Journal of Phonetics 64 (2017) 1–7

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

Mechanisms of regulation in speech: Linguistic structure and physical control system

Doris Mücke[a,*], Anne Hermes[a], Taehong Cho[b]

[a] IfL - Phonetik, University of Cologne, Herbert-Lewin-Str. 6, 50931 Cologne, Germany
[b] Hanyang Phonetics and Psycholinguistics Lab, Department of English Language and Literature, Hanyang University, Seoul 04763, South Korea

ARTICLE INFO

Article history:
Received 20 January 2017
Received in revised form 16 May 2017
Accepted 22 May 2017
Available online 5 July 2017

Keywords:
Dynamic systems
Variability
Mechanisms of speech regulation
Normal and adverse conditions

ABSTRACT

Speech variation is a naturally-induced phenomenon in human speech communication which can be attributed to the inevitably multifaceted nature of interactions between various higher-order linguistic and lower-order physiological factors. Speech is dynamic, and it is assumed that there are regulation mechanisms behind these complex interactions of structural, contextual and phonetic cues leading to an overwhelming variety of gradient phenomena in the speakers' linguistic behaviour. Recent years have increasingly witnessed the extensive development of dynamical theories which attempt to capture mechanisms of regulation that underlie speech production and perception in a unified way. In this introductory paper, we touch on some basic theoretical groundings of speech dynamics, and discuss the significance of the contributions made by each paper of the special issue under the rubric of mechanisms of regulation in speech. The special issue is interdisciplinary in nature, bringing together papers from different perspectives, ranging from tutorial and critical review papers on dynamic systems to original research papers on the regulation of speech in both normal and adverse (atypical) conditions. These selected papers, taken together, make considerable advancements in illuminating how variation in production and perception can be seen as a window to linguistic structure within and across languages.

1. Mechanisms of regulation in speech

One of the goals of linguistic phonetics is the understanding of underlying principles and mechanisms that regulate variation in speech production. A complex interplay between linguistic structure and the physical system leads to a huge amount of naturally-induced variability. Speakers generate an overwhelming variety of gradient phenomena in their linguistic behaviour. A bundle of factors playing a role in the natural process of human communication trigger and constrain variation in speech, most of them reaching deeply into human physiology, cognition and grammar. There are regulation mechanisms behind these complex interactions of structural, contextual and phonetic cues mediating between naturally-induced variability (e.g. due to prosodic marking) and the need for gestural coherence in order to preserve the phonological form of a given pattern. This is illustrated in Fig. 1. On the one hand, there is a need for prosodic marking (e.g. boundary marking) in the articulatory substance. The higher the prosodic domain, the stronger the spatial and temporal modifications on domain-initial consonants. Fougeron and Keating (1997) have shown in an EPG study of French that the linguopalatal contact for the alveolar nasal /n/ increases at strong boundaries and decreases at weak boundaries. On the other hand, there is a need for gestural coherence in order to preserve phonological form. Segments differ in terms of their degree of coarticulatory resistance (Farnetani & Recasens, 2010). Alveolar fricatives such as /s/, for example require a very precise predorsal activation. They show fewer degrees of articulatory and acoustic freedom (Bombien, Mooshammer, Hoole, & Kühnert, 2010:390) and therefore they are more resistant to prosodic changes than segments such as /n/, which have a lower degree of coarticulatory resistance. In a similar vein, Cho (2004) demonstrated that articulatory gestures for vowels in VCV context in English resisted coarticulation both at prosodic junctures and under prominence (e.g. receiving pitch accent), showing an interplay between naturally-induced variation and gestural coherence being modulated by prosodic structure.

* Corresponding author.
E-mail address: doris.muecke@uni-koeln.de (D. Mücke).

http://dx.doi.org/10.1016/j.wocn.2017.05.005
0095-4470/

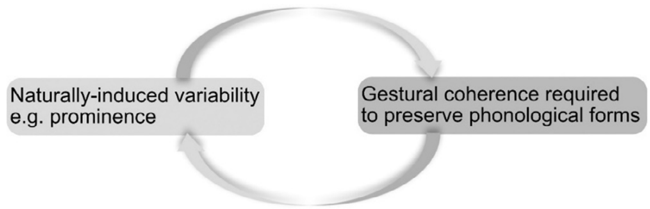

Fig. 1. Regulation mechanisms in speech.

Understanding how speech variation is regulated, however, appears to be never complete and in turn entails new questions. In his paper "*The devil is in the detail*", Nolan (1999:1) also notes that "the more detailed our knowledge of the properties of speech becomes, the more difficult it is to sustain our simplifying assumptions, our models which help us comprehend our universe".

The last few decades, however, have witnessed development of theories of speech dynamics which aim at illuminating mechanisms of regulation in speech in a unified way. In dynamic approaches, phonetic detail is seen as a direct window to linguistic structure rather than as the 'curse' of modern measuring techniques. Consequently, variability in the physical world is seen as a direct outcome of the system's behaviour, a challenge between naturally-induced pattern variability in human communication and the need for pattern stability to preserve phonological forms, mostly language-driven.

In this thematic special issue, we attempt to take a step forward toward disentangling the complexity in the mechanisms of speech regulation by bringing together different studies at different levels of description under the rubric of speech dynamics, exploring fundamentals of regulation in terms of dynamical systems, regulation of articulatory gestures and regulation of speech in adverse conditions. We assume that complexity in mechanisms of speech regulation can be best understood by allowing for a number of different perspectives (Wagner et al., 2015). In addition to a tutorial on the basics of dynamical systems, a review of dynamics in perception and production is presented. The special issue presents studies that show the need for speech systems that are flexible and adaptive to changes over time and that can respond to the complex interplay of grammatical, prosodic and physiological demands. Intonation as well as supralaryngeal articulation are presented in terms of continuous parameters that should be – in the future – integrated in a unified system. A combined focus on clinical and non-clinical speech in terms of the interaction between dynamical diseases and the requirements of a phonological system in a given language displays the complexity of speech regulation and deals with the breakdown of phonological patterns as a gradient phenomenon.

The papers in this special issue are organized into three sections: Section 1.1, Dynamic systems; Section 1.2, Regulation of articulatory gestures; Section 1.3, Regulation in adverse conditions. Each section starts with a brief overview of the research field and then deals with the individual contributions to the special issue.

1.1. Regulation in terms of dynamic systems

Dynamic systems describe the evolution of the complex behaviour of a system—that is, the language in this case. In a linguistic system, phonological information (low-dimensional) can be mapped directly onto continuous phonetic cues (high-dimensional) without the need for an interface between phonological form and phonetic substance (Browman & Goldstein, 1992; Gafos & Beňuš, 2006; Goldstein, Byrd, & Saltzman, 2006; Mücke, Grice, & Cho, 2014; Saltzman & Munhall, 1989). A dynamic system changes its behaviour in a lawful manner such that *rules of change* in terms of mathematical laws can be captured by the use of differential equations. Those dynamic approaches bridge the gap between discrete phonological description and continuous phonetic representations by modelling them in a unified system.

The differential equation of a dynamic system specifies the continuous behaviour of the system over time. While the equation is invariant, the physical output is not (Browman & Goldstein, 1989, 1992; Gafos, 2006; Gafos, Charlow, Shaw, & Hoole, 2014; Kelso, 1995; Spivey, 2007). Once a dynamic system is set into motion, it evolves towards a specific (linguistic) target (i.e. equilibrium position). This target is defined by an attractor, which defines values or regions of values in the possible phase space of the system. Attractors are often compared with a marble rolling to the bottom center of a bowl (Haken, Kelso, & Bunz, 1985; Nam, Goldstein, & Saltzman, 2010). In the rolling marble metaphor, a marble rolls into a bowl. The bowl defines all possible values of the phase space, and the attractor would be the bottom center of the bowl. Those dynamic systems always encode context-dependent variability. If the marble starts to roll next to the bottom center of the bowl, the path to the center is short. If it starts to roll from the bowl's margin, the path to the center is long. However, in both cases the marble is likely to roll towards the bottom center of the bowl, where the system eventually stabilizes and the marble comes to rest. If there is competition between multiple target attractors, the system evolves to one of the attractors as a function of the different starting conditions and attractor strengths (Tuller & Lancia, 2017).

Many skilled movements by humans have been characterized as being controlled by such a dynamical (point-attractor) system (see Goldstein et al., 2006 and the references therein). Speech production and perception can also be understood as dynamic systems using attractors that reflect linguistic structure. Let's assume that the attractor is a linguistic goal such as the lip closure during the production of an intervocalic consonant in a sequence such as /ibi/ and /aba/. The goal for the lips in /b/ is invariant (full closure of the lips), but the way the lips travel differs in the two conditions. The way is shorter in /ibi/ than in /aba/, because due to the different starting conditions the jaw is already higher in the high vowel /i/ than in the low vowel /a/. A dynamic system in speech needs to be flexible, and redundancy plays an important role (Browman & Goldstein, 1992; Fowler, Rubin, Remez, & Turvey, 1980; Goldstein & Pouplier, 2014; Hawkins, 1992; Saltzman & Kelso, 1987; Saltzman & Munhall, 1989). A great amount of context-dependent variability is generated in such a system, reflecting functional synergies of the articulators moving towards different competing attractors. Thus, phenomena like partial assimilation in English /n#g/ sequences (Barry, 1991; Ellis & Hardcastle, 2002) or incomplete neutralisation in German voiceless and "devoiced" stops (Roettger, Winter, Grawunder, Kirby, & Grice, 2014) tell us more about the system's behaviour rather than being subphonemic noise.

All four papers in this section on "*Regulation in terms of dynamic systems*" deal with basic concepts of dynamic systems and their application in theories of speech production and perception. The paper by *Khalil Iskarous (2017)* is a tutorial that offers the reader basics on differential equations applied to dynamic systems. These equations are important for understanding how to bridge the gap between discrete phonological and continuous phonetic descriptions. They are exemplified in teems of the motions of spring-mass systems. The equations are invariant and predict the phonetic outcome in terms of a movement trajectory of the dynamic system. Furthermore, Jordan's theory of serial order is introduced (Jordan, 1986) to treat higher level problems such as word and syllable generation on the basis of a network of oscillators.

Betty Tuller and *Leonardo Lancia* (2017) present a critical review on dynamic models in speech production and perception. They argue for an integration of perception and production, discussing empirical work on dynamical models addressing syllabification and categorization over the last 25 years, including nonlinear dynamic models, Bayesian approaches and hybrid approaches. There is a competition between stability and variability also in the production-perception link. In the communication process, speakers and listeners rely on repeatable articulatory patterns. However, the system's behaviour needs to be flexible enough to deal with changing environments. There are transitions between target attractors. Those transitions can be understood as instabilities, reflecting the regulation mechanisms of a complex adaptive system. In these approaches, cognitive factors such as attention, expectation, and memory need to be incorporated into a system as they systematically influence the stability and variability of the system's output.

When capturing coordinative movements over time of the so-called independent systems of oral and glottal control and head movements, those fluctuations need to be addressed.

In a production study using Electromagnetic Articulography, *Sam Tilsen (2017)* questions the influence of exertive modulations such as attention, effort, motivation and arousal on the behaviour of the speech system when speakers are recorded over longer experimental sessions. He investigated fluctuations in movement patterns of oral and glottal control as well as head movements during the speech process. As an experimental paradigm, he used syllable repetition tasks to focus on regulation mechanisms in coordinated speech movements over large-scale investigations of speakers' behaviour. He found that exertive mechanisms such as attention, effort or motivation show global effects on speech control involving coordinated movements of the glottal and oral control systems as well as the head. Therefore, he extends the dynamic account of speech articulation and planning oscillators by proposing the inclusion of a phase-domain analysis of timing instead of a linear time-domain analysis.

Noah Nelson and *Andrew Wedel* (2017) investigate the influence of lexical neighbourhood relationships on the dynamics of phonetic behaviour in speech production. They show that hyperarticulation can be found to distinguish between the target word and its competitor. More specifically they examine lexical competition and phonetic specificity in terms of voice onset time predictions in word-initial stops in conversational speech in French. They use metrics for lexical competition when comparing the phonetic specificity of target word and competitor (neighborhood density, onset competition, cue-specific minimal-pairs), and demonstrate that contrastive hyperarticulation predicts opposite effects for voiced versus voiceless stops under the influence of lexical competition. Speakers decrease VOT in voiced stops and increase VOT in voiceless stops in order to expand the perceptual distance between the target word and the competitor.

1.2. Regulation of articulatory gestures

In Articulatory Phonology based on Task Dynamics, speech can be decomposed into a set of potentially overlapping units: articulatory gestures (Browman & Goldstein, 1989; Browman & Goldstein, 1992; Saltzman, 1986; Saltzman & Kelso, 1987; Saltzman & Munhall, 1989). Articulatory gestures integrate low-dimensional descriptions (the gesture as a discrete phonological unit) and high-dimensional descriptions (the gesture as a continuous physical action) in a unified system. They define articulatory gestures such as the full closure of the tongue tip for /t/ in terms of vocal tract constrictions at the alveolar ridge. Changing the value of a gesture's parameter set changes the temporal and/or spatial properties of the physical articulatory action and therefore the acoustic outcome. For example, changing the degree of constriction from "closure alveolar" into "critical alveolar" leads to the spirantization of /t/.

Parameter modifications of vocal tract actions are related not only to consonant and vowel production in the textual string, but also to prosody and intonation. Therefore, supralaryngeal and laryngeal gestures should not be modelled in isolation but rather as part of the same regulation mechanisms as coordinative structures. This is a task for future research. In modelling intonation, tonal structures should also be seen from a dynamic point of view. In intonation, for example, not only the paradigmatic choice of a discrete pitch accent type but also the continuous parameters of its realisation, such as peak height and alignment with the segmental material, play a role in marking prominence (Ladd, 2008). It has been described for rising pitch accents that both later and higher peaks mark contrastivity (Gussenhoven, 2004), pointing to the fact that intonational categories also consist of a bundle of multiple parameters modifying their phonetic realisation. Indeed, there is no fundamental split between what is referred to as text and tune, and the two systems are intricately coordinated with each other. Thus, changes in phonetic parameters cannot be modelled in isolation (Shaw, Gafos, Hoole, & Zeroual, 2011) – instead, they are attributed to relations and interactions between different phonetic parameters changing over time in a lawful way. We assume that there is a regulation mechanism behind the complex interactions of structural, contextual and phonetic cues, which is important to capture when speakers, for example, "turn up the volume" to mark prominent elements in a speech continuum via intonational and articulatory cues. This means, intonation and articulation should not be investigated in isolation but viewed as a single unified system.

But what does it mean to “turn up the volume” in terms of consonant–vowel production? Signatures of prominence can be found in the supralaryngeal articulation of different linguistic constituents ranging from the syllable to the intonational phrase. Previous research has revealed a more distinct articulation of prosodic units such as syllables in prominent positions (e.g. under contrastive focus), involving larger, longer and faster movements of the vocal tract. Fig. 2 shows lip trajectories (inter-lip distance) from one speaker during the production of the target word <Bahber> in the German utterance <Melanie will Dr. Bahber treffen> (lit.: Melanie wants Dr. Bahber to-meet) in maximally diverging focus structures (black: background condition, dashed: contrastive focus). Low values indicate that the lips are closed during the production of the consonant, while high values indicate that the lips are open during the vowel. In contrastive focus condition, lip movements are considerably larger, longer and faster compared to background condition. These temporal and spatial expansions of articulatory movements are collectively referred to as *prosodic strengthening* (Beckman, Edwards, & Fletcher, 1992; Cho, 2006; Cho, 2016; Cho & Keating, 2009; de Jong, 2004; Harrington, Fletcher, & Beckman, 2000; Mücke & Grice, 2014). Under accent, for example, speakers produce louder and longer vowels by opening the mouth wider over a longer time (see Fig. 2). A more open oral cavity allows for greater radiation from the lips, leading to an increase in the overall acoustic energy.

However, prosodic modifications of vocal tract actions are highly complex, affecting multiple cues to prosodic prominence (Cho, 2006). In addition, effects of speaker-specific behaviour and segmental context also play an important role. Speakers use multiple cues in different combinations to express the same degree of prominence (e.g. contrastive focus), but the use of these cues is also dependent on the coarticulatory sensitivity of the segmental material (i.e. certain sounds are more resistant to coarticulation than others). Therefore, highlighting mechanisms are rather complex and require parameter adjustments in multiple dimensions. This kind of prosodically-conditioned complexity and modulation of speech may be captured well within the framework of dynamic systems evolving over time.

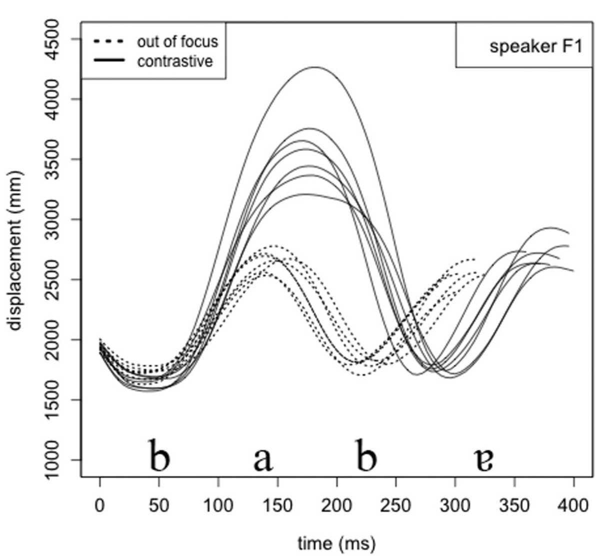

Fig. 2. Lip trajectories for the German target word ‘Bahber’ /babɐ/ for one speaker (background = black lines; contrastive focus = dashed lines); adapted from Mücke and Grice (2014).

This section on “*Regulation of articulatory gestures”* consists of four papers which address issues related to coordinative structures of laryngeal and supralaryngeal gestures, respectively. They focus on the interplay of physiology, prosody and grammar. All of them use acoustic data, with three of them in combination with EMA data.

Taehong Cho, Daejin Kim and Sahyang Kim (2017) investigate the effect of prosodic strengthening and linguistic contrast on the temporal realization of nasals and adjacent vowels in English. More specifically, they shed light on the variation in nasal coarticulation – a longstanding problem – as a function of focus and boundary strength in English. The effect of focus reveals symmetric coarticulatory resistance effects at both edges of prosodic domain in that it decreases vowel nasalization in CVN# and #NVC sequences despite the fact that the nasal consonant itself is temporally expanded in both conditions. Thus, this type of prominence entails an enhancement of the nasality of the consonant and at the same time also of the orality of the vowel (through coarticulatory resistance to nasalization). Boundary strength, on the other hand, *increases* vowel nasalization showing coarticulatory vulnerability but only in the domain-final position for CVN# sequences. The results reveal different types of enhancement patterns depending on prosodic position. In initial position (#NVC), the CV contrast is enhanced by reducing the nasality of the consonant driven by a boundary-induced syntagmatic enhancement, whereas in final position (CVN#), the increased vowel nasalization is arguably attributable to a phrase-final weakening of the articulatory linkage of the oral constriction and the velum lowering gesture. This study supports the view that speakers do indeed control for phonetic granularity in their expression of prosodic prominence, leading to the assumption that this behaviour is part of the phonetic grammar, reflecting dynamics of speech on different levels of the prosodic hierarchy.

This view of regulation mechanism can also be adopted for the paper by *Martine Grice, Simon Ritter, Henrik Niemann and Timo Roettger* (2017). This paper supports the claim for an integration of discrete phonological pitch accents and the continuous phonetic parameters. The authors provide an analysis of the variation in the production of pitch accents under three different focus conditions (broad, narrow and contrastive focus) in German. In their paper they analyse several phonetic parameters, such as F0 peak alignment, tonal onglide and target height. The analysis reveals differences as well as similarities across speakers: There are clear speaker-specific differences, but the relative patterns were similar, i.e. parameters were modulated in the same direction in that e.g. F0 peaks were later aligned in contrastive focus than in narrow focus. The intonational analyses are linked to results of supralaryngeal dynamics (Electromagnetic Articulography) and to a perception task. The study supports the claim that there is no one-to-one mapping between intonational categories and the pragmatic function. This study argues for an integrated analysis of discrete phonological categories and the modulation of continuous phonetic parameters and emphasizes the need for a dynamical approach linking quantitative and qualitative effects.

Mariann Pouplier, Stefania Marin, Philip Hoole and Alexei Kochetov (2017) capture the influence of speaking rate on the durational and temporal properties of Russian word-initial clusters. The authors ask how speech rate interacts with auditory cue robustness and lexical frequency. Therefore, they analyse intra-gestural parameters, such as constriction formation duration of each consonant, as well as inter-gestural parameters, such as the plateau lag between consonants in order to shed light on the flexibility of these clusters under speech rate changes. The effect of speech rate on the the consonants is not consistent, the second consonant is affected more. Furthermore, the data provide evidence that from high to low frequency clusters, rate effects decrease. However, the claim that certain clusters are not affected by speech rate changes due to their auditory robustness could not be supported. The authors describe these results for Russian word initial clusters as involving "conditional flexibility of relative timing patterns". This flexibility also fits the view of claiming that we are dealing with regulation mechanisms mediating between the variability, in this case due to frequency and speech rate, and the need to preserve the phonological form, i.e. the syllable. This means that the syllable coordination pattern behaves flexibly when mediating between prosodically-induced changes such as speaking style and pattern coherence in terms of the auditory recoverability of different cluster types.

Anne Hermes, Doris Mücke and Bastian (2017) simulate the degree of prosodic variability in the phonetic outcome of phonological syllable parses in Polish and Tashlhiyt Berber by pushing the system to its limits. The critical point in time when a pattern switches from complex to simple syllable organisation or vice versa differs in the languages under investigation and is likely driven by linguistic constraints. More specifically, the authors explore how temporal realization of continuous phonetic parameters may be regulated by a given linguistic system by examining the relationship between intergestural timing of word-initial consonant clusters and syllable structure in Tashlhiyt Berber vs. Polish. Results of their articulatory study are consistent with the simple onset coordination hypothesis for Tashlhiyt Berber (showing a 'tighter' gestural overlap and anchoring stability) but with the complex onset coordination hypothesis for Polish (showing a temporal accommodation of the initial consonant in the cluster). Moreover, the data in Tashlhiyt Berber fit well with the idealized stimulation pattern for simple onset parse across different consonant types even with an increase in anchor variability, although the simulation pattern for Polish is less clear presumably due to variability caused by segmental make-up. Interestingly, however, the observed temporal stability in Tashlhiyt Berber runs counter to the temporal variability previously observed for consonant clusters in Moroccan Arabic which is also characterized by simple onset coordination. The authors propose that such a language-specific degree of temporal variability may not be attributable to phonological syllable parse itself (simple vs. complex), but rather as a result of different regulation mechanisms between prosodic marking and the need to preserve phonological form, which should be specified in a dynamical system in a language-specific way.

1.3. Regulations in adverse conditions

Natural variability in human communication is determined by a number of different factors, such as segmental context or pragmatic functions, among others. However, as we have discussed so far, languages differ in terms of the degree of variability tolerated by the phonological system. In atypical speech, the term *dynamical disease* has been established. Dynamical disease refers to a qualitative change of the system's dynamics (*Mackey & Milton, 1987*), or in other words, a striking change in bodily rhythms (Glass, 2015):

"A dynamical disease is defined as a disease that occurs in an intact physiological control system operating in a range of control parameters that leads to abnormal dynamics. The signature of a dynamical disease is a change in the qualitative

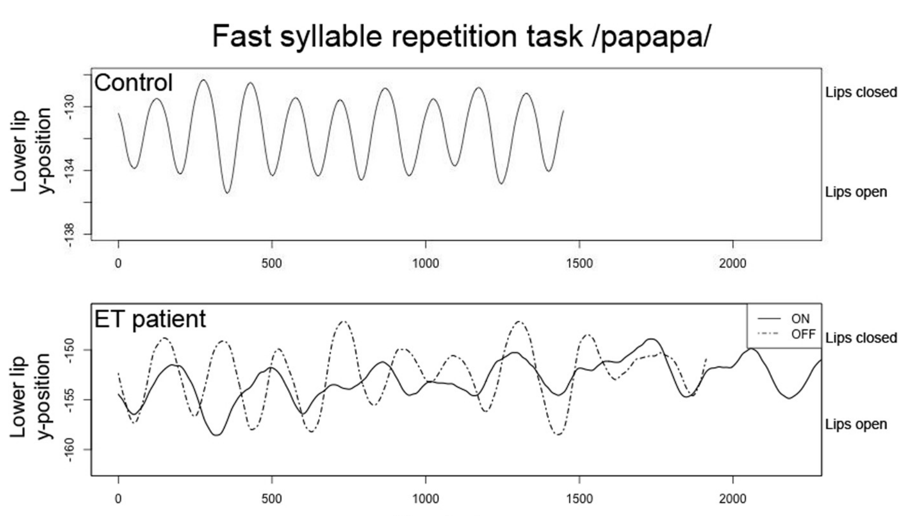

Fig. 3. Lower lip trajectory for fast syllable repetition task /papapa/ for one control speaker (top) and one ET patient (below; deep brain stimulation: dashed = stimulation-OFF; black = stimulation-ON), from Mücke et al. (2016).

dynamics of some observable nature as one or more parameters change" (Mackey & Milton, 1987:16)

Fig. 3 shows the production of ten syllable repetition cycles /papapa/ in an oral diadochokinesis task. The trajectories in the lower graph show the lower lip movement of an Essential Tremor patient treated with Deep Brain Stimulation. The patient was recorded in one recording session with stimulation-OFF (dashed trajectories) and a stimulation-ON (black trajectories) at the IfL Phonetics Lab in Cologne (Mücke et al., 2016). The trajectories in the upper graph show the lower lip movements of a healthy control speaker. Both the Essential Tremor patient and the healthy control speaker were asked to produce fast syllable repetitions of /papapa/-sequences on a single breath. High values indicate that the lips are closed during the consonant production /p/, and low values indicate that the lips are open during the vowel production /a/. A visual inspection of Fig. 3 immediately reveals that the syllable cycles for the Essential Tremor patient in the stimulation-OFF condition are highly irregular in comparison to the healthy control speaker. However, under stimulation (black trajectories) this effect is intensified, leading to a poor performance of opening and closing gestures within each syllable cycle. In particular, in the production for the same patient the rather small values for maximum displacement in the stimulation-ON condition compared to stimulation-OFF indicate incomplete closure of the lips during the production of /p/. On the acoustic surface this leads to spirantisation or even to a breakdown of the syllable pattern when the speech motor system is pushed to its limits in stimulation-ON condition. The stimulation induces qualitative changes in dynamics in the fast syllable task indicative of a dynamical disease. However, timing differences induced by changes in dynamics are not always categorical and can also be gradient in nature. A certain degree of variability defined by the phonological system of a given language is tolerated before the pattern breaks down.

In this section on *Regulations in adverse conditions*, two papers focus on the breakdown of phonological patterns and the use of compensatory strategies. Through a simulation task, *Wolfram Ziegler, Ingrid Aichert* and *Anja Staiger* (2017) investigate characteristics of articulatory accuracy in individuals with impairment of speech motor planning, i.e. Apraxia of Speech (AOS). In their clinical model, they capture to what extent an individual's performance of "easy" and "difficult" words (in terms of the motor proficiency requirements) fits into the patterns established for apraxia, raising the need for dynamic accounts and individual case studies in clinical studies. Therefore, the authors provide computational simulations based on speech data from patients with AOS. They aim to test the influence of phonological structure on word production in order to predict the ease of articulation patterns from apraxic speech error patterns. With their latent trait approach, they want to complement the previously suggested non-linear gestural model (NLG, Ziegler and Aichert, 2015) as a clinical model in order to allow for single case testing in the area of speech motor planning. In the present study, they calculated a Rasch model (also referred to as Item-Theory Response model) of speech planning ability based on accuracy data of clinical cases on German words. The probability is modelled as a logistic function and thus, allows one to estimate how well a single case fits with the overall shape of the model. In a following step, the simulation results provide different error mechanisms, indicating that only the NLG simulation fits with the latent trait approach for speech motor planning ability. This model is a promising tool for clinical use to test whether speakers' performance conforms to AOS. With this comprehensive approach that is especially related to adverse conditions, they aim to bridge the gap between neurophonetic evidence and phonological theories. This paper provides insights into regulation mechanisms in speakers who are not able to produce fluent speech, thus extending our knowledge of the interplay of linguistic structure and the physical control system in adverse conditions.

Serge Pinto, Angel Chan, Isabel Guimarães, Rui Rothe-Neves and *Jasmin Sadat* (2017) present a review on dysarthria in Parkinson's disease from a cross-linguistic perspective. They show that in atypical speech, regulations are constrained by the grammar of a given language, a fact that is often overlooked in clinical research of motor speech disorders. The paper covers different speech dimensions related to dysarthric Parkinsons' disease, i.e. phonation, articulation, prosody. Furthermore, the aspect of speech intelligibility is put into focus, since for example modification of certain phonetic parameters, such as breathiness or elisions, could still entail a 'normal' speech intelligibility. They provide three main approaches to assess speech intelligibility: the evaluation of oromotor activity, the description of articulatory deficits and the evaluation of speech recordings by native listeners. They argue for a need to focus on the link between speech disorders and speech intelligibility. In terms of regulation mechanisms in speech, this cross-linguistic approach would provide a new window into dysarthric research in general and more specifically into how linguistic structure affects the physical control system.

1.4. Conclusion

We have shown that the mechanisms involved in the regulation of speech play an important role when investigating speech production and perception. Natural variation can be seen as a window to linguistic structure rather than as noise in the human communication process. Systems in perception and production are not static, i.e. they do not trigger canonical forms from which all types of variation must be seen as variants or noise. Moreover, speech is dynamic and flexible in nature, truly interacting at different levels within the linguistic system. The competition of prosodic marking and the need to preserve phonological form within a given language can be seen as just one of many examples showing mechanisms of regulation between different levels of description. This becomes clear when looking, for example, at where those mechanisms break down in typical and atypical speech. Future research will have to fulfill the task of integrating different linguistic levels and linking different disciplines in order to understand the core of the dynamic process.

Acknowledgements

This work was supported by the National Research Foundation of Korea Grant to T. Cho funded by the Korean Government (Grant No. NRF- 2013S1A2A2035410) and by the German Research Foundation (DFG), Collaborative Research Centre 1252/4 Prominence in Language.

References

Barry, M. C., 1991. Temporal modelling of gestures in articulatory assimilation. *Proc. 12th International congress of phonetic sciences* (pp. 14–17) Aix-en-Provence, France.

Beckman, M. E., Edwards, J., & Fletcher, J. (1992). Prosodic structure and tempo in a sonority model of articulatory dynamics. In G. J. Docherty & D. R. Ladd (Eds.), *Papers in laboratory phonology II: Segment, gesture, prosody* (pp. 68–86). Cambridge: Cambridge University Press.

Bombien, L., Mooshammer, C., Hoole, P., & Kühnert, B. (2010). Prosodic and segmental effects on EPG contact patterns of word-initial German clusters. *Journal of Phonetics, 38*(3), 388–403.

Browman, C. P., & Goldstein, L. (1989). Articulatory gestures as phonological units. *Phonology, 6*(2), 201–251.

Browman, C. P., & Goldstein, L. (1992). Articulatory phonology: An overview. *Phonetica, 49*(3–4), 155–180.

Cho, T. (2004). Prosodically-conditioned strengthening and vowel-to-vowel coarticulation in English. *Journal of Phonetics, 32*(2), 141–176.

Cho, T. (2006). Manifestation of prosodic structure in articulation: Evidence from lip kinematics in English. *Laboratory Phonology, 8*, 519–548.

Cho, T. (2016). Prosodic boundary strengthening in the phonetics-prosody interface. *Language and Linguistics Compass, 10*(3), 120–141.

Cho, T., & Keating, P. (2009). Effects of initial position versus prominence in English. *Journal of Phonetics, 37*(4), 466–485.

Cho, T., Kim, D., & Kim, S. (2017). Prosodically-conditioned fine-tuning of coarticulatory vowel nasalization in English. *Journal of Phonetics, 64*, 71–89.

de Jong, K. (2004). Stress, lexical focus, and segmental focus in English: Patterns of variation in vowel duration. *Journal of Phonetics, 32*, 493–516.

Ellis, L., & Hardcastle, W. J. (2002). Categorical and gradient properties of assimilation in alveolar to velar sequences: Evidence from EPG and EMA data. *Journal of Phonetics, 30*, 373–396.

Farnetani, E., & Recasens, D. (2010). Coarticulation and connected speech processes. In W. J. Hardcastle, J. Laver, & F. E. Gibbon (Eds.), *The handbook of phonetic sciences*. Oxford UK: Blackwell Publishing Ltd.

Fougeron, C., & Keating, P. A. (1997). Articulatory strengthening at edges of prosodic domains. *The Journal of the Acoustical Society of America, 101*, 3728–3740.

Fowler, C. A., Rubin, P. E., Remez, R. E., & Turvey, M. T. (1980). Implications for speech production of a general theory of action. In B. Butterworth (Ed.), *Language production, Vol. I: Speech and talk* (pp. 373–420). New York: Academic Press.

Gafos, A. I., & Beňuš, Š. (2006). Dynamics of phonological cognition. *Cognitive Science [2005 Rumelhart Prize Special Issue Honoring Paul Smolensky: Optimization and Grammar in the Cognitive Science of Language], 30*(5), 905–943.

Gafos, A., Charlow, S., Shaw, J. A., & Hoole, P. (2014). Stochastic time analysis of syllable referential intervals and simplex onsets. *Journal of Phonetics, 44*, 152–166.

Glass, L. (2015). Dynamical disease: Challenges for nonlinear dynamics and medicine. *Chaos: an Interdisciplinary Journal of Nonlinear Science, 25*(9), 097603.

Goldstein, L., Byrd, D., & Saltzman, E. (2006). The role of vocal tract gestural action units in understanding the evolution of phonology. In M. Arbib (Ed.), *Action to language via the mirror neuron system* (pp. 215–249). Cambridge: Cambridge University Press.

Goldstein, L., & Pouplier, M. (2014). The temporal organization of speech. In V. Ferreira, M. Goldrick, & M. Miozzo (Eds.), *The Oxford handbook of language production* (pp. 210–227). Oxford: Oxford University Press.

Grice, M., Ritter, S., Niemann, H., & Roettger, T. B. (2017). Integrating the discreteness and continuity of intonational categories. *Journal of Phonetics, 64*, 90–107.

Gussenhoven, C. (2004). *The phonology of tone and intonation*. Cambridge: Cambridge University Press.

Haken, H., Kelso, J. A. S., & Bunz, H. (1985). A theoretical model of phase transitions in human hand movements. *Biological Cybernetics, 51*, 347–356.

Harrington, J., Fletcher, J., & Beckman, M. E. (2000). Manner and place conflicts in the articulation of accent in Australian English. In M. Broe (Ed.), *Papers in Laboratory Phonology 5* (pp. 40–55). Cambridge: Cambridge University Press.

Hawkins, S. (1992). An introduction to task dynamics. In J. Docherty & D. R. Ladd (Eds.), *Papers in laboratory phonology II: Gesture, segment, prosody* (pp. 9–25). Cambridge: Cambridge University Press.

Hermes, A., Mücke, D., & Auris, B. (2017). The variability of syllable patterns in Tashlhiyt Berber and Polish. *Journal of Phonetics, 64*, 127–144.

Iskarous, K. (2017). The relation between the continuous and the discrete: A note on the first principles of speech dynamics. *Journal of Phonetics, 64*, 8–20.

Jordan, M. I. (1986). *Serial order: A Parallel distributed approach. ICS Report 8604*. San Diego: University of California.

Kelso, J. A. S. (1995). *Dynamic patterns: The self-organization of brain and behavior*. Cambridge, MA: MIT Press.

Ladd, D. R. (2008). *Intonational phonology* (2nd ed.). Cambridge: Cambridge University Press.

Mackey, M. C., & Milton, J. G. (1987). Dynamical diseases. *Annals of the New York Academy of Sciences, 504*, 16–32.

Mücke, D., Grice, M., & Cho, T. (2014). More than a magic moment – Paving the way for dynamics of articulation and prosodic structure. *Journal of Phonetics*, 441–447.

Mücke, D., Hermes, A., Barbe, M. T., Roettger, T. B. Niemann H., Becker, J., et al. (2016). Regulation in speech motor control in ET patients treated with DBS: Kinematics of the lingual and the labial systems. In: *20th International congress of Parkinson's disease and movement disorders*, Berlin, June 19–23, 2016.

Nam, H., Goldstein, L., & Saltzman, E. (2010). Self-organization of syllable structure: A coupled oscillator model. In F. Pellegrino, E. Marisco, & I. Chitoran (Eds.), *Approaches to phonological complexity* (pp. 299–328). Berlin: de Gruyter.

Nelson, N. R., & Wedel, A. (2017). The phonetic specificity of competition: Contrastive hyperarticulation of voice onset time in conversational English. *Journal of Phonetics, 64*, 51–70.

Nolan, F. (1999). The devil is in the detail. In J. J. Ohala, Y. Hasegawa, M. Ohala, D. Granville, & A.C. Bailey (Eds). *Proceedings of the XIV international congress on phonetic sciences*. Berkeley, 1–8.

Pinto, S., Chan, A., Guimarães, I., Rothe-Neves, R., & Sadat, J. (2017). A cross-linguistic perspective to the study of dysarthria in Parkinson's disease. *Journal of Phonetics, 64*, 156–167.

Pouplier, M., Marin, S., Hoole, P., & Kochetov, A. (2017). Speech rate effects in Russian onset clusters are modulated by frequency, but not auditory cue robustness. *Journal of Phonetics, 64*, 108–126.

Roettger, T. B., Winter, B., Grawunder, S., Kirby, J., & Grice, M. (2014). Assessing incomplete neutralization of final devoicing in German. *Journal of Phonetics, 43*, 11–25.

Saltzman, E. (1986). Task dynamic coordination of the speech articulators: A preliminary model. In H. Heuer & C. Fromm (Eds.), *Generation and modulation of action patterns* (pp. 129–144). Berlin: Springer-Verlag.

Saltzman, E., & Munhall, K. (1989). A dynamic approach to gestural patterning in speech production. *Ecological Psychology, 1*, 333–382.

Saltzman, E., & Kelso, J. A. S. (1987). Skilled actions: A task-dynamic approach. *Psychological Review, 94*, 84–106.

Shaw, J. A., Gafos, A., Hoole, P., & Zeroual, C. (2011). Dynamic invariance in the phonetic expression of syllable structure: A case study of Moroccan Arabic consonant clusters. *Phonology, 28*(3), 455–490.

Spivey, M. (2007). *The continuity of mind*. Oxford: Oxford University Press.

Tilsen, S. (2017). Exertive modulation of speech and articulatory phasing. *Journal of Phonetics, 64*, 34–50.

Tuller, B., & Lancia, L. (2017). Speech dynamics: Converging evidence from syllabification and categorization. *Journal of Phonetics, 64*, 21–33.

Wagner, P., Origlia, A., Avesani, C., Christodoulides, G., Cutugno, F., D'Imperio, M., et al. (2015). Different parts of the same elephant: A roadmap to disentangle and connect different perspectives on prosodic prominence. In: *Proceedings of the 18th International Congress of Phonetic Sciences*, Glasgow, Scotland.

Ziegler, W., Aichert, I., & Staiger, A. (2017). When words don't come easily: A latent trait analysis of impaired speech motor planning in patients with apraxia of speech. *Journal of Phonetics, 64*, 145–155.