved in the training study ($R^2_{adj} = .71$). In the cross-validation, the test dataset was predicted with a high accuracy as well ($R^2_{adj} = .67$). The model shape, as reflected by the coefficient estimates, was consistent with current phonetic theories and with clinical evidence. In accordance with phonetic and psycholinguistic work, a strong influence of word stress on articulation errors was found.
Conclusions: The proposed model provides a unified and transparent account of the motor planning requirements of word articulation.

1. Introduction

Articulating a word of one's native language is not particularly expensive in terms of motor planning costs, – at least for neurologically healthy adults. German speakers won't experience much difference in articulatory planning efforts between saying, for instance, [haI] (shark), [kvalm] (fumes), or [ʃɪmpanzə] (chimpanzee), because they have acquired and

* *Corresponding author*. EKN, Klinikum Schwabing, Haus 19, Kölner Platz 1, 80804 München, Germany.
E-mail address: wolfram.ziegler@ekn-muenchen.de (W. Ziegler).
http://dx.doi.org/10.1016/j.cortex.2015.04.001

extensively practiced the motor routines generating the sounds of their language over and over, so that any German-sounding vocal utterance gets across their lips equally easily, with little effort, and without noticeable error rates.

Patients with apraxia of speech after lesions to left anterior peri- or subsylvian cortical areas are different in this regard. Their speech is dysfluent, and halting, with sound distortions (i.e., phonetically ill-formed sounds) and phoneme errors, with visible and audible groping, and with many false starts, restarts, and self-corrections, and these symptoms are considered to express a loss of the speech motor planning skills they have acquired during childhood and practiced ever since until their stroke (McNeil, Robin, & Schmidt, 2009; Ziegler, 2008; Ziegler, Aichert, & Staiger, 2012). The left hemisphere brain regions responsible for apraxia of speech (AOS), which are activated regularly in functional imaging studies of speech production, are considered to house the motor planning centers for speech movements (e.g., Bohland & Guenther, 2006; Brendel et al., 2011; Shuster & Lemieux, 2005).

Apraxic speakers are known to be sensitive to the phonological make-up of words. Their likelihood of making an error is higher, for instance, on /ʃɪmpanzə/ or /kvalm/ than on /haI/ (e.g., Romani & Galluzzi, 2005). Hence, to the extent that their speech errors reflect a problem of phonetic planning, accuracy data from apraxic speakers can be informative about the motor planning demands of a word as a function of the word's phonological shape. They may therefore be helpful in answering the question of what in a word makes it more or less difficult to articulate, – a topic that was brought up as early as 1966 in this journal (Shankweiler & Harris, 1966). Empirical sources suggest that, for instance, word length, position effects (high vulnerability of word onsets), the presence of consonant clusters, or the properties of the involved speech sounds may play a role, but there is still a controversy about the relative contributions of, e.g., phonemic length and syllable complexity (e.g., Nickels & Howard, 2004; Romani & Galluzzi, 2005). Since these factors are intricately interwoven in the sound patterns of words, we here seek for an integrative and comprehensive approach to resolve this issue by developing a computational model which spans the hierarchy of the involved units, from articulatory gestures to the rhythmical structure of spoken words.

1.1. *Linear accounts of phonetic planning*

Traditional linguistic and psycholinguistic views of how articulation is organized are based on linear models of the linguistic structures that provide the input to a conceived motor planning system. Usually these structures are viewed as sequences of phoneme-sized or syllable-sized units (for an overview see Shattuck-Hufnagel, 2014). According to the classical generativist view, for instance, phonetic representations of utterances are conceived of as “sequences of discrete segments” (Chomsky & Halle, 1968, p. 5), each of which is considered to contain the prescriptive motor information necessary for the control of the articulator movements generating the respective speech sound. This view has long been influential in cognitive modular models of word form production, such as the standard model described in Ellis and Young (2013, chap. 5). Likewise, interactive activation accounts of word production, such as the one presented by Dell, Schwartz, Martin, Saffran, and Gagnon (1997), also represent the output of phonological encoding processes for words as linear sequences of phonemes which are fed into a not further specified articulator component. Typically in these models, the processing network activates, at its bottom-end, a linear arrangement of nodes termed *phonemes*, with a downward-arrow representing the motor implementation of the activated phoneme string, as in Fig. 1A (cf. Rapp & Goldrick, 2006, for an overview). Even though the generation of segment strings may incorporate suprasegmental (e.g., syllable structural) information, as in Dell et al. (1997), there is no explicit account of how this information is conveyed to and implemented in the motor system. Implicit to these theories is the view that the articulatory planning system deals with linear arrays of phoneme targets, arranged like beads on a string, with some interpolation mechanism serving to generate a continuous articulatory trajectory from a series of discrete phonemic sampling points. From this perspective, prosodic planning is considered as a strictly separate process which imputes rhythm and melody onto the phoneme strings only after their being generated (cf. Keating, 2006).

The Levelt et al. model of speech production (Levelt et al., 1999) takes a different – yet still strictly linear – stance, postulating linearity at a higher organizational level, i.e., between phonetic syllables (Fig. 1B). In this approach, phonological words retrieved from the lexicon are syllabified in a first encoding step, whereupon holistic phonetic plans (which are not labeled for their prosodic roles) are accessed sequentially for each phonological syllable from a repository of syllable-sized speech motor plans. Here again, phonetic plans are arranged as linearly ordered strings, and the prosodic modulation of these strings is construed as a separate process by which the articulatory plans for the different syllables are adjusted to their prosodic contexts (Keating & Shattuck-Hufnagel, 2002).

Still another approach was proposed recently by Hickok (2014), who focused on differential roles of phonemes and syllables in speech production. Hickok's account provides for a hierarchical arrangement of segments and syllables, but it still neglects how these units are embedded in the rhythm and prosody of spoken language (cf. Ziegler, 2014, for a discussion).

1.2. *Nonlinear accounts of phonetic planning*

This dualist view of segmental versus prosodic aspects of speech motor planning and articulation is challenged by observations according to which articulations at the segmental level strongly interact with prosody. As an example, the “strength” of segmental articulation is known to vary systematically with the position of a segment within a hierarchy of prosodic domains (Fougeron & Keating, 1997). In articulations of the consonant [n], for instance, the tongue makes increasingly stronger contact with the roof of the mouth when it occurs initially in a syllable, a word, or a phrase, respectively, as compared to positions within or at the right edge of these prosodic units. Or, as another example, the laryngeal abduction movements (“glottal aperture” movements) through which consonants receive a voiceless quality (e.g., /p, t, k, f, s/ etc.) may vary as a function of prosodic strength (for a