Journal of Phonetics 64 (2017) 108–126

Contents lists available at ScienceDirect

Journal of Phonetics

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

Special Issue: Mechanisms of regulation in speech, eds. Mücke, Hermes & Cho

Speech rate effects in Russian onset clusters are modulated by frequency, but not auditory cue robustness

Marianne Pouplier [a,*], Stefania Marin [a], Philip Hoole [a], Alexei Kochetov [b]

[a] Institute of Phonetics and Speech Processing, Schellingstr. 3, LMU, 80799 Munich, Germany
[b] Department of Linguistics, University of Toronto, 100 St. George Street, Toronto, Ontario M5S 3G3, Canada

ARTICLE INFO

Article history:
Received 20 April 2016
Received in revised form 17 January 2017
Accepted 19 January 2017
Available online 22 February 2017

Keywords:
Russian onset clusters
Speech rate
Auditory cue robustness
Phonotactic frequency
Typological preference
Coarticulation
Durational control

ABSTRACT

This study presents data on the durational and timing characteristics of Russian onset clusters and their change as a function of speaking rate. The focus is on Russian due to it being known to have relatively less consonant overlap compared to languages like English and due to its unusual range of consonant clusters. Using articulography, we investigate whether these characteristics have implications for the flexibility of clusters under speech rate changes. In particular we ask whether a cluster's signal modulation profile, taken as an index of auditory recoverability, predicts the degree to which the overlap pattern of a cluster changes with rate. Previous research suggests that stop + stop clusters may be less susceptible to rate change than other, auditorily more robust clusters within the same language. Moreover, even though frequency and phonotactic preference are usually closely aligned, Russian also allows us to probe frequency effects on cluster timing, since for our data these factors are dissociated to a certain degree. Results show that both duration and relative timing of the consonants in a cluster are subject to change. Speech rate effects do not scale uniformly throughout the cluster but are carried predominantly by the constriction formation duration of C2. Clusters show a decreasing rate effect from high to low frequency clusters. Grouping clusters according to their assumed perceptual robustness does not lead to a clear result. We discuss these findings in the context of models of durational control of speech production.

1. Introduction

The goal of the current paper is to shed light on how timing and overlap in consonant clusters are affected by speech rate. More specifically, we ask whether either frequency or auditory cue robustness may interact with the relative flexibility clusters exhibit under rate changes in conjunction with language-specific phonotactic properties. To that effect, we present articulatory data on the durational characteristics and coarticulatory patterns of Russian onset clusters in two speaking rate conditions. The study addresses three overall research questions: For one, we investigate the extent to which onset clusters in Russian, a language with relatively little consonantal overlap, adapt to speech rate variation. This stands in the context of our increasing knowledge of systematic cross-linguistic differences in coarticulation and consonant-consonant timing. While languages like English, German, or French feature a relatively high degree of overlap between sequences of consonants, other languages like Russian have been reported to time their consonants comparatively far apart, which may in certain circumstances give rise to a transitional schwa between consonants (Bombien & Hoole, 2013; Davidson & Roon, 2008; Zsiga, 2000, 2003; relatedly Öhman, 1966). As explained in more detail below, there is reason to expect that clusters in a language with a low overlap pattern such as Russian may exhibit resistance to rate changes and this is one general hypothesis we put to test.

Secondly, we pursue the possibility that a cluster's susceptibility to rate changes is related to its cue robustness. In particular, we follow an idea originally proposed by Wright (1996) that clusters which are problematic in terms of auditory recoverability fail to increase overlap with rate. Russian offers an ideal opportunity to probe this possibility further since from a typological viewpoint, Russian onset cluster phonotactics are unusual. Russian not only features the cross-linguistically most common sonority-rising clusters like /bl-/, but also sonority pla-

* Corresponding author.
E-mail addresses: pouplier@phonetik.uni-muenchen.de (M. Pouplier), marin@phonetik.uni-muenchen.de (S. Marin), hoole@phonetik.uni-muenchen.de (P. Hoole), al.kochetov@utoronto.ca (A. Kochetov).

http://dx.doi.org/10.1016/j.wocn.2017.01.006
0095-4470/

teau (/tk-/) and even sonority falling sonorant-obstruent clusters like /lb-/. Henke, Kaisse, and Wright (2012) have argued that the traditional concept of sonority, designed to capture cross-linguistic segment-sequencing preferences (e.g., Clements, 1990), is epiphenomenal to auditory cue robustness. The current work empirically tests their prediction that clusters with low cue robustness should resist speech rate changes. Lastly, we ask whether differential cluster sensitivity to speech rate variation may be related to cluster frequency. While Russian does not permit us to completely de-correlate frequency and cue-robustness, the grouping of the clusters we have recorded is not isomorphic for these two factors and we can thus begin to assess their relative contribution to articulatory dynamics.

The following sections will consider in turn the state of the art for how speech rate changes are manifest in articulation in conjunction with known cross-linguistic coarticulatory differences, proposals on how rate may interact with auditory cue robustness, and finally current knowledge on the effects of phonotactic frequency.

1.1. Speech rate changes and the language-specific nature of coarticulation

A series of papers has investigated speech rate with the particular goal of understanding the fundamentals of durational control in speech production and aspects of movement optimization (e.g., Adams, Weismer, & Kent, 1993; Nelson, Perkell, & Westbury, 1984; Ostry & Munhall, 1985; Shaiman, Saltzman, & Tuller, 1995; Smith, Goffman, Zelaznik, Ying, & McGillem, 1995; Tasko & Westbury, 2004). All of these studies have revealed complex trading relations between velocity and amplitude, either of which may or may not vary under rate scaling (see Berry (2011) for an overview). The literature on speech rate effects has mainly been concerned with consonant–vowel transitions and rate effects on individual constrictions, in particular vowels (e.g., Lindblom, 1963). Speech rate scaling is not uniform throughout the CVC syllable in that the proportionally greatest durational compression occurs in the nucleus, and the least in the onset consonant (Gay, 1978, 1981). Adams et al. (1993) found that even the movement phases within a segment differ in how consistently rate changes occur: release gestures exhibited more uniform rate scaling compared to constriction formation gestures of the same segment. They attributed this to the CV transition being perceptually more important; therefore in their view control strategies may seek to minimize variation for perceptually critical movement phases (“islands of reliability”, Adams et al. (1993, p. 48)). Yet they acknowledge that stress may have been a confounding factor in their results. Another known effect of faster speech rate is increased coarticulation in terms of earlier anticipatory movement onsets between successive segments (Agwuele, Sussman, & Lindblom, 2008; Gay, 1978, 1981). Relatedly, inter-articulator phasing has been shown to change with rate (Nittrouer, 1991; Nittrouer, Munhall, Kelso, & Tuller, 1988).

Generally, there has been very little work on how speaking rate affects the timing of consonant sequences, which is the main focus of our present work. In an EPG study, Byrd and Tan (1996) found for American English cross-word consonant clusters that speaking rate changes have a linear effect on the duration of each member of the C1#C2 sequence. In contrast to this, rate conditioned changes to overlap were not consistently present in the data which lead the authors to suggest that listener-oriented factors may act as a constraint on overlap. The interaction of listener-oriented factors with rate variation was targeted specifically in a study on clusters in Tsou (Wright, 1996). Based on acoustic measures Wright observed a greater degree of rate-conditioned shortening for C1 compared to C2, but due to his experimental design (acoustic recordings of words in isolation) the duration measures could not be applied to stop-initial clusters. Hence his study could not distinguish whether this result was due a manner effect (C1 were non-stops and C2 was always a stop in his stimuli for this part of his analyses) or a position effect or an interaction between the two. For stop + stop clusters, C1 retained a release burst at all speaking rates with little variation between rates in the inter-burst-interval (temporal interval between the two stop bursts). At the same time, Wright could trace some acoustic influence of C2 on the release burst spectrum of C1, providing evidence for articulatory overlap between the consonants, despite a consistently released C1 even at the fast rate. Wright concluded that consonant cluster overlap generally is constrained by perceptual factors, since in his data only clusters with particularly vulnerable auditory cues (stop + stop sequences) barred an increase in overlap in fast speech.

Expanding on Wright (1996), Henke et al. (2012) proposed that gestural timing is a key factor in understanding how languages may achieve some level of cue robustness for clusters with perceptually sub-optimal signal modulation profile (discussed in detail in the next section). By extension, these clusters are predicted to be rather immune to any coarticulation-scaling factors such as speech rate. Their proposal thus concerns cluster-specific variation to overlap patterns within a given language: the existence of typologically exceptional (sub-optimal cue) clusters within a language is seen as conditional on these clusters showing a different coarticulation pattern. There are several proposals that go further in suggesting that languages featuring unusually complex consonant phonotactics are generally characterized by certain coarticulatory patterns. That coarticulation is language specific has been known at least since Öhman’s influential (1966) paper in which he observed a relative weakness of V-to-V coarticulation in Russian compared to Swedish and English. More recent work has accumulated more and more evidence for the language-specific nature of both V-to-V and C-C coarticulation, which we here take to be tantamount to temporal overlap (Beddor, Harnsberger, & Lindemann, 2002; Bombien & Hoole, 2013; Chitoran, Goldstein, & Byrd, 2002; Flemming, 2011; Kochetov, Pouplier, & Son, 2007; Ma, Perrier, & Dang, 2015; Manuel, 1990; Marin & Pouplier, 2014; Mok, 2010; Pouplier & Beňuš, 2011; Zsiga, 2000, 2003). In addition, there have been several proposals that there is a causal relationship between certain phonological phenomena and the coarticulation patterns within a language. For instance, Ma and colleagues (2015) recently attributed language specific effects in vowel-to-vowel coarticulation in Mandarin and French to the role of the syllable as a prosodic domain. In contrast to French, V-to-V coarticulation is effectively blocked in Mandarin, consis-