Jaargang 22, Supplement, juli 2017

Stem-, Spraak- en Taalpathologie

7th International Conference on Speech Motor Control Groningen: Abstracts

rijksuniversiteit groningen / bibliotheek

32.8310/2017/Supplement - i

Preface

This Supplement contains the abstracts of the seventh edition of the International Conference on Speech Motor Control, which is held in Groningen, The Netherlands, July 5 - 8, 2017. With this seventh conference, a well-established Nijmegen (5 editions) - Groningen (6^{th} edition) tradition is continued. This conference, like the ones before, highlights new trends and state-of-the-art approaches in theoretical and applied research in the area of normal and disordered speech motor control. The past decades have yielded a tremendous, multidisciplinary development in this dynamic research field, which is reflected in the Nijmegen - Groningen series of conferences.

In the first edition in 1985, focus was on motor control issues in stuttering. The second conference (1990) highlighted the development of more general motor control models and the inclusion of higher order psychomotor and psycholinguistic functions, broadening the scope to other motor speech disorders than stuttering. At the third conference (1996), more emphasis was put on the emerging field of brain imaging. In addition, development of speech motor control became a prominent topic. At the fourth edition in 2001, we witnessed the introduction of important theoretical neurophysiological and neurobehavioral concepts, and a growing interest in the interface between higher order cognitive/psycholinguistic processes and speech production.

The fifth and the sixth conferences in 2006 and 2011 focused on the development of interdisciplinary collaboration in the field of speech motor research on populations with and without speech disorders. Integration was the key-concept: integration of principles and models of perception-action relations in general and speech as an audio-visual-speech-motor performance in particular; biomechanical, and neurobiological aspects of motor control in general, and the biomechanics and neurological control mechanisms of speech in particular; the genetics of motor learning (automation) and of language disorders in general, and of speech motor learning and phonology in particular. Each of the three previous conferences has resulted in a book, containing chapters written by key presenters at those meetings. For the last book (van Lieshout, Maassen, & Terband, 2016), authors were invited to give their view on future research: where will the field be ten years from now? This book is published both as hardcopy and as e-book, and is available at the conference.

It has been the aim of the International Conference on Speech Motor Control to not only give an overview of the current state-of-the-art, but most importantly, to highlight (and where necessary, speculate) how in the next decade research in these areas will influence our notions about speech motor processes and how this will have implications for future studies and/or clinical procedures. A special topic of the 7th conference is inspired by the booming field of genetics, in particular with respect to the evolution of speech and language. Insight in the genetics of homo sapiens as compared to other primates, opens a window to study how speech and language came into being in our species, and what special characteristics distinguishes homo sapiens from other primates such that speech could emerge. The conference is opened by two keynote speakers addressing the history of evolution research and recent insights and theories. For the mainstream conference, we encourage participants to take speech research again a step further and report on their research on the origin, development and maintenance of cognitive, linguistic and motor processes that together determine humans most complex motor skill: speaking.

Books of previous editions of the conference

Maassen, B., Kent, R., Peters, H.F.M. van Lieshout, P.H.H.M. & Hulstijn, W. (Eds.) (2004). *Speech motor control in normal and disordered speech.* Oxford, UK: Oxford University Press.

Maassen, B., & Van Lieshout, P. (Eds) (2010). *Speech motor control: New developments in basic and applied research.* Oxford, UK: Oxford University Press.

Van Lieshout, P., Maassen, B., & Terband, H. (Eds.) (2016). *Speech Motor Control in normal and disordered speech: Future developments in theory and methodology.* Rockville, MD: ASHA.

Conference organization

In order to fulfil the main purpose of the conference a relatively large number of keynote speakers have been invited to present tutorials on specific topics. All presentations are plenary to stimulate a lively interaction. Due to time constraints, only a very limited number of submissions could be scheduled as oral presentations. Thematic poster sessions therefore form a major part of the conference program, offering a large variety of research in speech motor control in typical and atypical speech from all over the world. Many conferences advocate the policy to value oral presentations and posters equally, as do the orgazisers of this conference. In order to underscore this policy, ample time is scheduled for the poster sessions and a special prize is awarded for the most informative and well-designed poster.

The University of Groningen and the organizing departments are proud to attract such high-level researchers and clinical workers in the field to travel to Groningen and report on the results of their theoretical and empirical work at this platform of scientific exchange and discussion.

We look forward to a stimulating and productive conference,

Ben Maassen Groningen
Hayo Terband Utrecht
Pascal van Lieshout Toronto

June 2017

Program & Organizing committee

- Ben Maassen, chair (CLCG & BCN, University of Groningen)
- Hayo Terband, co-chair (Utrecht Institute of Linguistics - OTS, Utrecht University)
- Pascal van Lieshout (Oral Dynamics Lab, Department of Speech-Language Pathology, University of Toronto, Canada)
- Edwin Maas (Department of Communication Sciences and Disorders, Temple University, Philadelphia, USA)
- Aravind Namasivayam (Oral Dynamics Lab, Department of Speech-Language Pathology, University of Toronto, Canada & Prompt Institute, Santa Fé, NM, USA)
- Frits van Brenk (Department of Speech and Language Therapy, University of Strathclyde, Glasgow, UK)

Venue

The conference will be held in the Academy Building of the University of Groningen, in the heart of the city-centre of Groningen. Visiting address: Broerstraat 5, Groningen.

Secretariat of the Conference

Groningen Congres Bureau
Att. Ms. Christien Koopstra
Address: Griffeweg 5
Postcode: 9724 GE Groningen
The Netherlands
Telephone: +31 (0) 50 3168877
E-mail: info@groningencongresbureau.nl

Conference website

www.slp-nijmegen.nl/smc2017

Oral presentations

THE GENETIC STANCE IN STUDYING THE EVOLUTION OF SPEECH. SOME HISTORICAL REMARKS.

Willem Levelt
Max Planck Institute for Psycholinguistics, Nijmegen, The Netherlands

How did speech and language evolve? In their important book *Creating Language* (2016) Christiansen and Chater approach this issue from what I have called the "genetic stance". It is the comparative study of language genesis on three time scales: in speech (microgenesis), in language acquisition (ontogenesis) and in evolution (phylogenesis). Unknowingly, the authors recapitulate the perspective that was common ground for language scholars since the mid-18th century till the end of the 19th century.

That genetic stance provided two windows on the evolution of speech. The first and dominant one was to use the ontogenesis of speech as a model for speech evolution. I will discuss the quite explicit theory by French encyclopedist Charles de Brosses, who took the emergence of speech in both cases as a "mechanical" (i.e. physiological, biological) necessity. This view was shared by Gottfried Herder. But others opposed this nativist view, considering speech and language as an art, a human invention. Most outspoken was Lord Monboddo, who claimed that "articulation is not natural to man". Wolfgang von Kempelen, the greatest speech scientist of the 18th century, argued that language is invented time and again and that children have an important role to play. He had observed the stepwise creation of a new sign language in the Institute for the Deaf in Paris. The children "speak in their sign language with admirable skill about the most abstract entities."

The ontogenetic perspective on the evolution of speech strongly resurged during the second half of the 19th century. There was Fritz Schultze's theory of least effort, which explains the order in which consonants are acquired by the child, and potentially by ever more civilized languages. A short-lived, but hefty theoretical movement, triggered by Haeckel's "biogenetic law", reversed the ontogenetic perspective: the ontogenesis of speech and language recapitulates the phylogenesis of language. George Romanes, however, carefully turned the argument around again, stressing the role of the child's "language instinct" in the ever new creation of languages. Here, he adopted the inventive theory of Canadian anthropologist Horatio Hale, who sketched a scenario in which the primitive babbles and speech of children could cause the emergence of new languages. This was full circle back to De Brosses, who had however been forgotten by then.

The second window on speech evolution was the microgenetic one. Speech evolved from the expressive movements of our primordial ancestors. This perspective was introduced by Heymann Steinthal during the 1850s, but then fully reformulated by Wilhelm Wundt in Die Sprache (1900). Wundt tried to explain how intrinsically meaningless vocal sounds would acquire meaning by association to the universally used meaningful gesturing in primordial societies. I will discuss the (un)tenability of Wundt's explanation by means of a unique video from Rossel Island in the Pacific.

THE EVOLUTION OF VOCAL CONTROL

John L. Locke
Lehman College, City University of New York, New York, United States of America

Although language is universally spoken, most evolutionary proposals say little about any changes that may have facilitated the emergence of a trait famously lacking in the other primates - vocal control - a precondition to the emergence of words and grammar, thus of critical importance. To achieve such a proposal, one must identify factors that may have played a role in the evolution of a vocal delivery system, a process that, for me, has been facilitated and constrained by data from nonlinguistic species and prelinguistic human infants as well as speaking practices in traditional societies. Since new traits can only appear in development, I will suggest that evolutionary changes in human life history increased the importance of vocalization and vocal interactions in two different stages of development.

In the first stage of life history, I propose that paired increases in infant helplessness and maternal fertility produced new levels of competition for care, causing infants to vocalize in ways that gave receptive caregivers needed information about fitness. This change effectively placed the vocal behavior of infants under the perceptual control of adults. I also suggest that a second adaptation, the adoption of cooperative breeding, increased the diversity and functionality of social vocalization; and that parental tendencies to mimic the more complex vocalizations of infants, and to vocalize in synchrony with them, facilitated the ability of their offspring first to repeat, then to initiate articulated vocalizations, that is, to produce speech-like forms intentionally. These interactions between infants and caregivers, I claim, collectively emancipated the voice from affective control.

My second claim is that the remodeling of life history, in conjunction with intensified sociality, continued to enhance vocal flexibility and complexity in later stages of development, including adolescence and early adulthood, most evidently in the context of dominance displays and mate selection. Recently it was reported that in gelada monkeys - as in songbirds - males display significantly more complex vocalizations than females, and in playback experiments, female geladas display a preference for these more complex vocalizations. In our own species, there is evidence that young men tend to "ornament" their speech in a mating or competitive context - whether vocally, lexically, or structurally - and thereby enjoy unusual advantages in their pursuit of social dominance and reproductive opportunities.

References

Locke, J. L., & Bogin, B. (2006). Language and life history: a new perspective on the evolution and development of linguistic communication. *Behavioral and Brain Science, 29*, 259-311.

Locke, J. L. (2017). Emancipation of the voice: vocal complexity as a fitness indicator. *Psychonomic Bulletin and Review, 24*, 232-237. doi:10.3758/s13423-016-1105-7

THE BOUNDARY BETWEEN PERCEPTUAL AND MOTOR LEARNING

David Ostry
McGill University, Montréal, Canada

There is accumulating evidence that perceptual and motor learning do not occur in isolation. Rather motor learning affects perceptual judgements and changes sensory brain areas. Similarly, perceptual learning changes movements and motor areas of the brain (Ostry and Gribble, 2016) . In this talk, I will summarize recent work in support of these ideas drawing on behavioral and neuroimaging data from studies involving both speech motor learning and human arm movement (Vahdat et al., 2010). I will also present the results of recent work on speech motor learning in which we use a robotic device to selectively alter somatosensory input during speech. This technique is combined with a new resting-state neuroimaging analysis based on partial correlation that identifies from among the set of areas that encode learning those whose functional connectivity is both strengthened with learning and cannot be attributed to activity in other parts of the speech network. When we remove the signal attributable to other brain regions, we observe novel roles for sensory and motor systems in learning. Specifically, we find that it is largely sensory areas that survive the partial correlation test in the context of speech motor learning. That is, we find that adaptation to altered somatosensory feedback is largely accounted for by changes in connectivity in nonmotor areas, between auditory and somatosensory cortex, and between the inferior parietal cortex and pre-supplementary motor area. In contrast, we find that changes in the perceptual classification of speech sounds that occurs in conjunction with learning are primarily attributable to functional connectivity changes in cortical motor areas, between primary motor cortex and the inferior frontal gyrus. There are several notable aspects of these new results. First, it is seen that the technique results in activity in an extensive set of areas that reflect brain functional reorganization due strictly to somatic inputs (with the auditory signal held constant). This points to an elaborate cortical network in speech learning that could be substantially somatosensory in origin. More generally, it is seen that the traditional assumptions about brain areas involved in speech perception and production are reversed. In particular, this new work supports the idea that plasticity in sensory rather than motor brain areas is substantially responsible for speech motor learning.

References

Ostry D. J., Gribble P. L. (2016). Sensory plasticity in human motor learning. *Trends Neurosci, 39*, 114-23.

Vahdat S., Darainy M., Milner T. E., Ostry D. J. (2010). Functionally specific changes in resting-state sensorimotor networks after motor learning. *J Neurosci, 31*, 16907-15.

THE MOTOR SYSTEM'S ROLE IN COGNITION AND UNDERSTANDING

Friedemann Pulvermüller
Freie Universtät Berlin, Berlin, Germany

A classic perspective on the brain is this: Perceptual systems filter information from the input and direct it to higher centers, which fabricate cognition; the cognitive modules, in turn, use motor systems as slaves to move the body. The discovery of Mirror Neurons active during perception sparked an interest in the cognitive functions of motor systems, although skeptics questioned their causal role. Today, much evidence supports a shared contribution of sensory, motor and multimodal areas of cortex to cognition. That motor systems play a key cognitive role receives support from at least two arguments:

1. When we move, neuronal activity is not only present in the motor system, but, because we always perceive aspects of our movements, in perceptual systems too, thus leading to correlated nerve cell activity across sensory and motor areas. Because in particular primates are richly endowed with long-range connections in the cortex, such correlated sensorimotor activity leads to the formation of functional links and distributed neuronal circuits in which motor neurons are embedded (Pulvermüller et al., 2014). Current neuroscience data do indeed support a functional role of motor areas in cognition, as, for example, in language understanding (Schomers et al., 2016).

2. The meaning of symbols in many cases draws on specific motor and action-related information. This can be exemplified by words such as "(to) grasp", which relates to a complex set of specific object-related movements, or "(to) free", which relates to a broad class of context-constrained actions. When understanding such words, and even when predicting them in sentence context, human subjects activate their motor system, with local motor cortex activity indexing aspects of the activated meanings (Grisoni et al., 2016). Further work demonstrated the causal role of motor cortex in semantic processing.

Taken together, these results show that the motor system contributes to much more than motor execution. However, this system does not seem to house a unique module for perception and understanding either. More likely, distributed neuronal circuits in which motor neurons play a key part are the carriers of cognition, specifically-human language mechanisms included (Pulvermüller et al., 2014).

References

Pulvermüller, F., Garagnani, M. & Wennekers, T. (2014). Thinking in circuits: Towards neurobiological explanation in cognitive neuroscience. *Biol Cybern, 108*, 573-593.

Schomers, M. & Pulvermüller, F. (2016). Is the sensorimotor cortex relevant for speech perception and understanding? An integrative review. *Front Hum Neurosci, 10*, 435.

Grisoni, L., Dreyer, F. R. & Pulvermüller, F. (2016). Somatotopic semantic priming and prediction in the motor system. *Cereb Cortex, 26*, 2353-2366.

ADVANCES IN MODELING SPEECH PRODUCTION AS STATE FEEDBACK CONTROL

John F. Houde[1], Srikantan S. Nagarajan[1], Benjamin Parrell[2], Vikram Ramanarayanan[3]
[1] *University of California, San Francisco, United States of America*
[2] *University of Delaware, Newark, United States of America*
[3] *University of Southern California, San Francisco, United States of America*

For several years now, our lab has been examining how well state feedback control (SFC) models the control of speaking (Houde and Nagarajan, 2011). In brief, speaking involves changing the dynamical state (the position and velocity) of the vocal tract articulators and in SFC, speech is controlled by maintaining a running estimate of this dynamical state. Importantly, this state estimate is maintained entirely within the CNS, with sensory feedback exerting only a corrective influence on this internal estimate.

Our prior work has shown how SFC accounts for many of the behavioral phenomena associated with the role of sensory feedback in speech production as well as many of the neural phenomena associated with sensory processing during speaking (Houde et al., 2014). We are now examining how SFC accounts for the abnormal speech feedback processing seen in various neurological conditions. We have found that many conditions are associated with abnormally large compensatory responses to unexpected pitch (16p11 autism, Alzheimer's disease, and cerebellar ataxia) and formant (cerebellar ataxia) feedback perturbations, as well as abnormally small adaptive responses to sustained formant feedback alterations (16p11 autism, cerebellar ataxia). These results suggest a degree of independence between the mechanisms mediating long-term sensorimotor adaptation and the mechanisms generating immediate compensatory responses. In cerebellar ataxia in particular, simulations with our SFC model have led us to hypothesize that cerebellar damage principally only affects the internal predictive models governing sensorimotor adaptation, and that the large compensatory responses we observe result from a subsequently-learned over-reliance on sensory feedback.

Most recently, we have also begun to develop a hierarchical extension of SFC by combining it with the Task Dynamics (TaDA) model of speech production (Ramanarayan et al., 2016). In the combined model, gestural scores of the TaDA model drive the state feedback control law governing utterance production, and current vocal tract state is expressed in terms of task (constriction) state, which is estimated from both somatosensory and auditory feedback. The resulting model combines the strengths of both TaDA and SFC: it accounts for the sensitivity that speakers exhibit to changes in sensory feedback (including auditory feedback), and it is able to generate a more complete range of speech output.

References

Houde, J.F. and Nagarajan, S.S. (2011). Speech production as state feedback control. *Frontiers in Human Neuroscience, 5,* 82.

Houde, J.F., et al., (2014). Simulating a state feedback model of speaking. In *10th International Seminar on Speech Production.* Cologne, Germany.

Ramanarayanan, V., et al. (2016). A New Model of Speech Motor Control based on Task Dynamics and State Feedback. In *Proceedings of Interspeech 2016.* San Francisco, CA: International Speech Communication Association.

TRANSFER OF SKILLED BEHAVIOR ACROSS DIFFERENT VOCAL TRACT ACTIVITIES: THE INFLUENCE OF SPEECH PRODUCTION ON BRASS INSTRUMENT PERFORMANCE

Matthias Heyne[1], Bryan Gick[2,3], Donald Derrick[1]
[1] *University of Canterbury, Canterbury, New Zealand*
[2] *University of British Columbia, Vancouver, Canada*
[3] *Haskins Laboratories, New Haven, United States of America*

While speech motor control has traditionally been regarded as domain-specific or organized in a vertical manner, recent studies have postulated cross-system interactions that support a modular organization (d'Avella et al., 2015) of speech motor control. Extending claims that phylogenetically-encoded structures such as swallowing can influence speech learning, this presentation provides evidence for the transfer of vocal tract behavior from one form of skilled behavior (speech production) to another (brass playing).

We used ultrasound imaging of the tongue to record midsagittal tongue contours of 10 Tongan and 9 New Zealand English-speaking trombone players during speech production in their native language and while producing sustained notes on the trombone. Results normalized to account for differences in vocal tract shape and ultrasound transducer orientation suggest that motor memory of vowel tongue positions interacts with other constraints related to airflow, acoustical demands involving the interaction of vocal tract and instrument bore resonances, and motor efficiency considerations, to produce significant differences at the back and front of the tongue. These differences pattern independently, supporting accounts of the functional independence of various sections of the tongue, and a modular organization of speech motor control.

Gick and colleagues (Gick et al., 2016) have recently outlined how a modular theory of speech production might help solve the well-known degrees of freedom problem that has long posed a central challenge for theories of motor control. One important step for solving this challenge might be to abandon the notion of motor control as global optimization, regarding it instead as a search for locally-optimized solutions that address the demands of a given task (Ganesh et al., 2010), with the potential of explaining some of the abundant individual variation observed in speech production.

Our data provides evidence of such a process at work during the transfer of muscle synergies from speech production to trombone playing. When first starting to play a brass instrument, a beginning player's vocal tract musculature faces the challenge of coming up with a way of initiating, and channeling, the required airflow into the instrument. Developing a completely new sensorimotor program through a process of trial and error would be costly so that assuming a vowel tongue positon from one's native language, with subsequent local optimization, might provide the best strategy in such a situation (cf. Ganesh et al., 2010).

References

d'Avella, A., Giese, M., Ivanenko, Y. P., Schack, T., & Flash, T. (2015). Editorial: Modularity in motor control: From muscle synergies to cognitive action representation. *Frontiers in Computational Neuroscience, 9*, 126.

Gick, B. (2016). Ecologizing Dimensionality: Prospects for a Modular Theory of Speech Production. *Ecological Psychology, 28*(3), 176-181

Ganesh, G., Haruno, M., Kawato, M., & Burdet, E. (2010). Motor memory and local minimization of error and effort, not global optimization, determine motor behavior. *Journal of Neurophysiology, 104*(1), 382-390.

CHANGES IN RESTING STATE FUNCTIONAL CONNECTIVITY DUE TO SPEECH MOTOR LEARNING

Saloni Krishnan[1], Jennifer Chesters[1], John-Stuart Brittain[1], Ned Jenkinson[2], Daniel Lametti[1], Kate E. Watkins[1]
[1]*University of Oxford, Oxford, United Kingdom*
[2]*University of Birmingham, Birmingham, United Kingdom*

Introduction Brain regions continuously interact even when the brain is "at rest". These resting-state networks are determined by examining the correlations between spontaneous fluctuations in brain activity over time. The major resting-state networks are observed consistently and they closely correspond to brain dynamics during overt task behaviour (Smith et al., 2009). There is evidence to suggest that these networks are shaped by learning and consolidation. For instance, Albert, Robertson, and Miall (2009) demonstrated that activity in cerebellar and frontoparietal networks was modulated by learning of a new visuomotor relationship in a visuomotor adaptation task. In this study, we use a speech adaptation task to examine how learning a new auditory-motor mapping modulates resting network activity. We hypothesised that sensorimotor, auditory, and cerebellar networks are most likely to show changes in connectivity following learning.

Methods For all 13 participants in the study, functional brain activity during rest was recorded using MRI across two sessions. One session [Rest^{A}] immediately followed a speech motor-learning task in which participants compensated for altered auditory feedback (shift in the frequency of the first formant) by changing how they produced speech. The other was completed on a different day when no speech adaptation task was performed [Rest^{C}]. Session order was counterbalanced across participants. For each MRI dataset, we identified a set of reference networks previously described by Smith et al. (2009) using the FMRIB Software Library (FSL)'s dual regression tool. To identify networks that changed in strength of connectivity following the speech adaptation task we estimated the intrinsic connectivity in each of these subject-specific networks. In addition, a wholebrain voxel-wise comparison of each resting-state network before and after speech motor adaptation was made using a t-test and the significance of these statistics was evaluated using non-parametric permutation testing (1000 permutations) and threshold-free cluster enhancement ($p < 0.05$, corrected).

Results & Discussion Following speech adaptation, activity in left Heschl's gyrus (auditory cortex) showed increased connectivity with a right frontoparietal network (Figure 1A, osf.io/2gx8m) that is typically activated by action-inhibition and perception-somesthesis tasks. The degree to which individual participants adapted was negatively correlated with activity in another network of areas active in auditory tasks (Figures 1B & C, osf.io/2gx8m). Individuals with stronger connectivity in this resting-state network adapted less to the speech perturbation. These findings indicate that learning a new auditory-motor mapping relates to the functional connectivity of auditory regions, and results in auditory cortex activity becoming more strongly connected with functional components involved in action and perception.

References

Albert, N. B., Robertson, E. M., & Miall, R. C. (2009). The resting human brain and motor learning. *Current Biology, 19*(12), 1023-1027.

Smith, S. M., Fox, P. T., Miller, K. L., Glahn, D. C., Fox, P. M., Mackay, C. E., ... Beckmann, C. F. (2009). Correspondence of the brain's functional architecture during activation and rest. *Proceedings of the National Academy of Sciences, 106*(31), 13040-13045.

A MASSLESS 3D BIOMECHANICAL MODEL OF THE TONGUE AND ITS RELATION TO THE λ MODEL

Alan Wrench[1,2], Peter Balch
[1]*Articulate Instruments Ltd, Musselburgh, United Kingdom*
[2]*Queen Margaret University, Edinburgh, United Kingdom*

We report on an ongoing project which aims to find a robust, stable 3D biomechanical modelling methodology that allows arbitrary mesh configurations and muscle fibre alignments to be readily reconfigured and tested and one which exhibits the flexibility and movement observed in a real tongue. The widely favoured Finite Element Method (FEM) is known to exhibit stiffness and volumetric locking (Latash, 2012). We have chosen to take a different approach and are investigating a massless model based on a hexahedral mesh and a modified implementation of Hooke's law where the elasticity varies with spring (muscle) length. We have used our knowledge of the anatomy of the tongue to create a 3D hexahedral mesh and assign muscle fibres within it. These muscle fibres can be grouped and controlled as a group by specifying the muscle length at which no force is generated.

Validation experiments, where the model is fitted to midsagittal ultrasound data and the predicted parasagittal shapes compared with electropalatography data for the same utterance, are encouraging.

In this paper we set aside the spatial aspects of mesh and muscle assignments and focus on the dynamics of the model. The tongue consists of skeletal muscle, so we refer to historical research on the physiology of motor neurons, muscle spindles and the stretch reflex to build our model. Evidence summarized nicely by Latash (2012) suggests that the muscle spindle feedback loop, which is independent of the Central Nervous System (CNS), serves to find a balance between the competing contraction signals sent by the CNS to different functional neuromuscular compartments (NMCs). What is more, it takes a finite amount of time to reach this state of balance or equilibrium. We have implemented a Newton-Raphson type iterative approach to mimic the force balancing process. We note strong similarities between the way our model works and both the α-model and λ-model variants of equilibrium point control (Latash, 2012, Ch. 4). Our model further suggests that fluid control of movement can be achieved by a sparse series of CNS signals where the velocity profiles observed in real movements are not necessarily due to continuous updating of equilibrium points by sensory feedback to the CNS but might be due to a combination of the timing of CNS signals and the process of settling to equilibrium.

We ask whether an expanded and refined version of our model might provide new insight into the patterning of neuromuscular control signals in continuous speech.

References

Latash, M. (2012). *Fundamentals of motor control* (1st Ed.). Elsevier.

Rohan, P-Y., Lobos, C., Nazari, M. A., Perrier, P. & Payan Y. (2014). Finite element modelling of nearly incompressible materials and volumetric locking: a case study. *Computer methods in biomechanics and biomedical engineering. 17*(sup1), 192-193.

MULTIVARIATE ANALYSIS FOR LARGE ARTICULOGRAPHY DATASETS OF SPEECH AND INDUCED SPEECH ERRORS

Gabriel J. Cler, Jackson C. Lee, Talia Mittelman, Cara E. Stepp, Jason W. Bohland
Boston University, Boston, United States of America

Speech errors are often evaluated via auditory-perceptual analysis. However, errors judged to be categorical (e.g., perceived as exchanges of two phonemes) may actually represent blends of sounds or articulatory errors (Pouplier & Hardcastle, 2005). Analyses of the kinematics of error-free and errorful speech may help characterize the nature and genesis of individual errors and further elucidate speech motor control. Here, delayed auditory feedback (DAF) was used to induce disfluencies and speech errors (Fairbanks, 1955). Eight healthy speakers each generated a large dataset containing both error-free and errorful productions (mean 2939 productions of a closed set of six syllables) with and without DAF. Kinematics were captured with electromagnetic articulography using five sensors on the lips and tongue.

Time-locked acoustic recordings were transcribed, and the kinematics of utterances were analyzed with existing and novel quantitative methods. EMA data corresponding to individual speech tokens were converted to a high-dimensional feature vector, with each element representing the location of one sensor in one dimension at one time point.

For syllables perceived as error-free, a non-metric multidimensional scaling visualization approach was applied. In these vizualizations, the distance between each pair of data points (each representing one syllable) is related to the distance between data points in the high-dimensional feature space, reflecting the overall similarity of two productions. The visualization algorithm does not consider each token's transcribed identity, but the color and marker for each production were indicated *post hoc* to reflect their perceived class as well as to distinguish tokens produced with and without DAF. Non-error syllables produced with DAF show more kinematic variability than non-error syllables produced without DAF; this variability was compared to a commonly used measure, the spatiotemporal index (STI).

A machine-learning approach was used to classify errorful syllables into *error profiles*, in which each error production was classified against non-error data from all six syllable types. 43% of categorical errors (those perceived as a correctly-formed syllable that was different than the stimulus syllable) were classified as kinematic matches to only the perceived syllable. Interestingly, 20% of the categorical error syllables were classified as matches to both the perceived and stimulus syllable, suggesting that these may have been co-productions or articulatory blends of two syllables.

Relative strengths of the multivariate methods presented in comparison to typical methods (e.g., STI) are considered. The proposed measures, associated visualizations, and analysis approaches may be of general utility for visualization and characterization of other large kinematic datasets.

References

Fairbanks, G. (1955). Selective vocal effects of delayed auditory feedback. *The Journal of Speech and Hearing Disorders, 20*(4), 333-346.

Pouplier, M., & Hardcastle, W. (2005). A re-evaluation of the nature of speech errors in normal and disordered speakers. *Phonetica, 62*(2-4), 227-43.

AUDITORY MODULATION DURING SPEECH PLANNING IN TYPICAL SPEAKERS AND INDIVIDUALS WHO STUTTER

Ludo Max[1,2], Ayoub Daliri[3]
[1] *University of Washington, Seattle, United States of America*
[2] *Haskins Laboratories, New Haven, United States of America*
[3] *Arizona State University, Tempe, United States of America*

This presentation summarizes a series of electroencephalographic (EEG) studies investigating modulation of the auditory system during the planning phase immediately prior to speech onset (Daliri & Max, 2015, 2016). All studies included typical speakers; some studies also included stuttering speakers. Study I was a methodological study in which we developed the basic paradigm. In this paradigm, participants read aloud words presented during a delayed-response speaking task. Long-latency auditory evoked potentials (LLAEPs) are recorded in response to probe tones that are presented during speech movement planning in the speaking condition and during linguistic and non-linguistic control conditions without speaking. Typical adults show a statistically significant modulation of the N1 component in the LLAEP when preparing to speak as compared with reading the same words silently or seeing non-linguistic symbols. Study II included both typical and stuttering adults, and used high-density EEG recordings that allowed topographic analyses. Results replicated the prior finding of pre-speech N1 modulation for typical speakers but showed a lack of such modulation for the group of stuttering speakers. Study III investigated in typical speakers the effect of using pure tones versus truncated syllables as probe stimuli during the same delayed-response oral reading task. Results indicated that the auditory N1 component is modulated equally for tone and syllable stimuli, but a subsequent modulation for the higher-order P2 component is observed only with syllable stimuli. Study IV investigated the functional relevance of pre-speech auditory modulation, and again included both typical and stuttering speakers. Specifically, this study explored (a) a potential relationship between such auditory modulation and auditory feedback contributions to speech sensorimotor learning and (b) the effect on pre-speech auditory modulation of real-time versus delayed auditory feedback. Acoustic and electrophysiological data revealed that adults who stutter show deficits in both pre-speech auditory modulation and auditory-motor learning. However, limited pre-speech modulation is not directly related to limited auditory-motor adaptation. Of particular interest in light of clinical evidence indicating that many adults who stutter become more fluent when speaking with altered auditory feedback, delayed feedback paradoxically normalized the stuttering group's otherwise limited pre-speech auditory modulation. Lastly, in order to allow future studies with young children as participants, Study V demonstrated that the same pre-speech auditory modulation phenomenon can also be observed in a picture naming task. Overall, the results from this series of studies provide intriguing new insights into processes underlying sensorimotor interactions during speech motor planning.
[NIH/NIDCD R01DC007603, R01DC014510]

References

Daliri, A., & Max, L. (2015). Modulation of auditory processing during speech movement planning is limited in adults who stutter. *Brain and Language, 143*, 59-68.

Daliri, A., & Max, L. (2016). Modulation of auditory responses to speech vs. nonspeech stimuli during speech movement planning. *Frontiers in Human Neuroscience, 10*, 234.

EVOLUTION OF SPEECH: ANATOMY AND CONTROL

Bart de Boer
Vrije Universiteit Brussel, Brussels, Belgium

The modern human vocal tract is rather different from the vocal tracts of other primates (and other mammals for that matter). It has a rounder tongue, a lower larynx and has no air sacs. I have argued that these modifications are due to evolutionary pressure related to speech (e. g. de Boer, 2012). However, research modeling the vocal capacities of a macaque has shown that even a monkey vocal tract can in principle produce a sufficient range of different speech sounds for language (Fitch et al. 2016). These two findings are not contradictory: even though the ancestral vocal tract was capable of producing a reasonable range of speech sounds, it is still possible that the modern human vocal tract has been fine-tuned under selective pressure related to language and speech. However, it does indicate that the changes to the vocal tract are not the primary factor indicating or explaining the evolution of language. The true adaptations must have been cognitive: increased control, the ability to imitate and the ability to analyze and reproduce speech using a combinatorial set of building blocks (e. g. de Boer, 2017).

This talk will present on overview of work on modeling the evolution of the vocal tract (with a focus on my own work) as well as examples of ongoing work investigating whether we can find evidence for selective pressure related to speech on language on cognitive adaptations, for instance in our ability to learn the building blocks of speech, or in the ability to finely control the vocal tract.

References

de Boer, B. (2012). Loss of air sacs improved hominin speech abilities. *Journal of Human Evolution, 62*(1), 1-6.

de Boer, B. (2017). Evolution of speech and evolution of language. *Psychonomic Bulletin & Review, 24*(1), 158-162.

Fitch, W. T., de Boer, B., Mathur, N., & Ghazanfar, A. A. (2016). Monkey vocal tracts are speech-ready. *Science Advances, 2*(12).

NEW ADVANCES IN OUR UNDERSTANDING OF SPEECH EVOLUTION

Adriano R. Lameira
St. Andrews University, United Kingdom
Durham University, United Kingdom

The evolution of spoken language is one of the oldest questions in human thought, and science still has a very pale idea of how proto-speech stages may have looked like in human ancestors. Primate bioacoustics and vocal behaviour in our closest relatives - nonhuman great apes - represent a rich source of information on the probable composition of the ancestral great ape call repertoire that predated and putatively acted as precursor to human speech. Here, I illustrate and explain how the long-term inventory of the call repertoire of wild great apes, with a focus on orang-utans, together with innovative experimental paradigms in captivity, is offering new heuristic windows into the original building blocks of speech - proto-consonants and proto-vowels. Notably, great apes exhibit a much more advanced level of voluntary motor control than hitherto assumed - a trait that allows them to acquire new calls and expand their call repertoire through vocal learning. Although on-going research is still narrowing down the conditions wherein great apes are motivated to learn vocally and from whom, this capacity manifests in the form of individual-specific calls in captivity and population-specific vocal traditions in the wild. These findings spawn a new generation of questions that will help closing the gap between our last great ape common ancestor and humans, and bridge the vocal system of the former with the verbal system of the latter.

References

Lameira, A. R., Vicente, R., Alexandre, A., Campbell-Smith, G., Knott, C., Wich, S., & Hardus, M. E. (2017). Proto-consonants were information-dense via identical bioacoustic tags to proto-vowels. *Nature Human Behaviour, 1*(2), 0044. http://doi.org/10.1038/s41562-017-0044

Lameira, A. R., Hardus, M. E., Mielke, A., Wich, S. A., & Shumaker, R. W. (2016). Vocal fold control beyond the species-specific repertoire in an orang-utan. *Scientific Reports, 6*, 30315. http://doi.org/10.1038/srep30315

Lameira, A. R., Hardus, M. E., Bartlett, A. M., Shumaker, R. W., Wich, S. A., & Menken, S. B. J. (2015). Speech-like rhythm in a voiced and voiceless orangutan call., *PLOS ONE, 10*(1), e116136. http://doi.org/10.1371/journal.pone.0116136

MONKEY VOCALIZATIONS: NEW DATA, NEW TOOLS, AND NEW HYPOTHESES

Louis-Jean Boë
GIPSA-Lab, Grenoble, France
CNRS, Grenoble, France
Grenoble Alpes University, Grenoble, France

Since humans share common ancestors with both apes and monkeys, current vocalizations of these primates provide us with an underexploited window for exploring the nature of (human) speech. Recent new analyses of their anatomy and of the acoustics of their vocalizations allow us to challenge old hypotheses on the emergence of speech. In that vein, we can examine vocalizations of present-day monkeys as *relics* of earlier vocal abilities and, metaphorically, *fossil traces* of various aspects of the communication of our common ancestors. For instance, an orang-utan has been shown to imitate the pitch of a human instructor's voice, thus demonstrating voluntary vocal fold control (Lameira et al., 2016). In macaques, X-rays during facial displays, feeding, and vocalization show an array of vocal tract configurations achieved through tongue shaping and lip gestures (Fitch et al., 2016). Some of these are equivalent to vowel configurations, as has been shown by comparing swallowing and human feeding.

Dissection of male and female Guinea baboons showed that they have the same laryngeal and articulatory musculature as humans. We analyzed 1335 vocalizations produced by 3 male and 12 female Guinea baboons in different ethological situations (grunt, bark, yak, copulation call, & wahoo) (Boë et al., 217). The acoustic analysis reveals a wide F0 range (more than four and a half octaves) and five distinct classes of vowel-like. A new technique for modeling the potential capacities of their vocal tract, namely simulation of the Maximal Acoustic Space, showed they are comparable to human vowels [ɨ æ ɑ ɔ u], and are organized in a homologous proto-vocalic system.

This preponderance of evidence suggests the emergence of speech in humans was neither sudden nor *de novo*. Speech evolved from ancient articulatory skills already present in our last common ancestor with Cercopithecoidea, about 25 MYA.

References

Boë, L. J., Berthommier, F., Legou, T., Captier, G., Kemp, C., Sawallis, T. R., Becker, Y., Rey, A., & Fagot, J. (2017). Evidence of a vocalic proto-system in the baboon (*Papio papio*) suggests pre-hominin speech precursors. *Plos One*, doi:10.1371/journal.pone.0169321

Fitch, W.T., de Boer, B., Mathur, N., & Ghazanfar, A. A. (2016). Monkey vocal tracts are speech-ready. *Science Advances 2*, e1600723. doi: 10.1126/sciadv.1600723

Lameira, A.R., Hardus, M.E., Mielke, A., Wich, S. A. & Shumaker, R. W. (2016). Vocal fold control beyond the species-specific repertoire in an orang-utan. *Scientific Reports*, doi: 10.1038/srep30315

THE SPECIFICITIES OF CHIMPANZEE VOCALISATIONS AND THEIR RELATIONSHIP TO THE VOCAL APPARATUS

Marion Laporte[1], Philippe Martin[2]
[1]*Université Paris 3, Paris, France*
[2]*Université Paris Diderot, Paris, France*

The production of vocalisation in nonhuman primates is often taken as an example in the quest for the understanding of the evolution of human language. For example, models have been made to look for the potential of macaques' vocal tracts to produce speech-like sounds (Boë et al., 2017; Fitch et al., 2016) but meanwhile, the natural production of vocalisations in primates and especially apes, remains still poorly understood.

It is generally admitted that the source-filter theory of human phonation (Fant, 1960), also applies to most mammals. Essentially, the theory states that vocal signals are produced as a result of the independent contribution of the "source" (the laryngeal structure producing the glottal waveform or F0) and the "filter" (the supralaryngeal vocal tract, shaping the glottal waveform into formants). This theory thus permits to link the acoustic variation with the anatomy of the caller using the common formula Fn= (2n-1)c/4L, which takes each formant separately. However, the first three formants can better be considered as interrelated and the relationship takes the form of cotan-tan equality with the frequency imbedded in the trigonometric functions (Fant, 1960). It is with this equation that we analysed some typical vocalisations of one of our closest living relatives, the chimpanzee.

Data on wild chimpanzees were collected in 2007, 2008 and 2012 in the Sonso community of Budongo Forest, Uganda. Nine adult males and ten adult females were followed using focal animal sampling with continuous recording. The acoustical analysis was performed with the WinPitch software, with specifically dedicated functions. The chimpanzees' vocal repertoire presents both very harmonic and noisy vocalisations and very few formants. We calculated the 3 theoretical formants with the above equation from the measures of the vocal tract taken on the radiography of an adult female chimpanzee. We found formants mostly around high harmonics of the fundamental frequency. We discuss the possibility of the coupling between the source and the filter or the presence of another source (resulting of a tension of the cartilaginous glottis) in the light of the chimpanzee's vocal apparatus, which, while being broadly similar to that of humans, also presents key differences.

References

Boë, L.-J., Berthommier, F., Legou, T., Captier, G., Kemp, C., Sawallis, T. R., Fagot, J. (2017). Evidence of a Vocalic Proto-System in the Baboon (*Papio papio*) Suggests Pre-Hominin Speech Precursors. *PLOS ONE, 12*(1), 1-15.

Fant, G. (1960). *Acoustic theory of speech production.* The Hague: Mouton.

Fitch, W. T., de Boer, B., Mathur, N., & Ghazanfar, A. A. (2016). Monkey vocal tracts are speech-ready. *Science Advances, 2*(12), 1-7.

A MODEL OF THE EMERGENCE OF COORDINATIVE CONTROL

Sam Tilsen
Cornell University, Ithaca, United States of America

A variety of evidence supports the idea that speech motor learning involves the development of internal models of sensory feedback. In early stages of development, children rely primarily on external sensory feedback, and this promotes a competitive organization of articulatory movements. Subsequently children develop internal models of the predicted sensory consequences of motor actions, allowing for coordinative control of movement. This developmental *feedback internalization* process has broad consequences for phonetic and phonological patterns (Tilsen, 2016). However, a major challenge in implementing a computational model of feedback internalization arises due to conceptual differences between competitive and coordinative control (Tilsen, 2013), and empirical data require that the model should be flexible enough to exhibit both competitive and coordinative regimes of articulatory control.

To address these challenges, a new phenomenological model was developed, in which selection of articulatory gestures is governed by quantized step potentials. These potentials are hypothesized to emerge from pairwise inhibitory coupling between neural ensembles encoding gestural planning systems, and in the aggregate the potentials form a quantized macroscopic potential. In prototypical competitive control, gestures occupy energy levels in a mutually exclusive fashion because of strong inhibitory coupling forces. In prototypical coordinative control, multiple gestures can occupy the same energy level, due to weak inhibitory coupling. The transition between these regimes is modeled as a decrease in inhibitory coupling that occurs in parallel with increasing reliance on internal sensory feedback.

Simulations of competitive-to-coordinative control transitions for VC and VCC syllables were conducted with reinforcement learning applied to parameters for inhibitory coupling strength and reliance on internal feedback. Parameters characterizing initial activations of gestural planning systems were yoked to inhibitory coupling strengths, phase coupling strengths were yoked to reliance on internal sensory feedback, and a bias for reduced inhibitory coupling was imposed in each learning iteration.

The model transitions from competitive to coordinative organization of a pair of gestures when inhibitory coupling falls below a threshold, which corresponds to a lower barrier between energy levels. The transition is prevented when sensory targets are relatively distal in time, as in slow/careful speech. An advantage of using quantized potentials to regulate gestural selection is that sets of co-selected articulatory gestures do not need to be explicitly specified in the model. Hence the model exhibits greater flexibility behavior: by returning to a regime with stronger inhibitory coupling the model can regress to competitive control.

References

Tilsen, S. (2013). A Dynamical Model of Hierarchical Selection and Coordination in Speech Planning. *PLoS One, 8*(4), e62800. doi:10.1371/journal.pone.0062800

Tilsen, S. (2016). Selection and coordination: The articulatory basis for the emergence of phonological structure. *Journal of Phonetics, 55*, 53-77. doi:10.1016/j.wocn.2015.11.005

DOES SPEECH PRODUCTION REQUIRE PRECISE MOTOR CONTROL?

Saeed Dabbaghchian, Olov Engwall
KTH Royal Institute of Technology, Stockholm, Sweden

It is hypothesized that the overall transformation from motor command to the acoustic output has a quantal relationship (Perkell, 1996). This hypothesis predicts that there are regions of stability where some variation in motor command does not affect the acoustic output significantly; and that stable regions are connected through transient regions where the sensitivity of the acoustic output is very high. The quantal relationship either appears in the relation between motor command and vocal tract configuration or vocal tract configuration and acoustic output. Perkell (1996) shows that the contact mechanism between the tongue and other structures that are used by the speaker keeps the constriction degree relatively constant despite of the change in the motor commands. This mechanism keeps the acoustic output unchanged despite of its sensitivity to constriction degree in the vocal tract. In this study, we use a detailed 3D biomechanical model of the vocal tract (see Dabbaghchian et al. (2016) for more details) developed in ArtiSynth (Lloyd et al., 2012) to investigate the quantal theory between the motor commands (i.e. muscle activation) and the acoustic output. Using this computational model, activation of tongue muscles are changed systematically to move the tongue from neutral vowel towards [i] or [ɑ] configuration and the acoustic response in *F1-F2* plane (first and second formant) was calculated. The results clearly indicate that the acoustic output has the highest sensitivity to changes in muscle activation when the muscle activations are small (or correspondingly when the vocal tract is in neutral configuration). As the activations increase, the sensitivity decreases sharply. This means that, production of [ɑ] and [i] vowels may not need precise motor commands. For central vowels, there are two hypothesis. The first hypothesis is that the production of central vowels (such as [ə]) may require more precise motor commands. However, based on the second hypothesis, central vowels can be produced without precise motor commands; and the large variations in *F1-F2* plane (in comparison with cardinal vowels) is addressed in vowel perception. This needs further investigation to see if the perception area (e.g. in *F1-F2*) of central vowels is larger than cardinal vowels.

References

Dabbaghchian, S., Arnela, M., Engwall, O., Guasch, O., Stavness, I., & Badin, P. (2016). Using a Biomechanical Model and Articulatory Data for the Numerical Production of Vowels. In *Interspeech 2016* (pp. 3569-3573). San Francisco, USA.

Lloyd, J. E., Stavness, I., & Fels, S. (2012). ArtiSynth: a fast interactive biomechanical modeling toolkit combining multibody and finite element simulation. In *Soft tissue biomechanical modeling for computer assisted surgery* (pp. 355-394). Springer Berlin Heidelberg.

Perkell, J. S. (1996). Properties of the tongue help to define vowel categories: hypotheses based on physiologically-oriented modeling. *Journal of Phonetics, 24* (1), 3-22.

DECOMPOSING VOCAL TRACT CONSTRICTIONS INTO ARTICULATOR CONTRIBUTIONS USING REAL-TIME MAGNETIC RESONANCE IMAGING

Tanner Sorensen, Asterios Toutios, Louis Goldstein, Shrikanth Narayanan
University of Southern California, Los Angeles, United States of America

Real-time magnetic resonance imaging provides information about the dynamic shaping of the vocal tract during speech production. We present a method for decomposing the formation and release of a constriction in the vocal tract into the contributions of individual articulators such as the jaw, tongue, lips, and velum. The contours of speech articulators were identified in the real-time magnetic resonance imaging videos and tracked automatically over the course of their motion. We quantified the formation and release of constrictions by measuring the distance between the opposing structures (i.e., upper and lower lips for [p], tongue and alveolar ridge for [t], tongue and hard palate for [i], tongue and soft palate for [k], tongue and rear pharyngeal wall for [a]). Our method decomposed change in constriction degree into the contributions of individual articulators. We estimated the forward kinematic map, a nonlinear function which maps a vocal tract shape to the corresponding constriction degrees. The jacobian of the forward kinematic map quantifies how a small change in vocal tract shape changes the constriction degrees. By parameterizing vocal tract shape as the linear combination of jaw, lip, and tongue components, the jacobian mapped jaw motion, lip motion, and tongue motion to the corresponding changes in constriction degree at the phonetic places of articulation. The change in a constriction degree due to the jaw, lips, or tongue is the contribution of that articulator to the total change in constriction degree. The proposed method allows vocal tract constrictions to be decomposed into the contributions of individual articulators using real-time magnetic resonance imaging.

References

Sorensen, T., Toutios, A., Töger, J., Goldstein, L., & Narayanan, S. S. (2017). Test-retest repeatability of articulatory strategies using real-time magnetic resonance imaging. In *Proc. Interspeech*, Stockholm, Sweden.

Toutios, A., & Narayanan, S. S. (2015). Factor analysis of vocal tract outlines derived from real-time magnetic resonance imaging data. In *International Congress of Phonetic Sciences (ICPhS)*, Glasgow, United Kingdom.

LEARNING TO TALK: RELATIONSHIPS BETWEEN SPEECH MOTOR CONTROL AND PHONOLOGICAL DEVELOPMENT

Carol Stoel-Gammon
University of Washington, Seattle, United States of America

Phonological development has both a biologically based component associated with the development of speech-motor skills needed for adult-like pronunciations, and a cognitive-linguistic component, including processes of memory and pattern recognition, associated with storage and retrieval of words (Stoel-Gammon, 2011). This presentation focuses on two issues that must be considered when examining relationships between the two components: (1) What is the child's *target form*? (2) What is the nature of *variability/stability* in the child's productions?

For the early stages of development the issue is: Do the consonant-vowel (CV) syllables of babble have specified phonetic targets? If so, what is the nature of those targets? It is well documented that the consonants of babble are not random, but are influenced by articulatory factors: Most are produced with the lips or front of the tongue, with full oral closure or with an open mouth posture, thus yielding a small set of stops, nasals, and glides. The question is: How much *motor control* is involved in these CV productions; i.e., does the baby target [b] vs. [m] in producing [baba] and [mama]? Beyond *motor control*, we know that CV babble provides *motor practice* that influences meaningful speech. This practice increases the control and precision of movements – the more automatic the movements, the easier to execute them in targeted words (Kent, 1982).

When the child moves to the production of words, the issue of targets is more straightforward. In order to recognize and produce words, children must have articulatory, phonological, and semantic information stored in their "mental lexicon." Word recognition involves the ability to extract and store auditory information; word production requires linking a stored form with articulatory details. Despite decades of discussion, there is little agreement about the number and nature of underlying representations (URs) in the adult mental lexicon and even less agreement about the URs of children. Of interest here is research suggesting that the specification of phonological targets is affected by the size of the child's lexicon: When the vocabulary is small, representations may be stored as single, unanalyzed units; on this view, a word is stored and retrieved not as a sequence of phonemes, but as a single unit with relatively little detail. As vocabulary size increases, feature- and phoneme-based URs emerge and productions become more adult-like. Studies examining the phonological patterns of early words and of developmental trajectories related to the stability and variability word productions will be presented.

References

Stoel-Gammon, C. (2011). Relationships between lexical and phonological development in young children. *Journal of Child Language, 38*, 1-34

Kent, R.D. (1992) The biology of phonological development. In C.A. Ferguson et al. (Eds.) *Phonological development: Models, research, implications.* Timonium, MD: York Press.

HOW DOES THE TONGUE LEARN TO SPEAK A LANGUAGE FLUENTLY? A CROSS SECTIONAL STUDY IN GERMAN CHILDREN

Aude Noiray[1,2], Dzhuma Abakarova[1], Elina Rubertus[1], Jan Ries[1]
[1]*University of Potsdam, Potsdam, Germany*
[2]*Haskins Laboratories, New Haven, United States of America*

In the domain of spoken language acquisition, a large body of empirical research has focused on coarticulation mechanism, which regards the binding of articulatory gestures for neighboring phonemes. Coarticulation is an important mechanism to investigate as it engages multiple speech articulators (e.g., the lips, the tongue) whose actions must be finely coordinated in time and in the space of the vocal tract to produce fluent phonetic output in the native language. However, up to date, the development of temporal and spatial organization of speech gestures, in particular that of the tongue remains poorly understood in young children. Our study addresses this limitation by presenting a quantitative cross-sectional investigation of lingual coarticulation in German children that expands from the preschool years to the beginning of second grade. Unlike previous studies, we investigated the articulatory mechanisms from which differences in coarticulation may originate. Adapting the technique of ultrasound imaging to child study (SOLLAR, Noiray et al. 2015), we recorded movements of the main tongue articulator for vowels and consonants production in a series of $C_1VC_2ə$ nonwords. We tested whether the organization of intra-syllabic coarticulatory patterns not only varies as a function of age but also depends on the articulatory demands imposed on the tongue for consecutive phonemes. To achieve these goals, we measured the coarticulation degree (CD) of the tongue body between C_1 and V using consonants (/b, d, g, z/) known in adults to vary in their coarticulatory flexibility (Fowler, 1994). Results from linear mixed effects models highlighted significant age differences in lingual coarticulation with preschoolers showing larger CD than adults. Adults displayed more fine-grained modulations of CD as a function of consonants' articulatory signature compared to children. When further examining the temporal unfolding of the coarticulatory process within the consonant, we found that the coarticulatory span decreased with age. Preschoolers exhibited a strong encroachment of the vowel with the consonant, suggesting an organization of lingual gestures that encompasses both phonemes. School-age children showed less vocalic influence over the tongue configuration within the consonant but did not yet match adults' patterns. Overall, results show that in the second school year, children do not fully control the spatial and temporal organization of lingual gestures for fluently coarticulating the phonemes of their native language. As the degree of lingual control may be tightly intertwined with children's experience with their native language, we are currently testing for effects of phonological development on coarticulatory patterns.

References

Fowler, C. A. (1994). Invariants, specifiers, cues: An investigation of locus equations as information for place of articulation. *Percept. Psychophys. 55*, 597-610.

Noiray, A., Ries, J., Tiede, M. (2015). Sonographic & Optical Linguo-Labial Articulation Recording system (SOLLAR). *Ultrafest VII*, Hong Kong, 2015.

THE ROLE OF TONGUE CONTROL MATURATION FOR V-TO-V COARTICULATION

Elina Rubertus[1], Aude Noiray[1,2]
[1]*University of Potsdam, Potsdam, Germany*
[2]*Haskins Laboratories, New Haven, United States of America*

Tongue movements for speech segments vary depending on their phonetic context. For adults, it has been shown that these coarticulatory effects do not only occur between adjacent segments but can span several segments in both the anticipatory and carryover direction. Moreover, especially the two directions of vowel-to-vowel (V-to-V) coarticulation are claimed to originate from different underlying processes: While articulatory planning is the driving force for anticipatory coarticulation, carryover effects mainly result from mechanical constraints and articulators' inertia (Recasens, 1987). Lingual V-to-V coarticulation has also been investigated in children to address speech motor control development. However, most acquisition studies have focused on anticipation only and were restricted to acoustic measures. With this study, we shed more light on the development of speech motor control and articulatory planning by comparing lingual V-to-V coarticulation in anticipatory and car-ryover directions across different ages using articulatory measures.

We recorded 69 German children (3y, 4y, 5y, & 7y) and adults using SOLLAR, a child-friendly recording platform (Noiray, Ries, & Tiede, 2015). The technique of ultrasound imaging allowed us to trace tongue positions directly instead of inferring them from the acoustic signal. We used a symmetrical stimulus structure ($\text{ə}C_1VC_2\text{ə}$) to test for influences of the me-dial vowel (/i/, /y/, /u/, /a/, /e/, /o/) on both schwas - the preceding one for anticipatory and the following one for carryover coarticulation. The intervocalic consonants varied in their coarticulatory resistance (/d/ $>$ /g/ $\geq$ /b/) (Recasens, Pallarès, & Fontdevila, 1997). Congruent with Recasens' (1987) view, we hypothesized that highly resistant intervocalic consonants decrease lingual V-to-V coarticulation more extensively in the carryover than in the anticipa-tory direction.

Our data indicate that all age groups exhibit substantial anticipatory as well as carryover V-to-V coarticulation. In both directions, coarticulation magnitude decreases with age. However, the decrease is stronger and more linear in anticipatory coarticulation. The intervocalic consonant plays a greater role in carryover coarticulation than in anticipatory coarticulation, but its influence on V-to-V coarticulation is not uniform across cohorts. These first results provide more evidence for the two coarticulation directions to be guided by different underlying processes. Implications for the complex development of speech production and tongue control during childhood will be discussed in further detail.

References

Noiray, A., Ries, J., & Tiede, M. (2015). Sonographic & Optical Linguo-Labial Articulation Recording system (SOLLAR). *Ultrafest VII*, Hong Kong.

Recasens, D. (1987). An acoustic analysis of V-to-C and V-to-V coarticulatory effects in Catalan and Spanish VCV sequences. *Journal of Phonetics, 15*, 299-312.

Recasens, D., Pallarès, M. D., & Fontdevila, J. (1997). A model of lingual coarticulation based on articulatory constraints. *Journal of the Acoustical Society of America, 102*, 544-561.

NEURAL SIGNATURES OF CHILDHOOD STUTTERING PERSISTENCE AND RECOVERY

Soo-Eun Chang, Ho Ming Chow
University of Michigan, Ann Arbor, United States of America

Stuttering affects about 5% of all preschool children, but approximately 80% will recover naturally within a few years of stuttering onset. The pathophysiology and neurodevelopmental trajectories associated with persistence and recovery of stuttering are still largely unknown. To date, very few studies have examined structural and functional brain differences in children who stutter (CWS), and even fewer (if any) studies have examined neurodevelopmental trajectories that differentiate children who have persistent stuttering from those who naturally recover from stuttering. Elucidating these neurodevelopmental differences may lead to better prognostic markers of childhood persistent stuttering, and may inform intervention that modulates neural targets to improve speech fluency.

In this presentation, I will discuss results derived from the first longitudinal study of childhood stuttering, in which multimodal neuroimaging datasets (an average of 3 time points/child) from more than 100 children were acquired. We conducted studies examining structural (diffusion tensor imaging; DTI) and functional (resting state fMRI) imaging data to examine group- and age-related differences characterizing stuttering, and to ask whether brain connectivity patterns could predict persistence and recovery in later years. In terms of white matter development, relative to controls, CWS exhibited decreased fractional anisotropy (FA; a measure derived from DTI that reflects white matter integrity) in the left arcuate fasciculus, underlying the inferior parietal and posterior temporal areas and the mid body of corpus callosum. Further, a reduction in FA growth rate (i.e., decreased FA with age) was present in persistent children relative to controls, which was not evident in recovered children. We also conducted a whole-brain network analysis based on resting state fMRI data, in order to examine large-scale intra- and inter-network connectivity changes associated with stuttering. Moreover, we examined whether whole brain connectivity patterns observed at the initial year of scanning could predict persistent stuttering in later years.

Results showed that risk of stuttering, as well as stuttering persistence, was associated with aberrant network connectivity involving the default mode network and its connectivity with attention, somatomotor, and frontoparietal networks. Childhood stuttering risk and persistence was linked to developmental alterations in the balance of integration and segregation of large-scale neural networks that support proficient task performance, including fluent speech motor control. In sum, these findings of anomalous structural and functional brain connectivity in CWS provide novel insights into the possible neural mechanisms of childhood stuttering onset, persistence and recovery.

References

Chow, H., & Chang, S-E. (2017). White matter developmental trajectories associated with persistence and recovery of childhood stuttering. *Human Brain Mapping*. doi: 10.1002/hbm.23590.

Chang, S-E., Angstadt, M., Chow, H., Etchell, A.C., Garnett, E. O., Choo, A., Kessler, D., Welsh, R., & Sripada, C. (2017). Anomalous network architecture of the resting brain in children who stutter. *Journal of Fluency Disorders*. doi: 10.1016/j.jfludis.2017.01.002.

BROCA'S REGION IN PERSISTENT DEVELOPMENTAL STUTTERING WHERE STOP MEETS GO

Nicole E. Neef[1,2], Alfred Anwander[1], Christoph Bütfering[2], Angela D. Friederici[1], Walter Paulus[2], Martin Sommer[2]
[1]*Max Planck Institute for Human Cognitive and Brain Sciences, Leipzig, Germany*
[2]*Georg August University, Göttingen, Germany*

In 1861, Broca assigned a region in the human left frontal lobe to a designated function - articulated language. More than 150 years later it is still an open question how the human brain generates well-articulated fluent speech. Developmental stuttering is a speech fluency disorder that occurs without obvious reason during childhood and persists in one percent of the adult population. Previous neuroimaging studies frequently linked stuttering with Broca's region, thereby suggesting a causal, maladaptive, or compensatory role.

Here, we test the specific role of a part of Broca's region in stuttering; the left inferior frontal gyrus pars opercularis (IFGop). Functional MRI was used during two vocal motor imagery tasks. Adults who stutter (AWS) showed a reduced activity in the posterior-dorsal part of the left IFGop, a subarea that is involved in speech production, including phonological word processing, pitch processing, sequencing, and motor planning. In addition, AWS revealed a lack of task-related correlation between activity in left IFGop and activity in the left inferior parietal lobule, a link that likely represents a cortical network involved in implicit motor planning and motor control processes. According to motor control theory, this link implies the generation of predictive internal forward and inverse models that link motor intentions, motor commands, and sensory effects. Thus, our finding is in consonance with theories that postulate defective internal models as a possible source of stuttering. Furthermore, right IFGop activity was related to the inhibition of vocal responses, which leads to the proposal of a potential neuropathological principle: Stuttering might be caused by an overly active global response suppression mechanism mediated via the subthalamic nucleus-right IFG-basal ganglia hyperdirect pathway. Its amplified involvement might increase the system's tendency to globally inhibit motor responses. Inhibition via the hyperdirect pathway is unspecific and induces a global reduction of the thalamo-cortical drive. If this global inhibition is too strong or imbalanced, as proposed in stuttering, the stopping of an ongoing speech motor program or the selection of a succeeding speech motor program might fail, which might provoke sound prolongations, sound repetitions, and speech blocks. And indeed, the analysis of white matter structures that link the right IFGop into motor control networks revealed that the structural connectivity within these networks scales with stuttering severity. Altogether, our findings suggest an imbalanced organisation of IFGop networks in stuttering that adversely engage the left hemisphere in speech flow and the right hemisphere in the stopping of speech.

Further reading

Cai, W., Ryali, S., Chen, T., Li, C.-S. R., & Menon, V. (2014). *Journal of Neuroscience, 34*(44), 14652-14667.

Neef, N. E., Bütfering, C., Anwander, A., Friederici, A. D., Paulus, W., & Sommer, M. (2016). *NeuroImage, 142*, 628-644.

IMPROVING SPEECH FLUENCY IN ADULTS WHO STUTTER USING TRANSCRANIAL DIRECT CURRENT STIMULATION

Jennifer Chesters, Riikka Möttönen, Kate E. Watkins
University of Oxford, Oxford, UK

Non-invasive brain stimulation used as an adjunct to speech therapy has the potential to enhance fluency outcomes for adults who stutter (Chesters, Watkins, & Mottonen, 2017). Temporary fluency can be induced using an external cues (such as a metronome beat or second voice) but stuttering typically returns as soon as the inducers are withdrawn. We predicted that applying anodal transcranial direct current stimulation (tDCS) to the left inferior frontal cortex during speech production with temporary fluency inducers could extend fluency improvements beyond the intervention period. We chose to stimulate the left inferior frontal cortex because brain imaging studies have shown that developmental stuttering is associated with abnormalities in regions involved in speech planning and execution (Watkins, Smith, Davis, & Howell, 2008).

Thirty right-handed men aged 18-50 years with at least a moderate stutter were enrolled in a randomised sham-controlled trial. In the tDCS group, 1-mA direct current was applied for 20 minutes per day with the anode over the left inferior frontal cortex and the cathode over the right supraorbital ridge. In the sham group, the same electrode placement was used but the current was turned on only briefly to elicit the sensations that occur during tDCS. Both groups completed tasks fluently using choral reading and metronome-timed speech for 20 minutes on five consecutive days. Speech disfluency during reading and conversation was assessed at baseline, before and after the stimulation on each day of the five day intervention, one week and five weeks later.

As predicted, when the fluency-enhancing tasks were completed alone (sham) disfluency at the end of the five-day intervention was unchanged from baseline levels. When tDCS was combined with the fluency-enhancing tasks, fluency improvements were seen across the five-day intervention period and were significantly larger compared with the sham group. Most importantly, fluency was significantly improved for the tDCS group relative to the sham group one week after the intervention and this improvement was maintained for reading but not for conversation at five weeks. Scores on the Stuttering Severity Inventory (version 4) showed significantly greater reduction for the tDCS group relative to the sham group one week and five weeks after the intervention.

These results provide the first indication of the clinical potential of tDCS for increasing fluency outcomes in developmental stuttering. MRI measures were acquired before and after the intervention that may further our understanding of the mechanisms underlying the fluency improvements in the tDCS group.

References

Chesters, J., Watkins, K. E., & Mottonen, R. (2017). Investigating the feasibility of using transcranial direct current stimulation to enhance fluency in people who stutter. *Brain and Language, 164*, 68-76.

Watkins, K. E., Smith, S. M., Davis, S., & Howell, P. (2008). Structural and functional abnormalities of the motor system in developmental stuttering. *Brain, 131*(Pt 1), 50- 59.

MOTOR SPEECH DISORDER-NOT OTHERWISE SPECIFIED (MSD-NOS): PREVALENCE AND PHENOTYPE

Lawrence D. Shriberg
University of Wisconsin-Madison, Madison, United States of America

Background In addition to childhood dysarthria and childhood apraxia of speech, there are literature descriptions of a third pediatric speech sound disorder (SSD) proposed to be associated with a delay in motor speech development (Shriberg, Strand, & Mabie, 2017). A question is whether this putative clinical entity is the lower tail of the distribution of motor speech development or whether it is a subclinical or clinical type of motor speech disorder. Support for the latter two perspectives can be marshalled from research findings on the prevalence and characteristics of the behavioral phenotype. We have used the term Motor Speech Disorder-Not Otherwise Specified (MSD-NOS) as a placeholder for research on this potential clinical entity. The present report used new measures and a large multisite database of children ascertained for Speech Delay (SD) to estimate of the prevalence of MSD-NOS in SSD and to characterize the phenotype.

Method
Participants We analyzed continuous speech samples from six study groups recruited for SD of unknown origin over several decades. Sample sizes ranged from 21-188 participants, totaling 415 children and adolescents from 3-17 years of age assessed in cross-sectional and longitudinal designs.
Measures Participants recruited for SD at original assessment were classified into five mutually exclusive motor speech classifications: No Motor Speech Disorder (No MSD), Motor Speech Disorder-Not Otherwise Specified (MSD-NOS), Motor Speech Disorder-Dysarthria (MSD-DYS), Motor Speech Disorder-Childhood Apraxia of Speech (MSD-CAS), and Motor Speech Disorder-Dysarthria & Childhood Apraxia of Speech (MSD-DYS & CAS). Measures to identify the four motor speech disorders were: the Precision-Stability Index (for MSD-NOS), the Dysarthria Index (for MSD-DYS and MSD-DYS & CAS), and the Pause Marker (for MSD-CAS and MSD-DYS & CAS; Shriberg, Strand, & Mabie, 2017). The perceptual-acoustic measures were developed by operationalizing and standardizing signs of speech, prosody, and voice deficits in adult neurogenic disorders (Duffy, 2013) and signs developed in SSD research in children with idiopathic SD and SD in the context of complex neurodevelopmental disorders. Data reduction methods included narrow phonetic transcription, prosody-voice coding, and acoustic analyses.

Results & discussion We summarize prevalence and behavioral phenotype findings for MSD-NOS in comparison to prevalence findings for No MSD, MSD-DYS, MSD-CAS, and MSD-DYS & CAS. Findings are interpreted to support MSD-NOS as a substantially prevalent clinical entity characterized by spatiotemporal differences in articulatory, prosodic, and vocal precision and stability.

References
Duffy, J. R. (2013). Motor speech disorders: *Substrates, differential diagnosis, and management* (3rd ed.). St. Louis, MO: Mosby.

Shriberg, L. D., Strand, E. A., & Mabie, H. L. (2017). *Prevalence of speech and motor speech disorders in idiopathic speech delay and in complex neurodevelopmental disorders.* Manuscript submitted for publication.

THE AMPI CLASSIFICATION OF THE DYSARTHRIAS

Frank R. Boutsen, Justin D. Dvorak, Zachary Barrett
University of Oklahoma Health Sciences Center, Oklahoma City, United States of America

Introduction The dysarthrias traditionally have been classified using the auditory-perceptual method along what has come to be known as the Mayo Classification system. Recently, investigations have used prosodic (rhythm) matrices to see if they differentiate the dysarthrias. The results are mixed. Liss et al., (2009) found that a combination of these matrices can reasonably well distinguish the dysarthrias from each other and from normally healthy speakers. Lowit (2014) however reported that none of the metrices could do this despite clear perceptual differences. The present investigation evaluates the use of the newly developed Acoustic Multidimensional Prosody Index (AMPI) when classifying the dysarthrias.

Methods Audio recordings of the Grandfather Passage were taken from 57 participants (10 controls, 10 with ataxic dysarthria, 10 with flaccid, 7 with hypokinetic, 10 with hyperkinetic, and 10 with spastic) and then digitized at 44,100 Hz with 16-bit resolution. Vowel nuclei in a phonemically diverse, contiguous sample of the passage were identified using SpeechMark, followed by manual confirmation of onset/offset times and extraction of acoustic parameters in custom MATLAB software. Measurements for each vowel included duration (ms), intensity (dB), and pitch (Hz). For each pair of groups, prosody indices were computed at the patient level following Dvorak, Boutsen, & Ding (submitted), with weight vectors optimized for AUC via grid search. Those weights which maximally separated each pair of groups were extracted, and the corresponding AUC was assessed for significant predictive ability via permutation test with 1,000 iterations.

Results & Discussion Overall AUC of each binary classification was 80%, and ranged from 68% (flaccid vs. spastic; $p=0.1820$) to 93% (control vs. hypokinetic; $p<0.0001$). Of the 15 group comparisons, 11 were significant at $\alpha=0.05$. Adjusting for multiple testing, the prosody index distinguished the controls from spastic, hypokinetic, and ataxic dysarthria with duration driving the difference in spastic and ataxic dysarthria vs. normal comparisons and pitch driving the difference in hypokinetic dysarthria vs. normal controls. The trend towards a difference (.09) between flaccid dysarthria and normal control was driven by pitch. Subtypes also were shown to differ from one another ((7/10) contrasts, or 70%).

References

Liss, J. M., White, L., Mattys, S. L., Lansford, K., Spitzer, S., Lotto, A. J., & Caviness, J. N. (2009). Quantifying speech rhythm deficits in the dysarthrias. *Journal of Speech, Language, and Hearing Research, 52*, 1334-1352.

Lowit, A . (2014) Quantification of rhythm problems in disordered speech: a re-evaluation. *Phil. Trans. R. Soc. B 369*, 20130404

FACIAL KINEMATICS IN THE ASSESSMENT AND DIAGNOSIS OF BULBAR ALS

Yana Yunusova[1,2,3], Andrea Bandini[3], Lorne Zinman[2], Jordan R. Green[4]
[1]*University of Toronto, Toronto, Canada*
[2]*Sunnybrook Research Institute, Toronto, Canada*
[3]*UHN: Toronto Rehabilitation Institute, Toronto, Canada*
[4]*MGH Institute of Health Professions, Boston, USA*

Amyotrophic later sclerosis (ALS) is a fast-progressing motor neuron disease affecting orofacial (bulbar) musculature and resulting in an inability to speak and swallow. Although only approximately 30% of patients present with signs and symptoms in the bulbar musculature, nearly 80% of the remaining patients develop bulbar signs as disease progresses. Diagnosis of ALS is often delayed (Turner et al., 2010). This delay leads to the inability of patients to access necessary life-prolonging health care services and experimental drug trials. There is a significant need to expedite the detection of bulbar signs and diagnosis of ALS.

We turn to facial kinematics to aid detection of bulbar changes in ALS. Our early work suggested that jaw and lip kinematics changed with disease progression. An analysis of a large group of patients with ALS reported sensitivity of facial kinematic measures to both speaking rate and speech intelligibility decline (Rong et al., 2015). The goals of our recent work are to (1) expand our previous analyses in order to classify symptomatic and pre-symptomatic patients; (2) understand changes in facial kinematics within a patient as disease progresses; and (3) validate novel technology as means of clinical assessment of facial kinematics.

In this presentation, we will report on a series of studies performed on a large number of patients diagnosed with ALS (N=66), who participated in a longitudinal study of bulbar dysfunction (average months in study = 35.1, SD=25.3). All participants underwent a neurological evaluation; functional status assessment with ALS-Functional Rating Scale - Revised (ALSFRS-R); bulbar function testing with measures of speaking rate and speech intelligibility. Facial movement measurements (e.g., speed, distance, acceleration, jerk, surface) during a sentence task were obtained by using Optotrak Certus (NDI, Canada), Wave Speech Research System (NDI, Canada), and Intel®RealSense™ 3D camera. Analyses methods included classification approaches, graphical methods, and inferential statistics.

The results revealed the ability of kinematic measures to detect bulbar symptoms with 85% accuracy. Cumulative path distance, peaks of acceleration, speed, and jerk of the lower lip and jaw contributed the most to classification results, with the overall tendency for slower and smoother movements with disease progression. Changes overtime within a speaker were in the predicted direction. Kinematic data obtained with 3D camera was able to detect the same path in disease course as more expensive research grade methods.

References

Turner, M. R., Scaber, J., Goodfellow, J. A., Lord, M. E., Marsden, R., & Talbot, K. (2010). The diagnostic pathway and prognosis in bulbar-onset amyotrophic lateral sclerosis. *Journal of the neurological sciences, 294*(1), 81-85.

Rong, P., Yunusova, Y., Wang, J., & Green, J. R. (2015). Predicting early bulbar decline in amyotrophic lateral sclerosis: A speech subsystem approach. *Behavioural neurology*, 2015.

PREDICTING SPEECH INTELLIGIBILITY BASED ON TONGUE-JAW COUPLING IN PERSONS WITH ALS

Panying Rong[1], Jordan R. Green[2]
[1]*University of Kansas, Lawrence, United States of America*
[2]*MGH Institute of Health Professions, Boston, United States of America*

Speech is produced through the coordinated actions of the articulatory, phonatory, resonatory, and respiratory subsystems. Previous studies on amyotrophic lateral sclerosis (ALS), a progressive motor neuron disease, have shown that among all subsystems, impairment to the articulatory subsystem has the most severe effect on speech intelligibility (Rong et al., 2016). In particular, disease-related changes in tongue-jaw coupling were detected prior to the decline of speech intelligibility (P. Rong, Yunusova, Berry, Zinman, & Green, 2014) and might thus, serve as early predictors of bulbar motor involvement and the impending speech loss.

We quantitatively derived the pattern of spatial coupling between four semi-independent regions of the tongue (anterior tongue blade, posterior tongue blade, tongue body, tongue dorsum) and the jaw in various speech contexts (3 vowels: /ɔ/, /i/, /u/; 2 consonants: /d/, /k/) across 10 individuals with a range of severities of ALS from the XRMB database (Westbury, 1994). Based on the quantitative pattern, we assessed the effect of disease-related changes in tongue-jaw coupling on speech intelligibility.

To quantify the pattern of tongue-jaw coupling, the Parallel Factor Analysis (PARAFAC) was used, which derived two articulatory modes that characterized the pattern of independent tongue movement relative to the jaw (1^{st} mode) and the pattern of jaw contribution to composite tongue-jaw movement (2^{nd} mode). The extent of independent tongue movement and jaw contribution were (1) used to predict speech intelligibility of the 10 individuals with ALS using a nonlinear regression, and (2) compared with the corresponding coefficients derived from 10 age-matched healthy controls.

The results suggested that (1) independent tongue movement decreased progressively throughout the disease course, which led to a significant decline of speech intelligibility, whereas (2) the extent of jaw contribution showed non-monotonic changes during early versus late ALS. Specifically, the extent of jaw contribution was increased during early ALS, which partially compensated for the effect of reduced tongue movement on speech intelligibility. This compensation effect became unavailable during late ALS due to the decrease of tongue-jaw coupling, leading to precipitous declines of speech intelligibility. These findings could help us to understand the physiological basis of speech loss in ALS and anticipate the impending needs for assistive communication.

References

Rong, P., Yunusova, Y., Berry, J. D., Zinman, L., & Green, J. R. (2014). *Parameterization of articulatory pattern in ALS.* Paper presented at the Interspeech, Singapore.

Rong, P. Y., Yunusova, Y., Wang, J., Zinman, L., Pattee, G. L., Berry, J. D., . . . Green, J. R. (2016). Predicting Speech Intelligibility Decline in Amyotrophic Lateral Sclerosis Based on the Deterioration of Individual Speech Subsystems. *Plos One, 11*(5). doi:ARTN e0154971/10.1371/journal.pone.0154971

Westbury, J. R. (1994). On coordinate systems and the representation of articulatory movements. *J Acoust Soc Am, 95*(4), 2271-2273.

VOWEL FORMANT DISPERSION REFLECTS SEVERITY OF APRAXIA OF SPEECH AFTER STROKE IN ADULTS

Dirk-Bart den Ouden, Elena Galkina, Alexandra Basilakos, Julius Fridriksson
University of South Carolina, Columbia, United States of America

Introduction Apraxia of speech (AOS) is challenging to differentiate objectively from aphasic phonological planning deficits and from dysarthria. Deviations in consonantal voice-onset-time (VOT) have been associated with AOS, but studies of vowel acoustics have yielded conflicting results (Jacks et al., 2010), leading to a common assumption that vowels are spared in AOS. However, if the functional deficit of AOS is at the peripheral stages of speech production, there is no obvious reason why vowels should be left unaffected. Impaired articulatory control should lead to inconsistency, reflected in increased formant variance. We measured vowel formants produced by a large sample of stroke survivors, and assessed to what extent dispersion of F1, F2 and F3 was predictive of AOS severity, based on a rating schale that uses clinical observations to assess the gradient presence of AOS, aphasia, and dysarthria (Strand et al., 2014).

Method Picture-description samples were taken from 53 stroke survivors (20 female; mean age=61, range 32-82), grouped by impairment: unimpaired (11), primarily aphasia (20), aphasia with AOS (12), primarily AOS (2), aphasia with dysarthria (2), aphasia with AOS and dysarthria (6). We identified 7171 vowel tokens bearing main stress in open-class words, extracted the first three formants, and calculated their standard deviations. We measured VOTs for voiced and unvoiced consonants, with standard deviations, and estimated vowel space based on the first two formants of three vowels, /ɑ, i, u/. Multiple correlations (Spearman's *rho*) were used to assess relations between variables, Stepwise Linear Discriminant Analyses to predict group membership, and ordinal regression to predict AOS severity.

Results AOS severity correlated with *Voiced-VOT dispersion, mean F1* and *F1 dispersion*. A binary classification between speakers with and without AOS was achieved with 77.4% success by entering *Voiced-VOT dispersion, mean F1*, and *F2 dispersion*. The four acoustic factors that contributed significantly ($p<.05$) to the prediction of AOS severity were *F2 dispersion, mean Voiceless VOT, Voiceless VOT dispersion*, and *Voiced VOT dispersion*.

Discussion & Conclusion Presence and severity of AOS were most consistently predicted by F2 dispersion and (voiced) VOT dispersion. These phonetic-acoustic measures do not correlate with aphasia severity or dysarthria, and can be obtained objectively, reducing the subjectivity of AOS diagnosis. In addition, these results suggest that the apraxic deficit is across-the-board, not sensitive to phonological categories.

References

Jacks, A., Mathes, K., & Marquardt, T. (2010). Vowel acoustics in adults with apraxia of speech. *Journal of Speech, Language, and Hearing Research, 53*(1), 61-74.

Strand, E. A., Duffy, J. R., Clark, H. M., & Josephs, K. (2014). The Apraxia of Speech Rating Scale: a tool for diagnosis and description of apraxia of speech. *Journal of Communication Disorders, 51*, 43-50.

WINDOWS ON FLUENT AND DYSFLUENT SPEECH PRODUCTION

Martin Sommer[1], Annika Primaßin[1], Nicole Neef[1,2]
[1]*University of Göttingen, Göttingen, Germany*
[2]*Max Planck Institute for Human Cognitive and Brain Sciences, Leipzig, Germany*

Fluent speech is a fundamental prerequisite for social interaction in humans and a challenging task for the brain. Not all children acquire fluent speech, and stuttering persists after puberty in about 1 percent of adults. In recent years several windows have opened that allow a better view on kinematics, underlying brain structure and neurophysiology of fluent and dysfluent speech.

One window on kinematics is made possible by real-time very fast structural MR imaging observing the articulators while speaking. This gives a substantiated insight into dysfluent speech patterns and will allow answering the question whether fluent speech units in adults who stutter are normally articulated. With regard to brain structure, MR imaging of water diffusion in the brain has allowed uncovering a reduced integrity of white matter tracts, mostly in left inferior frontal areas. This has opened a whole new window to view stuttering as a speech area disconnection syndrome. It also provides a substrate for the known heritability of stuttering on which studies on whole genome sequencing shed further light. Clinical cases on cerebellar lesions that modulate speech fluency further substantiate the structural basis of fluent and dysfluent speech production.

Anticipation of dysfluent speech events is another aspect of stuttering. It is the premonitory awareness that something is about to go wrong with speech fluency. We have assessed this in detail by establishing a questionnaire derived from Tourette's syndrome to assess reliability and linguistic specificity of the premonitory awareness of dysfluencies in adults who stutter.

Another aspect is the sensitivity of stuttering to external pacing, which is reminiscent of the freezing of gait, a clinical phenomenon in Parkinson's disease. Here, electroencephalography allows dissociating brain areas responsible for self-triggered speech in contrast to externally triggered speech, thereby substantiating circuits likely involved in the pathogenesis of speech dysfluencies.

A window on the pathophysiology of stuttering has been opened by electrophysiological techniques such as transcranial magnetic stimulation and electroencephalography. These techniques allow investigating speech preparation with a very high time resolution, in order to assess speech preparatory processes of motor areas representing articulatory muscles. This allowed discovering an imperfect speech preparation in adults who stutter as a pathophysiological basis of dysfluent speech. Transcranial magnetic stimulation also gives a window on the hemispheric distribution of speech and non-speech related activations of articulatory muscles and as well as of hand muscles.

As a synopsis, we will illustrate how these windows allow views on fluent and dysfluent speech thereby refining and detailing current models of speech production.

THE VENTRAL SENSORIMOTOR CORTEX IN SPEECH ARTICULATION

Edward Chang
University of California, San Francisco, United States of America

The ventral sensorimotor cortex (vSMC) is important for the production of fluent speech. Injuries to this brain region cause dysarthria and other speech difficulties. In this talk, I will present new results from our investigations on the representations of vocal tract during speaking. I will address three topics: 1) the encoding of supra-laryngeal kinematic parameters during vowel production, 2) the cortical control of the larynx with regard to voicing and prosodic modulation, and 3) the vSMC cortical representation of articulatory gestural features in natural, continuous speech. These results will then be discussed in the context of current models of speech motor control.

References

Bouchard, K.E., Mesgarani, N., Johnson, K., & Chang, E.F. (2013). Functional organization of human sensorimotor cortex for speech articulation. *Nature, 495*(7441), 327-332.

Conant, D., Bouchard, K.E., & Chang, E.F. (2014). Speech map in the human ventral sensory-motor cortex. *Current Opinion in Neurobiology, 24C*, 63-67.

IDENTIFYING FUNCTIONAL REGIONS OF INTEREST WITHIN THE SPEECH MOTOR CONTROL NEURAL NETWORK

Jason A. Tourville, Alfonso Nieto-Castanon, Frank H. Guenther
Boston University, Boston, United States of America

Neuroimaging has revealed a reliable core network of cortical regions that contribute to speech production (Guenther, 2016). However, relatively little progress has been made toward identifying functional "units" within this network, due, in part, to relatively small sample sizes and resulting low statistical power. Here we describe efforts to identify functionally homogenous regions of interest (fROIs) within the speech motor control network. fROIS were derived from pooled "mega-analysis" (Costafreda, 2009) of a large set of overt speech production functional magnetic resonance imaging (fMRI) data using hierarchical clustering.

Data were pooled over 13 fMRI experiments in which blood oxygen level dependent responses (BOLD) were measured while subjects overtly produced single and bi-syllabic words and pseudowords. In total, 180 individual datasets from 153 unique subjects were included in the analysis. Estimates of the contrast between BOLD responses for Speech and Baseline conditions (silently viewing letter or symbol strings) were derived for each subject, then mapped to an average template using FreeSurfer. The cortex was then parcellated using between-subjects hierarchical clustering that minimized within-ROI variability in the Speech - Baseline contrast (Seghier & Price, 2009). Population-level boundary distributions were constructed to find the optimal number and location of region boundaries and an estimate of the reliability of these boundaries. A representative whole-brain parcellation was obtained from these boundary probability maps using watershed-based segmentation with adaptive smoothing to maximize the joint likelihood of the parcellation boundaries. ROI-level statistics were computed to assess within- and between-region functional homogeneity and functional sub-networks within the speech motor control network. Boundary-level statistics were also computed to determine between-subject variability in boundary positions and their relationship with macroanatomical cortical landmarks.

The resulting functional cortical parcellation provides insights into functional "maps" and networks within the speech motor control network. More generally, the method has the potential to improve the detection of BOLD response effects by providing a better means of aligning data across subjects. This was explored by comparing the sensitivity of detecting group-level differences in the Speech - Baseline contrast between fluent controls and persons who stutter using typical vertex-wise comparisons, ROIs based on macro-anatomical landmarks, and our fROIs. [Supported by NIH grants R01 DC002852 and R01 DC007683.]

References

Costafreda, S. G. (2009). Pooling fMRI data: meta-analysis, mega-analysis and multi-center studies. *Frontiers in Neuroinformatics, 3*(33). https://doi.org/10.3389/neuro.11.033.2009

Guenther, F.H. (2016). *Neural control of speech.* Cambridge, MA: MIT Press.

Seghier, M.L. & Price, C.J. (2009). Dissociating functional brain networks by decoding the between-subject variability. *NeuroImage, 45*(2), 349-359. https://dx.doi.org/10.1016/j.neuroimage.2008.12.017

ENHANCING SPEECH MOTOR LEARNING WITH NONINVASIVE BRAIN STIMULATION

Adam Buchwald, Mara Steinberg Lowe, Holly Calhoun, Rebecca Wellner, Stacey Rimikis
New York University, New York, United States of America

Introduction Research has revealed enhancement of non-speech motor learning from combining behavioral motor learning tasks with noninvasive neuromodulation provided by transcranial direct current stimulation (tDCS; Brunoni et al., 2012). While tDCS has been studied as a tool to facilitate word production in aphasia and unimpaired speakers (Monti et al., 2013), it has not been used to facilitate speech motor learning. We report a study successfully using tDCS as an adjunct to a speech motor learning protocol to enhance speech motor learning.

Methods 80 participants with no history of impairment completed a two-day protocol.
Speech Motor learning task L1 English participants produced nonwords with 8 different non-native onset clusters (e.g., /fm/, /gd/). Following brief pre-practice with detailed feedback, participants performed 18 minutes of structured practice producing nonwords with auditory and orthographic presentation (e.g., audio: /fmiku/, written: FMEEKOO). Short-term retention (R1; 30 mins after practice) and long-term retention (R2; 2 days after practice) tested both trained and novel nonwords.
Neuromodulation For active tDCS, a 1x1 Soterix battery-driven current stimulator delivered 20 minutes of 1mA current. The anode was placed over the left motor cortex (C3) and the cathode over the right supraorbital area. Sham tDCS used the same the electrode montage with current ramped up to 1mA over 30 seconds and then immediately decreased. Stimulation was administered before or during the practice session, with participants randomly assigned to one of four groups based on timing (before-during) and type (active-sham) of stimulation.

Results The primary outcome measures were change in word and cluster accuracy (assessed with perception and acoustics) between the first half of practice (P1) and: P2 (second half of practice); R1; and R2. Logistic regression mixed effects models were used to examine changes in whole-word accuracy and cluster accuracy. All groups improved on the task. For word accuracy, there was a significant 2-way interaction between stimulation type and session for all timepoints, with the active group improving more than the sham group (for cluster accuracy, this was only significant for R1). There were significant 3-way interactions for P1 compared to R1 for both measures where participants receiving active stimulation prior to the task improved more (10%) than other groups (all <4%).

Discussion This research represents a novel approach to enhancing speech motor learning. Our data suggest that non-invasive neuromodulation, prior to a task, can facilitate learning of novel sound structure sequences. Possible extensions to impaired speakers will be discussed.

References

Giacobbe, V. et al. (2013). Transcranial direct current stimulation (tDCS) and robotic practice in chronic stroke: The dimension of timing. *NeuroRehabilitation, 33*(1), 49-56.

Monti, A. et al. (2013). Transcranial direct current stimulation (tDCS) and language. Journal of Neurology, *Neurosurgery & Psychiatry, 84*(8), 832-842.

OUTCOMES OF SPEECH TRAINING EXAMINED WITH MRI OF THE BRAIN AND VOCAL TRACT

Carolyn McGettigan[1,2], Daniel Carey[1,3], Matthieu Ruthven[4], Clare Lally[1], Bronwen G. Evans[2], Patti Adank[2], Marc E. Miquel[1,4]
[1] *University of London, London, United Kingdom*
[2] *University College London, London, United Kingdom*
[3] *Trinity College Dublin, Dublin, Ireland*
[4] *Barts Health NHS Trust, London, United Kingdom*

It is a well-documented phenomenon that achieving native-like pronunciation in a second language becomes increasingly challenging in adulthood, with highly variable outcomes across listeners. Producing speech in a non-native language requires novel articulations to be accurately learned and deployed flexibly in a variety of word-level contexts. In a novel methodological approach to vocal learning, we have been using a speech training paradigm combined with real-time imaging of the vocal tract (rtMRI) and functional MRI of the brain (fMRI) to explore vocal imitation skill and chart the articulatory outcomes of learning L2 speech.

In the first such experiment (Carey et al., 2017), participants trained on imitating a native vowel and a similar non-native vowel that required lip rounding. Later, participants imitated these vowels and an untrained vowel pair during separate rtMRI and fMRI runs. Behavioural results showed that acoustic imitations of non-native vowels were less accurate than for the corresponding native vowels, with considerable individual differences in imitation success. Vocal tract rtMRI data revealed evidence of increased lip protrusion following training in lip rounding, which was related to variability in acoustic learning success. Univariate fMRI analyses revealed that regions including left inferior frontal gyrus were more active during sensorimotor transformation (ST) and production of non-native vowels, compared with native vowels; further, ST for non-native vowels activated somatomotor cortex bilaterally, compared to ST of native vowels. We further probed the data using multivariate Representational Similarity Analysis (RSA), which allowed us to directly map the acoustic and articulatory properties of the test vowels onto their underlying neural representations, using searchlight analyses. Using models constructed from participants' vocal tract images and from stimulus formant distances, we found that both types of description of the vowel categories were represented within somatomotor, temporal, cerebellar, and hippocampal neural activation patterns during ST. We thus provided the first evidence of widespread and robust cortical and subcortical neural representation of vocal tract and formant parameters during prearticulatory ST. Ongoing work on these data has additionally explored the effects of in-scanner learning on articulation, acoustics and neural activations, indicating significant effects of training on the profile of vocal learning over time.

This talk will introduce our novel methodology for the investigation of L2 speech learning, using examples from Carey et al. (2017) and follow-up experiments.

References

Carey, D., & McGettigan, C. (2017). Magnetic resonance imaging of the brain and vocal tract: applications to the study of speech production and language learning. *Neuropsychologia 98*, 201-211.

Carey, D., Miquel, M. E., Evans, B. G., Adank, P., & McGettigan, C. (2017). Vocal Tract Images Reveal Neural Representations of Sensorimotor Transformation During Speech Imitation. *Cerebral Cortex* doi: https://doi.org/10.1093/cercor/bhx056

A FRAMEWORK FOR INVESTIGATING SPEECH PRODUCTION AND ITS SENSORIMOTOR CONTROL

Vincent L. Gracco[1,2], Hiroki Ohashi[1]
[1]*Haskins Laboratories, New Haven, United States of America*
[2]*McGill University, Montréal Canada*

Current neural models and theories of speech production rely to a significant extent on empirical data obtained from functional magnetic resonance imaging (fMRI) studies. These models and the associated data are being used increasingly to understand and speculate on the neural mechanisms for speech production and, subsequently explain and design treatment approaches for speech disorders. However, for the most part, brain based models for speech production and its sensorimotor control are incomplete due to: 1) an almost exclusive focus on a single component of the neurovascular signal, the positive blood-oxygenation-leveldependent (BOLD) response (PBR) to the exclusion of the negative BOLD response (NBR), and, 2) limited consideration for network level interactions and their associated functions. With the application of more advanced statistical approaches to brain function it is becoming clear that behavior emerges from a complex interaction between distributed *functional networks* (FN). These interactions are dependent on a balance of activation (excitatory) and deactivation (inhibitory) mechanisms in task-positive and task-negative networks that often overlap within the same cortical region and are modulated by context or state-dependent conditions. The result is that network-level interactions represent a range of brain-based actions that initiate and control speech production, engage neuroplastic change for new and adaptive speech behaviors and maintain and tune overall network properties. Here we present data and analysis techniques for a number of speech production tasks that capture a broad and representative conceptualization of the neural mechanisms and processes for sensorimotor and cognitive control of speech production. The approach relies on a combination of traditional general linear modeling (GLM) to examine task-related activation and deactivation and spatial independent components analysis (sICA) to separate BOLD signal mixtures from each voxel into functional networks. We present fMRI data from three different speech production datasets in which speech is elicited under a variety of conditions including speech motor learning/adaptation. We highlight the ubiquitous presence of the NBR and task-negative networks, their variation with the PBR and task-positive networks and how combining sICA and GLM analyses can be used to inform on how different FN contribute to behavior. Overall, we suggest that in order to develop more explanatory, neurobiological models of speech production and its sensorimotor control, an integrated approach is needed focusing on functional networks, their interaction and relationship to behavior combined with explicit consideration of excitatory and inhibitory mechanisms.

Further reading

Fox M.D., Snyder A.Z., Vincent J.L., Corbetta M., Van Essen D.C., Raichle M.E. (2005) The human brain is intrinsically organized into dynamic, anticorrelated functional networks. *Proc Natl Acad Sci A, 102*, 9673-9678.

Shmuel A., Augath M., Oeltermann A., Logothetis N.K. (2006) Negative functional MRI response correlates with decreases in neuronal activity in monkey visual area V1. *Nat. Neurosci., 9*, 569-577.

QUANTITATIVE ASSESSMENT OF A NEUROCOMPUTATIONAL MODEL OF SPEECH PRODUCTION WITH NEUROIMAGING DATA

Frank H. Guenther[1], Ayoub Daliri[1,2], Alfonso Nieto-Castanon[1], Jason A. Tourville[1]
[1]*Boston University, Boston, United States of America*
[2]*Arizona State University, Tempe, United States of America*

Advances in neuroimaging techniques such as functional magnetic resonance imaging (fMRI) and electrocorticography (ECoG) over the past two decades have resulted in a greatly improved understanding of the neural mechanisms underlying human sensory, motor, and cognitive capabilities, leading to increasingly sophisticated neural models of these functions. Within the domain of speech, we have developed and refined a large-scale neurocomputational model, called the Directions into Velocities of Articulators (DIVA) model (Guenther, 2016; Guenther, Ghosh, & Tourville, 2006; Golfinopoulos, Tourville, & Guenther, 2010), which provides a unified mechanistic account of acoustic, kinematic, and neuroimaging data on speech. Functional neuroimaging has been a powerful means for evaluating and refining such models. To date, however, these evaluations have been almost exclusively qualitative. Quantitative evaluations have been hampered by the absence of a general computational framework for (i) generating predicted functional activation from a model that can be directly and quantitatively compared to empirical functional neuroimaging data, and (ii) testing between models to identify the model that best fits experimental data.

Here we present a general computational framework to overcome these issues, with a specific application to the DIVA model. Within this framework, the brain network responsible for a task is broken into a set of computational nodes, each of which is localized to an MNI stereotactic coordinate in the brain. Associated with each node is a computational load function that links the node's activity to a computation involving quantifiable measures from the task. The instantaneous neural activity at each location in the brain (e.g., each voxel of an fMRI image, or each electrode of an ECoG array) is then calculated by summing the contributions of all model nodes at that location, with each node treated as a Gaussian activity source centered at the node's location. The parameters of the Gaussians (i.e., spread and magnitude of activation) are optimized to produce the best fit to the functional data. Model comparisons are based on the overall fit level and number of free parameters using the Akaike Information Criterion (AIC).

This framework was used in conjunction with a large fMRI database of speech production studies (116 speakers) to illustrate the DIVA model's ability to provide a unified account for whole-brain activity patterns seen during speech production under normal and perturbed conditions. [Supported by NIH grants R01 DC002852, R01 DC007683.]

References

Guenther, F. H. (2016). *Neural control of speech.* Cambridge, MA: MIT Press.

Guenther, F. H., Ghosh, S. S., & Tourville, J. A. (2006). Neural modeling and imaging of the cortical interactions underlying syllable production. *Brain and Language, 96*, 80-301.

Golfinopoulos, E., Tourville, J. A., & Guenther, F. H. (2010). The integration of large-scale neural network modeling and functional brain imaging in speech motor control. *NeuroImage, 52*, 862-874.

A "PHONOLOGICAL MIND" IN OUR BRAINS? CLINICAL EVIDENCE.

Wolfram Ziegler, Katharina Lehner, Hanna Jakob, Ingrid Aichert
Ludwig-Maximilians-Universität, Munich, Germany

Background According to a widely propagated theory, the human mind is equipped with a universal stock of "core phonological knowledge" that is purely algebraic and autonomous from the perception/production channels of linguistic communication (Berent, 2013). As a part of the human genetic endowment, this knowledge is claimed to be mediated by a dedicated neural network in the brain (Berent et al., 2014). The present study was based on the assumption that sound production impairments resulting from lesions to this network provide insight into the architecture of the "Phonological Mind". We examined (i) whether phonological impairment destroys the alleged "core phonological knowledge" postulated by Berent, and (ii) to what extent phonological impairment conforms to a model based on principles of intergestural coordination (Ziegler, Aichert & Staiger, 2016).

Methods
Subjects: 15 speakers with phonological impairment after left hemisphere stroke, 15 speakers with apraxia of speech after left hemisphere stroke.
Materials: Single word repetition task.
Procedure: Transcripts were analyzed for violations of phonological regularities and of markedness constraints, and for schwa-insertions in consonant clusters. Furthermore, item-response-analysis techniques were applied to map patterns of phonological errors onto an existing gesture-based model of sound production impairment, with the aim of testing substance-based explanations of phonological processing (Ziegler et al., 2016).

Results The results are incompatible with the assumption that aphasic phonological impairment interferes with the "core knowledge" and the abstract algebraic processing mechanisms postulated in the *Phonological Mind* theory.

Discussion Substance-based interpretations of the clinical data are proposed.

References
Berent, I. (2013). The phonological mind. *Trends Cogn Sci., 17*(7), 319-327.
Berent, I., Pan, H., Zhao, X., Epstein, J., Bennett, M. L., Deshpande, V., . . . Stern, E. (2014). Language Universals Engage Broca's Area. *PLoS ONE, 9*(4), e95155.
Ziegler, W., Aichert, I., & Staiger, A. (2016). When words don't come easily: A latent trait analysis of impaired speech motor planning in patients with apraxia of speech. *Journal of Phonetics.* doi:10.1016/j.wocn.2016.10.002

Poster session I

ACQUISITION OF SPEECH RHYTHM IN ONTOGENESIS AS AN ANALOGUE WINDOW ON DEVELOPMENT OF SPEECH RHYTHM IN PHYLOGENESIS AND EVOLUTION OF SPEECH FACULTY

Mikhail Ordin[1,2], Leona Polyanskaya[1]
[1] *Basque Center on Cognition, Brain and Language, San Sebastián, Spain*
[1] *Ikerbasque - Basque Foundation for Science, Bilbao, Spain*

In this project, we focused on the development of speech rhythm in first and second language acquisition, which is an analogue window on the development of speech rhythm in phylogenesis. We analyzed the rhythmic patterns in controlled utterances produced by learners of English with rhythmically contrastive languages: German (exhibiting rhythmic patterns similar to those in English) and French (exhibiting rhythmic patterns contrastive to those in English). Also, we analyzed rhythmic patterns in the same utterances spoken by English children at different ages. We found that the rhythm becomes progressively more stressed-timed (typical of an adult English native speaker) as acquisition progresses, and the direction of the developmental trajectory did not depend on the native language of the learners in second language acquisition (rhythmically similar or contrastive to the target language), nor on the mode of acquisition (L1/L2). As adult German learners of English do not transfer the rhythmic patterns from the native to the target language, and instead have to develop the stress-timed rhythm of the target language anew, following the developmental trajectory of L1 acquirers, we conclude that regular, syllable-timed rhythm is a default rhythm that is easier to produce.

An AX-discrimination experiment shows that processing syllable-timed rhythm is also easier to process. The experiment was set up with native speakers of rhythmically contrastive languages - French and German - to understand whether the predominant rhythms in native languages affect the accuracy and the reaction time (RT) in detecting rhythmic change in acoustic stream. The analysis of sensitivity and bias measures was conducted for accuracy. For the RT, we performed a 2-way ANOVA with native language of the listeners as a between-subject factor and the rhythm type of the A stimuli in AX pairs as a within-subject factor. Also, we computed the Bayes factors for comparing the statistical models. The analyses revealed that all listeners performed better when the rhythm in the A stimulus was syllable-timed (French-like): it is easier to detect the change (or the lack of change) after tuning into a syllable-timed rhythm, irrespective of the native language of the listener. This indicated that the detection of rhythmic changes draws on general auditory mechanisms more than on linguistic experience. These general mechanisms could potentially be sought in coupling of environmental rhythms with neural oscillations, and it can also explain the universality of more regular rhythmic patterns in the world languages, and the default mode - both in perception and production - of syllable-timing.

References

Botha, R. (2006). On the Windows approach to language evolution. *Speech Communication 6*(2), 29-143.

Ordin, M., & Polyanskaya, L. (2015). Acquisition of speech rhythm in a second language by learners with rhythmically different native languages. *Journal of Acoustical Society of America, 138*(2), 533-545.

QUANTIFICATION OF COARTICULATION RESISTANCE IN GERMAN WITH ULTRASOUND

Dzhuma Abakarova[1], Khalil Iskarous[2], Aude Noiray[1,3]
[1]*University of Potsdam, Potsdam, Germany*
[2]*University of Southern California, Los Angeles, United States of America*
[2]*Haskins Laboratories, New Haven, United States of America*

A successful characterization of vocal tract control during speech needs to account for regular variability in the degree of coarticulatory overlap allowed by different speech segments. While some segments allow for large degree of articulatory overlap, others show high coarticulation resistance (CR) i.g. ability to resist influence from neighbors and retain control over articulators across contexts. Despite the importance of the CR phenomenon for theories and models of speech production, a unified quantitative measure of coarticulation resistance is not established yet. The most prominent description of CR, DAC scale (Recasens & Espinosa, 2009), has some limitations: first, it subjectively divides segments into several categories based exclusively on the degree of lingual coarticulation they exhibit across contexts. Second, this categorization is based on the measures of coarticulation that quantify only linear dependencies. Generally, a wide variety of experimental techniques and corresponding quantification methods make it difficult to directly compare CR estimates across studies.

Recently, Iskarous et al. (2013) suggested measuring CR with Mutual Information (MI), or the amount of information shared by a given segment with other segments across contexts. This is a non-parametric method that does not make assumptions about distribution but rather estimates it from data. The MI values have been shown by Iskarous et al (2013) to capture the CR effects of place and manner, as well as time differences in CR.

Our study extends the application of MI to quantifying coarticulation from ultrasound images. To investigate the effect of consonantal context on the amount of vowel-to-consonant coarticulation in German we compare two different methods of quantifying tongue shape: the position of the highest point on the tongue dorsum and the description of the whole tongue contour based on discrete Fourier transform (DFT). The findings suggest that the highest point of the tongue reflects the CR at the point of maximum constriction forngue consonants while DFT coefficients capture fine spatial and temporal distinctions in whole tongue CR. This suggest that MI measure allows for crossmethodological comparisons and generalizations of quantitative findings.

References

Iskarous, K., Mooshammer, T., Recasens, D., Hoole, Shadle, C., Saltzman, E, and Whalen, D.H. (2013). The Coarticulation/Invariance Scale: Mutual Information as a measure of coarticulation resistance, motor synergy, and articulatory invariance, *J. Acoust. Soc. Am. 134*, 1271-1293.

Recasens D., and Espinosa A. (2009). An articulatory investigation of lingual coarticulatory resistance and aggressiveness for consonants and vowels in Catalan, *J. Acoust. Soc. Am. 125*, 2288-2298.

DISSOCIATING INPUT- AND OUTPUT-RELATED REPRESENTATIONS OF SPEECH IN SYLLABLE REPETITION

Jason W. Bohland[1], Christopher J. Markiewicz[2]
[1]*Boston University, Boston, United States of America*
[2]*Stanford University, Stanford, United States of America*

Repetition tasks require speakers to map auditory representations onto learned speech items, maintain those items in working memory, interface that representation with the motor system, and articulate the target sounds. The neural mechanisms underlying these processes are tightly intertwined. For example, brain areas traditionally associated with speech planning and motor output are commonly activated in perception tasks not requiring overt or covert speech. Here we describe an fMRI study, which builds on our previous work (Markiewicz & Bohland, 2016), and attempts to disentangle cortical regions that contain predictive information about the speech sounds heard, remembered, or planned for production in different stages of a novel syllable repetition task.

Trials began with visual presentation of two CVC words or nonwords, each containing different vowels (/æ/, /ɪ/, or /ʌ/). Subjects were asked to remember both (non)words, which were then removed from the screen. Next, an acoustic stimulus corresponding to one of the two (non)words was presented. This was followed by a visual cue, instructing subjects to prepare to produce *either* the auditory stimulus (REPEAT trials) *or* the unheard visual stimulus (CHANGE trials). After a delay, a GO signal instructed subjects to overtly produce the planned (non)word. In this design, the spoken and heard stimuli differed in 50% of trials. 18 healthy adults participated, most of whom completed 288 repetition trials. Data were analyzed using univariate and multivariate approaches. Cortical searchlight multi-voxel pattern analysis was used to determine voxel clusters whose response patterns significantly predicted the vowel identify of the auditory stimulus or the vocal target at different temporal events in individual trials.

Results indicated that the left inferior frontal sulcus and surrounding regions, hypothesized to serve as a phonological output buffer (Bohland, Bullock, & Guenther, 2010), were activated during the input stage of the task and predicted the vowel heard. When subjects were given a "change" cue requiring them to update their motor output plan, this region was additionally engaged, and clusters of nearby voxels became predictive of the *new* vocal target. Another prominent cluster in the posterior superior temporal sulcus, on the other hand, maintained a representation of the heard syllable throughout the trial, suggesting a more direct role in processing and encoding auditory inputs rather than planning speech output. Additional contrasts of classification accuracies revealed differential involvement of multiple areas in representing words and nonwords. These results help clarify neural correlates of phonological encoding and provide new data to constrain neural models of speech.

References

Markiewicz, C.J., & Bohland, J.W. (2016). Mapping the cortical representation of speech sounds in a syllable repetition task. *NeuroImage, 141*(1), 174-190.

Bohland, J.W., Bullock, D., Guenther, F.H. (2010). Neural representations and mechanisms for the performance of simple speech sequences. *Journal of Cognitive Neuroscience, 22*(7), 1504-1529.

COORDINATIVE SPEECH STRUCTURES: SPATIOTEMPORAL TONGUE PART MOVEMENT PATTERNS IN AMERICAN ENGLISH /ɹ/.

Suzanne E. Boyce, Ashley Walton, Sarah Hamilton, Allison Joy Garbo, Neeraja Mahalingam, Douglas Mast, Michael Riley
University of Cincinnati

The act of speech production demands the ability to successfully coordinate the movements of multiple articulators while adapting to a wide range of conversational contexts. This coordination must be flexible but precise, while simultaneously controlling the spatiotemporal patterns of many articulatory components. It has been proposed that this control is made possible due to the individual muscles and joints forming a functional unit called a coordinate structure where the variability in their movements cooperate and compensate in a manner specific to the speech production task at hand (Kelso et al., 1985). Magnetic resonance imaging (MRI) studies of typical adults have revealed at least twenty-two different tongue shapes characteristic of perceptually-correct American English /ɹ/ sound (Tiede et al., 2004). Within those shapes, there are consistent observations of constrictions in pharyngeal and oral spaces. Thus, there may be invariant movement parameters for /ɹ/, described as relative tongue part displacement from one speech segment to the next. We aim to examine how the tongue blade, tongue dorsum, and tongue root move relative to each other during the production of /ɹ/. We compared productions of /ɑɹ/ in 30 typical American English adult speakers to productions from 10 children with residual /ɹ/ errors. Data were recorded using a Siemens Acuson X300 ultrasound machine with a head-stabilization unit. Tongue blade, dorsum, and root displacements were measured using a novel method that tracks the location of local brightness maxima from images representing /ɑ/ and /ɹ/ production. We identified ranges of relative blade, dorsum and root displacement associated with normal /ɑɹ/ production and found that misarticulating children showed significantly different displacement of tongue parts compared to adults.

This quantitative description of relative tongue part movements may capture an invariant parameter that correlates with clinical perceptual ratings of being “closer” and “farther” from an acceptable /ɹ/ sound, and thus may have clinical utility in treating speech production errors. Methods that provide objective, quantitative sensory feedback and focus on how tongue parts are functionally organized as a coordinative structure (rather than on the parts’ independent movements) may result in motor learning that is more robust to speech contexts outside the clinical setting. Our method for quantifying tongue displacements is also well suited for developing simplified, real-time, visual feedback about articulation, which may further enhance learning and thus intervention effectiveness.

References

Kelso, J.A.S., Tuller, B., Vatioksis-Bateson, E., & Fowler, C. A. (1985). Functionally specific articulatory cooperation following jaw perturbations during speech: Evidence for coordinative structures. *Journal of Experimental Psychology: Human Perception and Performance, 10*, 812-832.

Tiede, M. K., Boyce, S. E., Holland, C. K., & Choe, K. A. (2004). A new taxonomy of American English /r/ using MRI and ultrasound. *The Journal of the Acoustical Society of America, 115*(5), 2633- 2634.

MUSCLE ACTIVATION PATTERNS IN LABIAL STOP CONSONANTS

Thibault Cattelain[1,2], Maëva Garnier[1,2], Christophe Savariaux[1,2], Pascal Perrier[1,2]
[1] *CNRS, GIPSA-lab, Saint Martin d'Hères Cedex, France*
[1] *University Grenoble-Alpes, Grenoble, France*

Introduction The goal of this study is to describe precisely orofacial muscle activities during labial stop consonant production (/p/, /b/), as a function of the vowel context and level of articulatory effort. This work raises several methodological problems: First, the high muscle density of the lip region raises some concerns with the use of surface electromyography to investigate lip movements, since there may be some signal crosstalk between the different pairs of electrodes positioned on different lip muscles. Second, for the production of a labial stop consonant the lip movement during the closing and opening gestures is crucial for speech acoustics. We propose to decompose these gestures into several basic phases, lip closing, lip compression and lip opening, during which the precise muscle coordination will be investigated.

Material & methods Four adult subjects without speech disorders were recorded in laboratory conditions while producing the logatoms /lepa/, /leba/, /lepi/, /lebi/, /lepu/ and /lebu/ with 5 levels of articulatory effort. Twelve repetitions of each logatom were recorded for each level of articulatory effort. Three types of signals were simultaneously recorded : the audio signal (at 44.1kHz), high-speed images of the lips (at 200Hz) and the EMG signal of 6 different orofacial muscles recorded at 200Hz with bipolar surface electrodes. The seven pairs of surface electrodes were placed on the superior orbicularis oris (OOS), the inferior orbicularis oris (OOI), the zygomatic (ZYG), the buccinator (BUC), the depressor labii inferioris (DLI), the mentalis (MNT) and the anterior belly of the digastric (DIG).

The external and internal lip contours, extracted from the high speed images, were used to segment the speech gesture into four time intervals corresponding to 1- lip closing, 2- lip compression increase, 3- lip compression decrease and 4- lip re-opening. The activity of each muscle was analyzed during these 4 times intervals, considering the full-wave rectified EMG signal.

Results Preliminary results suggest that the mentalis plays a major role in the lip closing phase, while the orbicularis oris is crucial for holding the compression. Simultaneous signals observed of electrodes informing on agonist/antagonist muscles suggest the existence of co-activation patterns that will be further investigated. The amplitude of muscle activation depends on the level of effort in a speaker-dependent way.

Further reading

Lapatki, B. G., Stegeman, D. F., & Jonas, I. E. (2003). A surface EMG electrode for the simultaneous observation of multiple facial muscles. *Journal of neuroscience methods, 123*(2), 117-128.

McClean, M. D., & Tasko, S. M. (2003). Association of orofacial muscle activity and movement during changes in speech rate and intensity. *Journal of Speech, Language, and Hearing Research, 46*(6), 1387-1400.

ARTICULATORY MOTOR CORTEX CONTRIBUTES TO PREDICTIVE CODING DURING SPEECH PERCEPTION: A COMBINED TMS AND EEG STUDY

Tatsuya Daikoku[1,2], Riikka Möttönen[1]
[1]*University of Oxford, Oxford, United Kingdom*
[2]*Max Planck Institute for Human Cognitive and Brain Sciences, Leipzig, Germany*

According to a popular hypothesis the brain constantly generates probabilistic predictions of 'what' is going to happen in our environment. In the Bayesian framework, the brain generates multiple predictions that can have different probabilities. The lower the probability of the sensory input, the greater the prediction error and reaction to the stimulus. This type of 'predictive coding' has been proposed to be important for speech perception and comprehension, since speech signals are periodic and show strong regularities. The motor system plays a central role in predictive coding especially during controlling of actions. During speech production, the motor cortex controls the movements of different articulators and generates predictions of sensory consequences of these movements, which are sent to sensory systems and compared to incoming sensory inputs. It has been suggested that the articulatory motor cortex also generates predictions during speech perception. However, so far there is no direct evidence supporting this idea. The main aim of the current study was to investigate whether the articulatory motor cortex contributes to predictive coding during speech perception. Since neural rhythms have been proposed to provide a mechanism for sensory predictions, we focused on oscillatory responses (e.g., Arnal and Giraud, 2012, Trends in Cognitive Sciences). We reanalyzed a set of four experiments from a combined TMS and EEG study (Möttönen et al., 2013, Cerebral Cortex). We compared oscillatory responses to the sounds that were highly predictable (probability = 1.0) with responses to the same sounds when they were less predictable (probability = 0.1). First, we found that reduction of predictability of phonetic ('ba' and 'ga') and acoustic (duration and intensity) content of sounds increased oscillatory power delta (1-4 Hz) and theta (4-8 Hz) bands. Second, we found that TMSinduced disruption of the lip motor cortex further increased theta power to low-probability 'ba' sounds, but not for low-probability 'da' sounds. This suggests that predictions for lip-articulated sounds became weaker and less accurate when the lip motor cortex was disrupted. This effect was highly specific: disruption of hand motor cortex had no effect on theta responses to low-probability 'ba' sounds and the disruption of the lip motor cortex had no effect on theta responses to reductions in predictability of duration and intensity. Our findings provide evidence that the articulatory motor cortex generates articulator-specific predictions during speech perception and that theta oscillations provide a mechanism for this predictive coding.

References

Arnal, L.H., Giraud, A.L. (2012) Cortical oscillations and sensory predictions. *Trends in Cognitive Science, 16*, 390- 398.

Möttönen, R., Dutton, R., Watkins, K.E. (2013) Auditory-Motor processing of speech sounds. *Cereb Cortex, 23*, 1190 -1197.

INFLUENCE OF ALTERED AUDITORY FEEDBACK ON ORAL-NASAL BALANCE IN SPEAKERS OF BRAZILIAN PORTUGUESE

Gillian de Boer[1], Viviane Cristina de Castro Marino[2], Evelyn Spazzapan[2], Larissa Cristina Berti[2], Eliana Maria Gradim Fabbron[2], Tim Bressmann[1]
[1] *University of Toronto, Toronto, Canada*
[2] *Universidade Estadual Paulista, Marília, Brazil*

Background & Purpose Adaptation studies explore the role of auditory feedback on speech. The compensatory response is often in the direction opposite to the changed parameter, indicating that auditory feedback mechanisms help the speaker reach a constant motor target. A prior study built on this research paradigm by exploring the role of auditory feedback in the regulation of oral-nasal balance in speakers of English. It found increased nasal feedback led to a reduction of nasality. A decrease in nasal feedback led to a smaller and inconsistent increase in nasality. The present study repeated the methodology among speakers of Brazilian Portuguese, whose language includes phonological and phonetic vowel nasalization.

Method Twenty normal speakers of Brazilian Portuguese (10F,10M) wore a Nasometer headset and headphones. Oral-nasal balance was quantified as a nasalance score. Two additional microphones were attached to the Nasometer sound separation plate, one for the nasal sound signal and one for the oral sound signal. The signals from the additional oral and nasal microphones were fed to a digital multitrack recorder. The participants continuously repeated two oral-nasal stimuli sentences. The volume level from the nasal microphone was gradually altered to change the relative loudness of the nasal channel in the mix, so that the speakers heard themselves as more or less nasal.

Results A repeated measures ANOVA was run for the mean nasalance scores of both stimuli by gender across five nasal feedback level conditions (three control, one maximum and one minimum nasal feedback). There was a significant main effect of nasal feedback level ($F(4,72) = 32.10$, $p < 0.0001$) and stimuli ($F(1,17) = 48.98$, $p < 0.0001$). There were no other effects. The maximum condition produced the lowest mean nasalance scores and the minimum condition had scores higher than two of three control conditions.

Conclusions As with speakers of English, the increased nasal feedback led to a compensatory adjustment in the opposite direction, and a lack of nasal feedback led to a smaller and inconsistent increase in nasalance. This suggests that even in Brazilian Portuguese, a language with phonological and phonetic vowel nasalization, decreased nasality was not perceived as critically as increased nasality by the speakers.

References

De Boer, G., & Bressmann, T. (submitted). Influence of altered auditory feedback on oral-nasal balance in speech. *Speech Language and Hearing Research.*

Mitsuya, T., Samson, F., Ménard, L., & Munhall, K. G. (2013). Language dependent vowel representation in speech production. *The Journal of the Acoustical Society of America, 133*(5), 2993-3003.

MODULATION OF THE SPEECH MOTOR CONTROL NETWORK THROUGH HIGH-DEFINITION TRANSCRANIAL DIRECT CURRENT STIMULATION

Dirk-Bart den Ouden, Karim Johari, Keiko Bridwell, Caroline Hayden, Roozbeh Behroozmand
University of South Carolina, Columbia, United States of America

Introduction Neural interactions between sensorimotor mechanisms play a critical role in speech motor control, yet the precise nature of these interactions is still poorly understood. To study these sensorimotor interactions, we applied high-definition transcranial direct current stimulation (HD-tDCS) to modulate compensatory responses to pitch shifts in voice auditory feedback. This study was motivated by the prediction that a modulatory effect of neurostimulation would support the use of ventral motor cortex as a potential target region for neural stimulation therapy in speakers whose speech motor control is impaired, such as patients with Parkinson's Disease.

Methods Speech and electroencephalography (EEG) signals were recorded in three groups (total n=30; 10 per group), before and after HD-tDCS, while subjects maintained steady vowel vocalizations and received pitch perturbations in their auditory feedback. We recorded participants' vocalizations and measured the amplitude of pitch shifts in response to manipulated feedback. Two groups received either anodal or cathodal HD-tDCS at 2mA to the left ventral motor cortex while the third group received sham stimulation at 1mA to adjacent electrodes on occipital cortex.

Results Results indicated that the magnitude of compensatory vocal responses to downward pitch shifts was significantly reduced following both anodal and cathodal HD-tDCS, with stronger effects associated with cathodal HD-tDCS. However, no such effect was observed following sham (Group by Stimulation interaction: $F(2,27)=4.45$, $p=0.022$). In addition, event-related potential (ERP) amplitudes time-locked to the pitch perturbations were significantly increased following anodal and cathodal HD-tDCS, but no such modulation was observed following sham. Source estimation revealed that cathodal HD-tDCS resulted in increased neural activity in a bilateral network within the inferior parietal lobule (IPL), whereas for anodal HD-tDCS, this effect was limited to the left IPL.

Discussion & Conclusion These findings suggest that neurostimulation of left ventral motor cortex modulates sensorimotor mechanisms underlying speech motor control. We propose that this effect is associated with the increased contribution of feedforward motor mechanisms, leading to a reduced modulatory effect of feedback perturbations and greater vocal motor stability during speech. We also suggest that the IPL provides a neural interface for sensorimotor integration during speech motor control.

Further reading

Behroozmand, R., Korzyukov, O., Sattler, L., & Larson, C. R. (2012). Opposing and following vocal responses to pitch-shifted auditory feedback: Evidence for different mechanisms of voice pitch control. *Journal of the Acoustical Society of America, 132*(4), 2468-2477

Larson, C. R. (1998). Cross-modality influences in speech motor control: The use of pitch shifting for the study of F0 control. *Journal of Communication Disorders, 31*(6), 489-503.

Wolpert, D. M., Diedrichsen, J., & Flanagan, J. R. (2011). Principles of sensorimotor learning. *Nature Reviews Neuroscience, 12*(12), 739-751.

RELATIVE PERCEPTUAL WEIGHTING OF PROSODIC CUES: AN EYE-TRACKING INVESTIGATION

Justin D. Dvorak, Frank R. Boutsen
University of Oklahoma, Oklahoma City, United States of America

Introduction Prosody, defined as the planned modulation in pitch, intensity, and duration, is a critical side-stream component of the speech signal. Aside from signaling contextual or pragmatic information, prosody is also used by listeners to organize meaning and make stress judgements, and can thereby affect lexical access in languages such as English, which use stress to distinguish semantically unrelated homophones. Research into the relative importance of the abovementioned acoustic dimensions in stress perception has yielded conflicting results, as have studies aiming to elucidate the trading relationships that exist among these cues.

Methods The present study used eye tracking to assess decision confidence in word-stress judgements under fully-crossed combinations of prosodic cues differing by successive JNDs (just-noticeable differences) in each dimension. Twenty native English-speaking adult participants (6 males, 14 females; mean age 28.1 yrs) with no history of disorders in speech, hearing, language, or vision were asked to listen to a series of bisyllabic nonsense words in which one syllable received emphasis in one or more prosodic dimensions. They were then asked to indicate which syllable they perceived as stressed by clicking on the number '1' or '2'. Decision confidence was evaluated by estimating the proportion of gaze time spent on the screen region allocated to the number eventually selected. From these measures, trading curves were computed among dimensions by estimating the level-set contours in decision confidence.

Results & discussion The behavioral results of this study (response selections and response time) support the primacy of pitch and intensity over duration in stress perception. Both of these dimensions were significantly associated with increased response accuracy ($p=0.0030$ for intensity, $p<0.0001$ for pitch) and decreased response times ($p<0.0001$ for intensity, $p=0.0169$ for pitch). Higher relative pitch was also significantly associated with greater response confidence ($p=0.0395$), and was allocated the greatest weight (0.4855, vs. 0.2417 for duration and 0.2729 for intensity) when the three dimensions were examined using the acoustic multidimensional prosody index (AMPI). Curiously, cue trading was not observed in the behavioral responses, but a complex pattern of trading emerged when comparing curves generated from confidence measures. This pattern, in addition to being nonlinear and nonmonotonic, also varied significantly by syllable position ($p<0.0001$). This suggests that the perceptual trading relationships among the prosodic cues are not only nonlinear, but also context-dependent.

Further reading

Boutsen, F.R. (2003). Prosody: the music of language and speech. *The ASHA Leader, 8*, 6-8.

Chrabaszcz, A., Winn, M., Lin, C.Y., & Idsardi, W.J. (2014). Acoustic cues to perception of word stress by English, Mandarin, and Russian speakers. *Journal of Speech, Language, and Hearing Research, 57*(4), 1468-1479.

Fry, D.B. (1958). Experiments in the perception of stress. *Language and Speech, 1*, 126-152.

TONGUE PALATAL CONTACTS DURING SPEECH PREPARATION IN 7 LANGUAGES

Susanne Fuchs[1], Özlem ÜnalLogacev[2]
[1]*Leibniz Centre General Linguistics, Berlin, Germany*
[2]*Istanbul Medipol University, Istanbul, Turkey*

Pauses before starting to speak are temporal intervals of high cognitive activity. They have been associated with planning the upcoming speech (Krivokapić, 2014). Many psycholinguistic studies have taken these pause intervals up to the acoustic landmark for speech onset as reaction time measures (RT) to support or reject psycholinguistic models. A few authors, however, have pointed out that articulatory motions occur well before this acoustic speech onset, depending on the respective initial segments (Kawamoto et al., 2008).

We followed these attempts and investigated articulatory behaviour during pauses by using a large speech production database. Tongue palatal contact patterns were analysed for 45 speakers of 7 languages (6 Turkish, 6 Norwegian, 3 Polish, 2 Bulgarian, 2 Mandarin Chinese, 14 English, 12 German) by means of electropalatography. The experimental design consisted of a reading task, where speakers had to read target words that were embedded in a carrier sentence. So far we annotated the acoustic onset of speech (phonation for sonorants, burst for stops and noise for fricatives) for the first 30 sentences of each speaker. Depending on the language, the initial segment of the carrier sentence occurring just after the pause, was either /p, k, s, h, o, a, j/. If speakers silently prepared the initial segment during the pause, we would expect languagespecific differences with a larger amount of tongue palatal contacts for /s, j/ (English, Norwegian) then /p, k/ (Polish, Bulgarian) and /o, a, h/ (Mandarin Chinese, Turkish, German). However, we found two languageindependent patterns: First, the most frequent pattern (34 speakers) was that speakers moved their tongue from a large percent of contacts at the palate (75100% PC) to almost no contact (020% PC) before speech onset. Second, speakers kept the tongue relatively constant throughout the pause with a small amount of contacts. The large articulatory variations in the first case were interpreted with respect to articulatory manoeuvres for inhalation through the mouth, with breathing noise during the pause. The second case was interpreted with respect to some rather neutral articulatory position. We suggest that articulatory motion for speech preparation may not only involve the anticipation of the upcoming segments, but may also reflect a complex interplay with respiration.

References

Kawamoto, A. H., Liu, Q., Mura, K., & Sanchez, A. (2008). Articulatory preparation in the delayed naming task. *Journal of Memory and Language, 58*(2), 347-365.

Krivokapić, J. (2014). Gestural coordination at prosodic boundaries and its role for prosodic structure and speech planning processes. *Philosophical Transactions of the Royal Society B* (Biology), 369, 20130397.

PERCEPTUAL LEARNING OF SPEECH PRODUCED BY A SPEAKER WITH DOWN SYNDROME

Alexandre Hennequin, Amélie Rochet-Capellan, Jean-Luc Schwartz, Marion Dohen
Université Grenoble Alpes, GIPSA-Lab, Grenoble, France

Down Syndrome (DS) is a genetic disease involving, among other impairments, a reduction of speech intelligibility. Although the behavioral phenotype is highly variable in DS, decreased speech intelligibility is common to all individuals (Kent & Vorperian, 2013). The speech of people with DS is, among other features, often qualified as dysarthric, although it does not embrace all the aspects of dysarthria (Kent & Vorperian, 2013). Generally, speakers with reduced intelligibility are better understood by people familiar with them (family, caregivers; Kumin, 1994 for DS). Familiarization, through perceptual training, could thus help people better understand speakers with DS.

Perceptual learning is a listener-oriented means to improve a speakers intelligibility. Studies show that perceptual learning, through familiarization with the speakers speech, could improve comprehension of speakers with dysarthria (Borrie, McAuliffe, & Liss, 2012). Familiarization also improves comprehension for typical speakers with an unfamiliar accent, in particular when listeners are asked to imitate the speaker during training (Adank et al., 2010). In this context, our study aims at investigating if familiarization to the speech of a person with DS can improve its comprehension. In particular, it evaluates if an active imitation of the speaker by the listener during familiarization improves comprehension. The procedure consists of three experimental phases: 1. Pre-test phase, to evaluate a baseline of the participants ability to understand the speaker; 2. Training phase, in which participants hear the speaker with feedback on what was actually said; 3. Post-test phase, to measure potential improvement in comprehension. The study involves three groups of participants: 1. Control group: no training phase; 2. Audition group: participants simply listen to the stimuli during training; 3. Imitation group: they are instructed to imitate the speaker. The stimuli consist of isolated French words with varying phonetic features to examine the specificities of perceptual learning as a function of phonological contrasts. Comparisons between groups make it possible to evaluate both the use of training in the understanding of speech produced by a speaker with DS and the influence of motor involvement during training. We will present preliminary results and further perspectives regarding the role of the motor system in perceptual learning, the role of perceptual learning in comprehension and listener-based training methods to improve speech intelligibility of people with DS.

This research has received funding from the European Research Council under the European Community's Seventh Framework Programme (FP7/2007-2013 Grant Agreement no.339152- "Speech Unit(e)s").

References

Adank, P., Hagoort, P., & Bekkering, H. (2010). Imitation improves language comprehension. *Psychological Science, 21*(12), 1903-1909.

Borrie, S. A., McAuliffe, M. J., & Liss, J. M. (2012). Perceptual learning of dysarthric speech: a review of experimental studies. *Journal of Speech, Language, and Hearing Research, 55*(1), 290-305.

Kent, R. D., & Vorperian, H. K. (2013). Speech Impairment in Down Syndrome: A Review. *Journal of Speech, Language & Hearing Research, 56*(1), 178-210.

AGEING EFFECTS ON SPEECH MOTOR CONTROL

Anne Hermes, Doris Mücke, Jane Mertens
University of Cologne, Cologne, Germany

Introduction Ageing is an inevitable natural process which entails changes at several physiological levels, including the central nervous system, the musculoskeletal system, the skeletal system, the cardiovascular system and the respiratory system. Crucially, increasing age affects motor control in general, involving a slowing down of, for example, the movements of the limbs (Brown, 1996). However, very little is known about how ageing affects speech production. Speech motor control almost exclusively involves fine motor control of the articulators, with the millimetre precision and split-second timing required to perform this highly complex task. As in motor control in general, a commonly reported effect of ageing on speech is that the tempo is slower, leading to a general slowing down of articulation rates (Amerman & Parnell, 1992). However, our knowledge of how ageing affects specific patterns of speech motor control, such as the coordination patterns within the oral system in the production of consonants and vowels, is limited by the fact that most studies are primarily based on acoustics. This study investigates ageing in speech motor control within the framework of Articulatory Phonology (Browman & Goldstein, 2000) using Electromagnetic Articulography (EMA).

Methods & Results We recorded 6 healthy older subjects (70-80 years old) and 6 healthy younger speakers (20-30 years old) using EMA (AG 501). Target words with a CV.CV structure such as <Bina> (labial) and <Dina> (alveolar) were embedded in carrier sentences. We labelled landmarks of the consonantal and vocalic gestures and analysed parameters within a mass-spring model such as gestural duration, peak velocity, displacement and stiffness. Preliminary results for two older speakers (age 72, 74) and two younger speakers (age 23, 29) indicate that differences are not restricted to temporal modifications such as slowing down; spatial parameters of the oral system are also affected. Density plots for the peak velocity of the consonantal gesture before it reaches its target for the labial closure for /b/ and the alveolar closure for /d/ reveal that consonantal gestures in older speakers are slowed down. Further the results indicated that there is less displacement for older speakers compared younger speakers.

Conclusion Preliminary results shed light on the articulatory modifications of ageing in speech motor control in the temporal but also in the spatial domain. Thus, an articulatory analysis provides further insights into the effects of ageing on speech motor control and allows comparison with the effects of ageing on motor control in general.

References

Amerman, J. D., & Parnell, M. M. (1992). Speech timing strategies in elderly adults. *Journal of Phonetics, 20*(1), 65-76.

Browman, C. P., & Goldstein, L. (2000). Competing constraints on intergestural coordination and selforganization of phonological structures. *Bulletin De La Communication Parlée, 5*, 25-34.

Brown, S. H. (1996). Control of Simple Arm Movements in the Elderly. *Changes in Sensory Motor Behavior in Aging* (Vol. 114, pp. 27-52). Elsevier Masson SAS.

CHANGES OF SOMATOSENSORY EVENT-RELATED POTENTIALS DURING SPEECH PRODUCTION

Takayuki Ito[1,2,3] Hiroki Ohashi[3], Eva Montas[3], Vincent L. Gracco[3,4]
[1]*GIPSA-lab, CNRS, Saint Martin Dheres Cedex, France*
[2]*Université Grenoble Alpes, GIPSA-Lab, Grenoble, France*
[3]*Haskins laboratories, New Haven, United States of America*
[4]*McGill University, Montréal, Canada*

Gating of sensory information around the onset of movement is a characteristic of a range of voluntary behaviors (Chapman, Spidalieri, & Lamarre, 1984). For speech, the amplitude of auditory cortical potentials is reduced during speech production, the so-called speech induced suppression (Niziolek, Nagarajan, & Houde, 2013). However, it is unknown whether somatosensory processing (or perception) is also suppressed during speech production. The current study focused on somatosensory event-related potentials (ERPs) associated with facial skin deformation and examined whether the somatosensory ERPs are changed during speech production tasks. Specifically, we investigated changes in the magnitude of somatosensory ERPs during speech posture with and without voicing together with changes associated with different speech utterances. We recorded ERPs from 64 scalp sites in response to somatosensory stimulation associated with facial skin stretch. Participants engaged in; 1) a vowel production task, and, 2) a non-speech task similar in which participants maintain the same posture without voicing. Three vowels (/a/, /i/, and /u/) were used in both the voicing and posture tasks. Somatosensory ERPs in the six task conditions (3 utterances x 2 voicing tasks) were compared with the ERPs in a resting, baseline condition. We applied Global field power (GFP) measures to evaluate the global tendency of ERPs on the cortical surface. GFP showed two peaks consistently around 160 ms and 320 ms after somatosensory onset in all seven conditions. The first peak amplitude was significantly different according to the vowel task, but there were no significant changes in the second peak. The first peak for the speech utterance, /u/ were reliably smaller than those in the other two vowel conditions. This was found consistently in both voicing and posture tasks. Cluster analysis of the temporal pattern of GFP demonstrated significant reduction of somatosensory GFP in /u/ condition relative to the control condition, while the other two vowel conditions were not reliably different. The results suggest that the sensitivity of somatosensory ERP associated with facial skin deformation changes according to the task. More importantly, there is evidence for a short latency (around 160 ms) gating of somatosensory input during speech. It appears that sensory suppression mechanisms are reflected in both sensory modalities associated with speech production and for the somatosensory system may be modulated differently relative to phonetic identity.

References

Chapman, C. E., Spidalieri, G., & Lamarre, Y. (1984). Discharge properties of area 5 neurones during arm movements triggered by sensory stimuli in the monkey. *Brain Res, 309*(1), 63-77.

Niziolek, C. A., Nagarajan, S. S., & Houde, J. F. (2013). What does motor efference copy represent? Evidence from speech production. *J Neurosci, 33*(41), 16110- 16116. doi: 10.1523/JNEUROSCI.2137-13.2013

EFFECTS OF EXPOSURE TIME VS. PRACTICE RATE ON SPEECH AUDITORY-MOTOR ADAPTATION AND DE-ADAPTATION

Kwang Seob Kim[1], Takashi Mitsuya[1], Ludo Max[1,2]
[1] *University of Washington, Seattle, United States of America*
[2] *Haskins Laboratories, New Haven, United States of America*

Sensorimotor adaptation plays a critical role in learning, maintaining, and refining voluntary movements. Surprisingly, few research efforts have been directed at uncovering the processes that occur whenafter a period of sensorimotor adaptationthe feedback alteration or movement perturbation is removed (de-adaptation). The lack of research attention for de-adaptation may be related to the fact that subjects behavioral responses gradually return to their pre-perturbation form, and that this change in behavior is often interpreted as an expected "forgetting" of newly learned internal models. This interpretation is typically based on the claim that de-adaptation is a faster process than adaptation. For some tasks, however, the rate of de-adaptation is not faster than the rate of initial adaptation (Davidson & Wolpert, 2004, for limb movements; Purcell & Munhall, 2006, for speech production) even though the former involves re-adjusting movement parameters so that they once again account for sensorimotor transformations that previously had been experienced for many years. In addition, for limb movements, de-adaptation can take longer in children than inif de-adaptation represents forgetting, this result would mean that children have better motor memory retention than adults. Instead, it may be more likely that de-adaptation itself represents new sensorimotor learning. Hence, de-adaptation may be better conceptualized as requiring an active re-learning of the appropriate sensorimotor maps rather than a process of forgetting.

Addressing this question for auditory-motor learning in speech production, we investigated adaptation and de-adaptation as a function of the number of practice trials versus the amount of practice time during the corresponding phases of the experiment. For example, fewer trials performed per period of time in the post-exposure phase should lead to faster/increased de-adaptation if this phenomenon reflects mainly forgetting but to slower and decreased de-adaptation if the phenomenon reflects active sensorimotor re-learning of the non-perturbed environment.

Forty adult subjects produced target words in pre-exposure (baseline), exposure (adaptation), and post-exposure (de-adaptation) phases of the experiment. During the exposure phase, all formants in the real-time auditory feedback signal were shifted 250 cents up. Participants were randomly assigned to four groups which differed in the rate at which trials were produced (either 18/min or 6/min) during the adaptation and de-adaptation phases. Presentation of the findings will focus on the dissociation of effects due to the number of practice trials versus the amount of practice time and implications for speech motor learning and forgetting. [Funded by NIH/NIDCD R01DC014510]

References

Davidson, P. R., & Wolpert, D. M. (2004). Scaling down motor memories: de-adaptation after motor learning. *Neuroscience Letters, 370*(2-3), 102-107. https://doi.org/10.1016/j.neulet.2004.08.003

Purcell, D. W., & Munhall, K. G. (2006). Adaptive control of vowel formant frequency: Evidence from real-time formant manipulation. *The Journal of the Acoustical Society of America, 120*(2), 966-977. https://doi.org/10.1121/1.2217714

CAN PERCEPTUAL TRAINING ALTER THE ABILITY TO MATCH A VISUALLY PRESENTED ACOUSTIC TARGET?

Adam Klaus[1], Daniel Lametti[2], Doug Shiller[3], Tara McAllister Byun[1]
[1]*New York University, New York, United States of America*
[2]*University of Oxford, Oxford, United Kingdom*
[3]*University of Montréal, Montréal, Canada*

Studies investigating how speakers respond to perturbed auditory feedback have played a major role in illuminating mechanisms that underlie speech production, but direct clinical applications of this line of investigation have remained limited. The present study aims to extend the translational potential of this line of research, focusing on the finding that speech-motor learning in the context of perturbed auditory feedback can be enhanced when preceded by a period of auditory-perceptual training (e.g., Lametti et al., 2014). Specifically, we investigate whether a task of matching targets using *visual-acoustic biofeedback* (e.g., McAllister Byun & Campbell, 2016) is a more technologically accessible method that could recapture some of the crucial properties of auditory feedback perturbation. Biofeedback involves an explicit shift to match a target rather than an implicit shift to compensate for a perturbation, but it may nevertheless be alterable through perceptual training.

Participants were 40 female native English speakers aged 18-35. Following Lametti et al. (2014), a synthetic 11-step continuum from /ɛ/ *(head)* to /æ/ *(had)* was generated to test and train perception. After a baseline identification task, participants completed training in which they received feedback to shift the location of the /ɛ/ - /æ/ perceptual boundary. Participants were randomly assigned to a "*head*-shift" condition, which encouraged them to identify ambiguous tokens as *had*, or a contrasting "*had*-shift" condition. Participants then completed a baseline production task and a visual-acoustic biofeedback matching task. Participants were familiarized with a real-time LPC spectrum of their speech, focusing on the first formant (F1). They were then instructed to produce the word *head* while trying to match F1 to a visual target at a higher (more /æ/-like) frequency. This change should be more likely to cross the /ɛ/ - /æ/ perceptual boundary for the *head*-shift than the *had*-shift group. On the hypothesis that speakers should be less willing to alter vowel production when the output crosses the boundary into another phoneme, we predicted that the *had*-shift group would be more effective in matching the visual target. The two groups did not differ significantly in boundary location for *head-had* at baseline, but after perceptual training, boundary locations differed significantly in the predicted direction (all $p<.001$). In addition, on average, participants significantly shifted F1 toward the visual target. However, groups did not differ significantly in normalized F1 frequency during and after production training with biofeedback. Theoretical and empirical explanations for this finding will be discussed.

References

Lametti, D.R., Krol, S.A., Shiller, D.M., & Ostry, D.J. (2014). Brief Periods of Auditory Perceptual Training Can Determine the Sensory Targets of Speech Motor Learning. *Psychological Science, 25* (7), 1325 1336.

McAllister Byun, T., & Campbell, H. (2016). Differential effects of visual-acoustic biofeedback intervention for residual speech errors. *Frontiers in Human Neuroscience, 10.* http://dx.doi.org/10.3389/fnhum.2016.00567.

HOW FLEXIBLE ARE SPEAKERS' INTERNAL REPRESENTATIONS OF THE ARTICULATORY-TO-ACOUSTIC MAPPING?

Eugen Klein[1], Jana Brunner[1], Phil Hoole[2]
[1]*Humboldt-Universität zu Berlin, Berlin, Germany*
[2]*Ludwig-Maximilians-Universität München. Munich, Germany*

Rochet-Capellan and Ostry (2011) show that, within the same experimental session, speakers adapt their production of the English vowel /ɛ/ to three different auditory shifts of the first formant (F1). That means that speakers do not correlate the degree of jaw opening with the vowel highness as they rapidly map and remap different articulatory configurations to the intended acoustic output. To further investigate the flexibility and potential constraints of the articulatory-to-acoustic mapping, we conducted a bidirectional perturbation study where compensatory movements interacted with coarticulatory effects.

Speech of 18 native Russian speakers was acoustically recorded while they were producing CV syllables which contained the close central unrounded vowel /ɨ/. The second formant (F2) of the vowel /ɨ/ was perturbed in real-time during the production and fed back to the participants via insert headphones. In order to encourage participants to produce the target vowel with two different articulatory configurations, the F2 was shifted upwards on one half of the trials and downwards on the other half. The trials were presented in random order. The direction of the perturbation was dependent on the consonant preceding the target vowel (/d/ or /g/) and was counterbalanced for the place of articulation between the 18 participants. Due to coarticulatory effects, baseline F2 was expected to be higher for /dɨ/ compared to /gɨ/. For nine participants, F2 was decreased in /dɨ/ and increased in /gɨ/, such that the compensatory movements were expected to act in the same direction as the coarticulatory effects. In that case, baseline F2 values were expected to drift apart over the duration of the experiment. For the other nine participants, the perturbation direction was swapped for both syllables, putatively preventing effective compensation as it would be constrained by the coarticulatory effects and the baseline F2 values had to cross between the two syllables.

Our analyses show that, by the end of the experiment, participants produced significantly different F2 values for the two syllables depending on the direction of the perturbation. There was no compensatory advantage for one of the two experimental groups; the F2 either kept drifting apart or crossed between the two syllables, but there were inter- and intra-individual differences in the compensation strategies. We interpret our findings as support for the hypothesis that goals of the speech production are defined in the auditory space, but their achievement is dependent on "speaker-specific internal representation of articulatory-to-acoustic relationships" (Savariaux et al., 1995).

References

Rochet-Capellan, A., & Ostry, D. J. (2011). Simultaneous acquisition of multiple auditory-motor transformations in speech. *Journal of Neuroscience, 31*(7), 2657-2662.

Savariaux, C., Perrier, P., & Orliaguet, J. P. (1995). Compensation strategies for the perturbation of the rounded vowel [u] using a lip tube: A study of the control space in speech production. *The Journal of the Acoustical Society of America, 98*(5), 2428-2442.

SPEECH ACOUSTICS BEFORE AND AFTER COCHLEAR IMPLANTATION IN POST-LINGUALLY DEAFENED ADULTS

Xaver Koch[1,2], Thordis Neger[1], Wendy Huinck[3,4], Emmanuel Mylanus[3,4], Esther Janse[1,4]
[1]*Radboud University, Nijmegen, The Netherlands*
[2]*Humboldt University, Berlin, Germany*
[3]*RadboudUMC, Nijmegen, The Netherlands*
[4]*Donders Institute for Brain, Cognition and Behaviour, Nijmegen, The Netherlands*

Prolonged deprivation of auditory feedback may negatively affect speech production accuracy, as shown by reports of reduced vowel space, changed formant values and reduced sound contrasts for challenging sounds such as sibilants in post-lingually deafened adults, compared to normal-hearing controls (e.g. Waldstein, 1990; Lane and Webster, 1990). Post-lingually deafened adults who undergo cochlear implantation have been hypothesized to enter a phase of retuning of their feed-forward motor commands to account for the restored, altered auditory feedback (Lane et al., 2007). The present study investigates relatively early changes (two weeks to three months after cochlear implantation) in speech acoustics due to cochlear implantation. Nine post-lingually deafened novice CI users read Dutch target words containing the sibilant sounds /s/ and /ʃ/ and the five vowels [aː, ɛ, i, ɔ, u] during three test sessions: before implantation (T0), directly after implant activation (T1), and approximately three months after activation of their CI (T2). Their acoustic realizations of the target words were compared to those of an age and gender-matched control group (also tested three times, with the same amount of time elapsing between test sessions). Spectral means of the sibilants and vowel formants (F1, F2) were analyzed. Results show that, prior to CI surgery, sibilant contrasts were diminished and vowel formants showed a centralization tendency for the deafened patients relative to the control group. After cochlear implantation, formant values were produced more control-like. Across test sessions, both groups enlarged the sibilant contrast, yet in different ways. The CI users enlarged the sibilant contrast mainly by lowering the spectral mean of the post-alveolar sound /ʃ/, whereas the controls heightened the sibilant contrast more symmetrically by increasing spectral means of the /s/ productions as well as decreasing spectral means for /ʃ/ productions. The CI recipients less symmetrical way to enhance the contrast may be due to reduced spectral resolution for the high frequencies. Furthermore, among CI users, duration of hearing loss predicted the amount of change in sibilant contrast after CI activation. The absence of auditory feedback thus impairs phonemic contrasts for sibilants over time in deafened participants. Once auditory feedback is available again sibilant contrasts improve immediately, especially for participants with shorter duration of deafness.

References

Lane, H. and Webster, J. W. (1991). Speech deterioration in postlingually deafened adults. *J. Acoust. Soc. Am., 89*(2), 859-866.

Lane, H., Mathies, M., Guenther, F., Perkell, J. S., Stockmann, E., Tiede, M., and Vick, J. (2007). Effects of short- and long-term changes in auditory feedback on vowel and sibilant contrasts. *J. Speech Lang. Hear. R., 50*, 913-927.

Waldstein, R. S. (1990). Effects of postlingual deafness on speech production: Implications for the role of auditory feedback. *J. Acoust. Soc. Am., 88*(5), 2099-2114.

A MULTIDIMENSIONAL STUDY OF LOUD SPEECH IN GERMAN

Laura L. Koenig[1,2], Susanne Fuchs[3]
[1]*Adelphi University, Garden City, United States of America*
[2]*Haskins Laboratories, New Haven, United States of America*
[3]*Leibniz Centre General Linguistics, Berlin, Germany*

Loudness modulation is a natural part of typical speech production. Loud speech has also been proposed as a treatment method for certain speech pathologies, e.g., those subsequent to Parkinsons disease (e.g., Sapir et al., 2007). Clarifying the mechanisms that underlie the production of loud vs. normal speech is therefore important both for basic speech research and for evaluating the empirical underpinnings of therapeutic methods. Past work demonstrates that louder speech may a) be associated with changes across respiratory, laryngeal, and supralaryngeal systems, and b) affect speech over various temporal intervals. Thus, although many authors have evaluated loudness variation over long epochs such as the breath group, loudness also impacts the acoustic and articulatory realization of single segments (e.g., Schulman, 1989), and affects laryngeal function at the level of individual glottal cycles (e.g., Mooshammer, 2010).

Our aim is to take all these levels into account and investigate loud speech on a multidimensional basis across speech tasks. We gathered simultaneous recordings from eleven German females of acoustics (to analyse f0, duration, intensity, vowel formants), laryngography (to assess phonation), intraoral pressure (to infer subglottal pressure) and inductance plethysmography (to observe abdominal and thoracic volume changes). Loudness variation was elicited naturalistically in reading, question-answer tasks, and interactions with an interlocutor. In earlier analyses, we observed considerable variation in formant frequency changes as a function of loudness across tasks, speakers, and vowel types. Here we focus on the respiratory, aerodynamic, and laryngeal measures to quantify the degree to which they a) vary across tasks and speakers, and b) correlate with each other, positively or negatively. Do some speakers rely more on laryngeal than supralaryngeal mechanisms when producing loud speech? Do speakers with more extreme supralaryngeal variation do likewise for breathing and laryngeal contact? Do results based on read speech, one of the most common experimental paradigms, generalize to spontaneous speech? Having a fuller picture of loudness modulation in typical speakers may ultimately provide some insight into individual variation in treatment efficacy for loud speech approaches.

References

Sapir, S., Spielman, J. L., Ramig, L. O., Story, B. H., & Fox, C. (2007). Effects of intensive voice treatment (the Lee Silverman Voice Treatment [LSVT]) on vowel articulation in dysarthric individuals with idiopathic Parkinson disease: acoustic and perceptual findings. *Journal of Speech, Language, and Hearing Research, 50*, 899-912.

Schulman, R. (1989). Articulatory dynamics of loud and normal speech. *Journal of the Acoustical Society of America, 85*, 295-312.

Mooshammer, C. (2010). Acoustic and laryngographic measures of the laryngeal reflexes of linguistic prominence and vocal effort in German. *Journal of the Acoustical Society of America, 85*, 127, 1047-1058.

PERCEPTION-PRODUCTION LINK IN ENGLISH LEXICAL STRESS ACQUISITION AMONG HONG KONG CANTONESE LEARNERS OF ENGLISH

Wience Wing Sze Lai[1,2], Manwa L. Ng[2]
[1] *The Hong Kong Polytechnic University, Hong Kong*
[2] *The University of Hong Kong, Hong Kong*

The present study investigated the link between perception and production in English lexical stress acquisition by Hong Kong Cantonese learners of English (CS) through examining F0, intensity and duration. The majority of perception-production studies on segmental instead of suprasegmental features, and on Mandarin instead of Cantonese learners suggested the need for this study, as an attempt to to verify the main cue connecting tone and stress.

To address the issue, three experiments were carried out: (1) an acoustic experiment examining production of seven disyllabic English noun-verb pairs, (2) a perceptual experiment in the form of a listening evaluation test of (1), and (3) a perception experiment with manipulation of F0, intensity and duration of a disyllabic English noun-verb pair (Lai & Ng, 2015). Twenty-two CS (11 highly proficient, HCS, and 11 less proficient, LCS), with 14 native English speakers (NS) serving as controls, were recruited for Experiment 1, and F0 (Hz), intensity (dB) and duration (ms) values were analysed from their productions. In Experiment 2, 13 NS, with 30 CS serving as controls, were recruited to listen to the CS tokens produced in Experiment 1 and to perform noun-verb distinction and provide acceptability ratings. 42 CS, with seven control NS, participated in Experiment 3 and performed noun-verb distinction.

The findings revealed that CS tended to rely on F0 but neglected intensity in both the production and perception of English lexical stress, as compared to NS, suggesting a possible perception-production link and L1 influence models (Lai & Ng, 2014). In addition, the overuse of F0 and proper use of intensity and duration by HCS compared to LCS suggested the impact of L2 proficiencies. The above appeared to support Motor Theory with a specialised perception-production module proposed. While Perceptual Assimilation Model explained the assimilation English lexical stress and non-stress contrast, Speech Learning Model and Native Language Model accounted for the warped perceptual space. With reference to Attention-to-Dimension Models, explicit teaching or audio-visual training may redirect HCSs attention from F0 to the other cues and LCSs attention to intensity and duration.

Vowel quality was suggested as an additional cue for investigation in future studies. Yet, the present findings inspire more effective regimens in teaching English lexical stress and future studies on the acquisition of English speech suprasegmentals such as rhythm and intonation.

References

Lai, W. W., & Ng, M. (2014). Identification of dominant cue(s) for the perception of English word stress by Cantonese ESL speakers. In *Proceedings of the 2014 International Conference on Phonetic Research and Language Learning (ICPRLL) & English Phonetic Conference in China (EPCC)*, Changsha, China: Hunan University.

Lai, W. W. S., & Ng, M. (2015). The relationship between perception and production in Cantonese ESL speakers English word stress acquisition. In *Proceedings of the 8th Annual International Conference on Language & Linguistics (AICLL2015)* (pp. 60), Athens, Greece: Athens Institute for Education and Research.

TESTING THEORETICAL MODELS OF INTER-SPEAKER COORDINATION BY USING A HUMAN DYNAMICAL CLAMP

Leonardo Lancia[1], Thierry Chaminade[2], Noël Nguyen[3], Laurent Prévot[3]
[1]*Laboratoire de Phonétique et Phonologie, Paris, France*
[2]*Institut de Neuroscience de la Timone, Marseille, France*
[3]*Laboratoire Parole et Langage, Aix en Provence, France*

In order to compare two different accounts of inter-speaker coordination, we use an artificial agent (AA) that behaves as a human dynamical clamp (Dumas et al., 2014) by adapting in real time its behaviour to that a human speaker (HS) during the simultaneous repetition of simple speech utterances. According to theoretical models based on prediction (e.g.: Pickering and Garrod, 1013), to anticipate their interlocutors behaviour, speakers adapt the forward models used to control their own speech production. This parameterization results in a tendency to imitate their partners behaviour each time they engage in a conversation. If however, the sensorimotor systems of speakers in conversation are coupled as subsystems of a complex dynamical system, coordination may not require that each speaker predicts his/her interlocutors behaviour (cf. Stepp and Turvey, 2010). Moreover, by propagating across time scales and levels of activity, coordination can constrain the behaviour of the individual sensorimotor systems and induce behavioural matching, if this favours the specific coordinative patterns produced. To test these predictions, we asked 10 Italian speakers to repeat without interruptions the utterance /topkop/ simultaneously with the AA during 12 seconds. In some uninterrupted sequences, the AA was parameterized to cooperate with the HS in producing its syllables simultaneously (in-phase) with those produced by the HS. In these sequences, both theoretical approaches predict some degree of imitation. In other sequences, the AA was parameterized to produce its syllables in-between those produced by the HS (in anti-phase). Since HSs spontaneously tend toward in-phase coordination, in these sequences the two partners compete to impose different coordinative relations. If in competitive interactions imitation does not favour coordination, accounts based on dynamical coupling would predict a lower degree of imitation with respect to cooperative interactions. In order to facilitate the quantification of imitation, we varied the prosody of the AAs utterances from one sequence to the other. We found that, in line with models based on prediction, speakers tend to imitate the AAs f0 during both kinds of interaction, with no significant differences between the conditions. However, in cooperative interactions we observed a positive correlation between the degree of imitation and the stability of inter-partner coordination. Therefore, given appropriate conditions, inter-speaker coordination can spread through time scales and levels of activity favoring behavioral matching.

References

Dumas, G., de Guzman, G. C., Tognoli, E., & Kelso, J. S. (2014). The human dynamic clamp as a paradigm for social interaction. *Proceedings of the National Academy of Sciences, 111*(35), E3726-E3734.

Pickering, M. J., & Garrod, S. (2013). An integrated theory of language production and comprehension. *Behavioral and Brain Sciences, 36*(04), 329-347.

Stepp, N., & Turvey, M. T. (2010). On strong anticipation. *Cognitive systems research, 11*(2), 148-164.

AUDITORY FEEDBACK, MOTOR PREDICTION AND SENSORY GOAL IN LEARNING NON-NATIVE SPEECH SOUNDS: EVIDENCE FROM ADULT MANDARIN CHINESE SPEAKERS

Xiaoluan Liu, Xing Tian
New York University Shanghai, Shanghai, China

In speech acquisition, three factorsauditory feedback, motor prediction and sensory goalcollectively contribute to the development and maintenance of accurate speech production (Guenther, 1995; Hickock, 2012). However, it remains unclear how the three factors interact in learning non-native speech sounds, especially by adults whose native language and dialect are already well established. The present study investigates the relation among auditory feedback, motor prediction and sensory goal in speech acquisition by examining adult Mandarin Chinese speakers learning of two types of non-native speech sounds: German vowel /ö/ and Tianjin Tone 1. German vowel /ö/ is a novel speech sound that does not exist in Mandarin vowel inventory. Tianjin Tone 1 is close to but still distinct from Mandarin Tone 4 in perception and production, making it characteristic of Tianjin dialect which is a variation of Mandarin Chinese. In this study, Mandarin speakers without prior knowledge of German and Tianjin dialect learned the two types of sounds under two conditions: with auditory feedback (unmasked condition) or feedback masked with noise (masked condition). The results, revealed by acoustic differences between the learned and target speech, showed that learning German vowel /ö/ was better in the unmasked condition than the masked condition; whereas learning Tianjin Tone 1 was equally bad in both masked and unmasked conditions. These results suggest that without auditory feedback, motor prediction and sensory goal alone are not sufficient for controlling speech production deviance when learning nonnative sounds that do not exist in ones native speech repertoire (e.g., German vowel /ö/). On the other hand, auditory feedback is not always effective, especially if the novel speech target (e.g., Tianjin Tone 1) is in perception and production similar to an existing sound in ones native speech inventory. Together, these findings imply a weighted contribution of auditory feedback, motor prediction and sensory goal to learning different types of non-native speech sounds by adults.

References

Guenther, F. H. (1995). Speech sound acquisition, coarticulation, and rate effects in a neural network model of speech production. *Psychological Review, 102*, 594-621.

Hickok, G. (2012). Computational neuroanatomy of speech production. *Nature Reviews Neuroscience, 13*, 135-145.

TONGUE- AND JAW SPECIFIC CONTRIBUTIONS TO INCREASED VOWEL ACOUSTIC CONTRAST IN RESPONSE TO SLOW, LOUD, AND CLEAR SPEECH

Antje Mefferd
Vanderbilt University Medical Center, Nashville, United States of America

Purpose Although it is well-known that slow, loud, and clear speech elicit increases in acoustic vowel contrast, articulator-specific changes in response to these speech modifications and their relative contributions to the increase in acoustic vowel contrast remain poorly understood. Based on previous kinematic findings, this study sought to test the hypothesis that increases in acoustic vowel contrast associated with slow speech are predominantly tongue-driven whereas those associated with loud speech are predominantly jaw-driven. Finally, increases in acoustic vowel contrast during clear speech were expected to be tongue- and jaw-driven.

Method Twenty typical talkers (18-28 years old, 9 males, 11 females) participated in the study. Speech kinematic data was recorded using 3D electromagnetic articulography (AG501). All data was corrected for head movements, re-expressed relative to an anatomically-based coordinate system, and low pass filtered. Posterior tongue movements (recorded approx. 4cm from tongue tip) and jaw movements (recorded at the front central incisors) were of interest in this study. Participants repeated the phrase "See a kite again" five times using typical speech, then five times at slow speech, loud speech, and clear speech. The onset and offset of the target /ai/ in "kite" were defined by the positional extrema of the posterior tongue sensor in the vertical dimension. 3D positions of the jaw sensor and posterior tongue sensor were extracted at the defined onset and offset and the 3D Euclidean distance between onset and offset was calculated for each articulator. A decoupling algorithm (a modified version of Westbury et al., 2002) was applied to the posterior tongue distance measure to determine displacements of the decoupled tongue and the jaw (at the posterior location). In the acoustic domain, the F2 minimum during /a/ and the F2 maximum during /i/ and their corresponding F1 values were extracted to calculate acoustic vowel contrast in F1-F2 vowel space.

Results & Conclusions Findings of linear regression analyses supported the hypotheses for clear and slow speech. That is, during slow speech increases in acoustic vowel contrast were predominantly tongue-driven whereas during clear speech increases in acoustic vowel contrast were driven by both, the tongue and jaw. However, contrary to our hypothesis, jaw displacement changes during loud speech accounted for a much lower percentage of the variance in acoustic vowel contrast than tongue displacement changes. Findings will be discussed within the context of previous kinematic and acoustic studies and with regards to their clinical implications for treatments.

Further reading

Schulman, R. (1988). Articulatory dynamics of loud and normal speech. *Journal of the Acoustical Society of America, 85*(1), 295-312.

Tasko, S. & Greilick, K. (2010). Acoustic and articulatory features of diphthong productions: A speech clarity study. *Journal of Speech, Language, and Hearing Research, 53*, 84-99.

THE EFFECT OF AUDITORY DELAY ON AUDITORY MOTOR LEARNING

Takashi Mitsuya[1], Kevin G. Munhall[2], and David W. Purcell[3]
[1] *University of Washington, Seattle WA. USA*
[2] *Queens University, Kingston, ON, Canada*
[3] *Western University, London, ON, Canada*

Adaptive learning of motor behavior using a perturbed sensory feedback paradigm assumes that the motor control system processes the perturbed feedback is self-generated. One factor that may play a crucial role for assessing the feedback is self-generated is its temporal (in)congruence to the causal action. The importance of this factor for auditory motor learning was demonstrated by a recent work by Max and Maffett (2015). In their study, an auditory delay of 100 ms eliminated compensatory production in response to perturbed formants. It was postulated that speech motor control system rejected the delayed feedback as self-generated. In the present study, we examined whether the effect of auditory delay on auditory motor learning would depend on types of formant perturbation used. Instead of perturbing all formants as Max and Maffetts study, only the first formant of auditory feedback was perturbed, while systematically delaying auditory delay. This perturbation changed only perceived jaw openness of the vowel being produced (e.g., Traunmüller, 1981; Fahey et al., 1996); thus, it was speculated that auditory error induced by this perturbation might be easier for the control system to map onto articulatory correction. Our results showed a small yet significant adaptation even with 100 ms of auditory delay, contrary to the past report. We concluded that tolerance of auditory delay is not fixed but it depends on the ease and/or feasibility of the correction estimates that the control system processes. Moreover, we found that auditory motor learning had a linear-inverse relationship with auditory delay, such that as auditory delay decreased, larger compensatory formant production was observed. It is speculated that undelayed and unperturbed somatosensory feedback may be given more preference as auditory delay increases.

References

Fahey, R. P., and Diehl, R. L. (1996). The missing fundamental in vowel height perception. *Percept. Psychophys. 58*, 725-733.

Max, L., and Maffett, D. G. (2015). Feedback delays eliminate auditory-motor learning in speech production. *Neurosci. Lett., 591*, 25-29.

Traunmüller, H. (1981). Perceptual dimension of openness in vowels. *J. Acoust. Soc. Am. 69*, 1465-1475.

ACOUSTIC AND NEURAL MEASURES OF SPEECH ERROR DETECTION AND CORRECTION IN PERSONS WITH APHASIA

Caroline A. Niziolek, Swathi Kira
Boston University, Boston, United States of America

Recent neuroimaging research in healthy speakers suggests that the auditory system constantly monitors its own speech for small deviations from intended speech sounds, and that successful monitoring may drive an unconscious correction of these deviations before they are realized as errors (Niziolek, Nagarajan, & Houde, 2013). Phonemic errors of this sort are a hallmark of speech in persons with aphasia (PWA), and are often accompanied by abnormalities in perception and feedback processing (Chapin, Blumstein, Meissner, & Boller, 1981; Robson, Grube, Lambon Ralph, Griffiths, & Sage, 2013). Here, we investigated the hypothesis that speech errors in aphasia coincide with impairments in using feedback to correct speech acoustics online.

Ten PWA and ten age-matched controls took part in a magnetoencephalography (MEG) study to measure auditory sensitivity to self-produced acoustic variability. Participants produced 200 repetitions of three monosyllabic words ("eat", "Ed", "add") while neural activity was recorded. This "speak" condition was interleaved with a "listen" condition in which recorded audio from the speak trials was played back to participants through earphones. We considered two measures of feedback control: *detection*, the cortical sensitivity to acoustic deviations in ones own speech, and *correction*, the extent to which these deviations are behaviorally corrected online. First and second formant trajectories were calculated to determine the deviation of each trial from the 2D formant median. Detection was defined as a differential modulation of the auditory M100 response for large (vs. small) deviations from typical vowel acoustics. Correction was defined as formant movement toward the median over the time course of the syllable, lessening acoustic deviation.

Evidence for detection of self-produced acoustic deviations was found in both controls and PWA. In controls, neural activity was consistently modulated by acoustic deviation only in the left hemisphere. In contrast, PWA showed only weak activation of left auditory regions, but right hemisphere activity was modulated by acoustic deviation. Behavioral correction was also largely intact: PWA had greater acoustic variability than controls at vowel onset, but both groups exhibited vowel centering, significantly decreasing variability over the course of the syllable. These analyses suggest that online feedback correction mechanisms are at work in person with aphasia, partially compensating for feedforward production deficits.

References

Chapin, C., Blumstein, S. E., Meissner, B., & Boller, F. (1981). Speech production mechanisms in aphasia: A delayed auditory feedback study. Brain and Language, 14(1), 106-113.

Niziolek, C. A., Nagarajan, S. S., & Houde, J. F. (2013). What Does Motor Efference Copy Represent? Evidence from Speech Production. The Journal of Neuroscience, 33(41), 16110-16116.

Robson, H., Grube, M., Lambon Ralph, M. A., Griffiths, T. D., & Sage, K. (2013). Fundamental deficits of auditory perception in Wernickes aphasia. Cortex, 49(7), 1808-1822.

OROFACIAL SOMATOSENSORY INPUTS MODULATE PERCEPTUAL SHIFTS DUE TO SPEECH MOTOR LEARNING

Hiroki Ohashi[1], Takayuki Ito[1,2,3]
[1] *Haskins laboratories, New Haven, United States of America*
[2] *Gipsa-lab, CNRS, Saint Martin Dheres Cedex, France*
[3] *Université Grenoble Alpes, Grenoble, France*

Orofacial somatosensory input during listening to speech changes speech perception in a manner that is consistent with the direction of the stimulation (Ito, Tiede, & Ostry, 2009). One interpretation of the link between somatosensory and auditory system is that the somatosensory input may tune auditory representations during speech motor learning. The present study tested this prediction by examining whether somatosensory input during speech motor learning modulates speech perception associated with altered auditory feedback. Native speakers of American English participated in a speech motor adaptation study in which perceptual discrimination was assessed before and after speech motor adaptation (Shiller, Sato, Gracco, & Baum, 2009) with and without additional somatosensory stimulation. We examined changes in perception and production in /sh/-/s/ utterances in which horizontal lip motion contrasts between the two phonemes. Subjects repeated the phrase a shed and altered-auditory feedback changed the spectral centroid on-line to sound more like /s/. Auditory feedback was accompanied by either; 1) somatosensory stimulation in the forward direction (ASF); 2) stimulation in the backward direction (ASB); or 3) no stimulation (A). We also included a no perturbation (control) condition. Adaptation was quantified as changes in the spectral centroid frequency of the produced /sh/ sounds. Before and after the altered feedback conditions, phoneme identification tests were carried out to characterize the boundary between each subjects /s/ and /sh/ categories. The identification test employed a set of acoustic continuum from /s/ to /sh/. The thresholds prior and post speech adaptation were compared to quantify the perceptual shift due to the training. The three groups (ASF, ASB and A) showed similar adaptation effects regardless of the additional somatosensory stimulation. However, the phoneme identification revealed that the perceptual shift due to the speech training with the backward stretch (ASB) was different from the other two conditions with altered auditory feedback (ASF and A). While the two conditions (ASF and A) showed the reliable perceptual shifts, the shift associated with ASB was eliminated and was similar to the control condition. The findings indicate that somatosensory input during speech motor learning is capable of modulating perceptual representations and that somatosensation during speech motor development may be a key element in both establishing and tuning the representations for speech.

References

Ito, T., Tiede, M., & Ostry, D. J. (2009). Somatosensory function in speech perception. *Proc Natl Acad Sci U S A, 106*(4), 1245-1248.

Shiller, D. M., Sato, M., Gracco, V. L., & Baum, S. R. (2009). Perceptual recalibration of speech sounds following speech motor learning. *J Acoust Soc Am, 125*(2), 1103-1113.

RECRUITMENT OF THE ARTICULATORY MOTOR CORTEX DURING SPEECH PERCEPTION IN OLDER ADULTS

Muriel T.N. Panouillères, Riikka Möttönen
University of Oxford, Oxford, United Kingdom

Speech perception is effortless for young adults, but older adults have often difficulties to understand speech. It has been suggested that the decrease in speech perception skills is a general feature of ageing that would be mostly related to the brain senescence (Wong et al., 2009). The articulatory motor cortex has been shown to interact with the auditory cortex during speech processing and to contribute to speech perception in young adults (Möttönen & Watkins, 2012). We hypothesized that changes in the recruitment of the motor cortex could be related to the difficulties in perceiving speech in older adults. Here, we investigated the effects of aging and hearing loss on the recruitment of the articulatory motor cortex during listening to speech in noise. We used Transcranial Magnetic Stimulation (TMS) to evaluate the excitability of the articulatory motor cortex when listening to speech in 18 young adults (21.9 $\pm$ 2.9 years old) and 21 older adults (68.7 $\pm$ 3.2 years old). Single-pulse TMS was applied to the tongue representation in the left primary motor cortex, eliciting motor evoked potentials (MEPs) in the tongue muscle, while participants listened to sentences with and without noise, speech-spectrum signal-correlated noise and white noise. As a control, TMS was also applied to the hand representation in the left primary motor cortex in each participant. We also assessed hearing abilities in both ears using pure-tone audiometry and speech-in-noise perception using the Quick Speech-In-Noise test in each participant. Eleven older adults had mild to moderate hearing loss at the speech frequencies (250 Hz to 4KHz) while the 10 had normal hearing. Our MEP results show that the excitability of the tongue motor cortex was enhanced when listening to speech with and without noise in young and older adults with normal hearing. However, the older adults with hearing loss did not show any increase in the tongue MEPs size when listening to speech. Moreover, there was a correlation in the older adults group between the ability to perceive speech-in-noise and the facilitation of the tongue motor cortex, demonstrating that the older adults with difficulties in perceiving speech in noise show the lowest facilitation of the tongue motor cortex. Our results demonstrate that the older adults with hearing loss and speech perception difficulties do not recruit the articulatory motor cortex to the same extent as older adults with normal hearing and speech perception.

References

Möttönen, R., & Watkins, K. E. (2012). Using TMS to study the role of the articulatory motor system in speech perception. Aphasiology, 26(9), 1103-1118. https://doi.org/10.1080/02687038.2011.619515

Wong, P. C. M., Jin, J. X., Gunasekera, G. M., Abel, R., Lee, E. R., & Dhar, S. (2009). Aging and cortical mechanisms of speech perception in noise. *Neuropsychologia, 47*(3), 693-703. https://doi.org/10.1016/j.neuropsychologia.2008.11.032

REWARD LEARNING IN SPEECH

Benjamin Parrell
University of Delaware, Newark, United States of America

There are two established pathways for motor learning. In sensorimotor learning, differences between expected and perceived sensory outcomes of actions (sensory error) drives recalibration of motor commands to realign these signals. In reinforcement learning, reward (positive or negative) drives alterations of behavior without changing the relationship between actions and sensory consequences. In speech research, sensorimotor learning has received the vast majority of attention, and it is well established that people adapt to consistent perturbations of auditory feedback [e.g. Houde and Jordan, 1998]. However, the ability of speakers to alter their behavior in response to reward has not been established. Despite this lack of experimental evidence, a number of recent models of infant speech acquisition rely on reinforcement learning [e.g. Howard and Messum, 2011; Warlaumont and Finnegan, 2016].

Two experiments were conducted to examine the ability of external reinforcement to change speech behavior in healthy adults. Each experiment started with a baseline phase where participants read a set of monosyllabic words with the vowel /ɛ/ and their average F1 (μ_{F1}) was measured. A test phase followed, where participants again read words with /ɛ/. Productions with an F1 between μ_{F1}-10 and μ_{F1}-110 were positively rewarded with +10 points. Productions outside this range were negatively rewarded with -10 points. Participants were instructed that their goal was to get the most points possible. A washout phase followed the test phase, where the reward was removed.

In experiment 1, participants heard their own voice while speaking. Positive rewards were accompanied by a chime sound; negative rewards, by spoken productions of the target word, with the vowel replaced with either [ɪ] or [æ]. In Experiment 2, masking noise blocked auditory feedback. Negative rewards were similar to Experiment 1. Positive rewards were the participants own production of the target word from the baseline phase, with F1 shifted by - 60 Hz to the mean of the reward zone.

In both experiments, roughly 2/3 of participants were able to learn from the feedback, lowering their produced F1 so that it fell in the rewarded zone. This change persisted throughout the entire 100-trial washout phase, unlike the return to baseline seen in sensorimotor learning studies. We found that no participants adopted an appropriate explicit strategy beyond "speak[ing] clearly." Together, these results suggest that reward learning in speech can occur without conscious awareness and regardless of the presence sensory information, and may lead to long-lasting changes to speech motor behavior.

References

Houde, J., & Jordan, I. M. (1998). Sensorimotor adaptation in speech production. *Science, 279*, 1213-1216.

Howard, I. S., & Messum, P. (2011). Modeling the development of pronunciation in infant speech acquisition. *Motor Control, 15*(1), 85-117.

Warlaumont, A. S., & Finnegan, M. K. (2016). Learning to produce syllabic speech sounds via reward-modulated neural plasticity. *PloS One, 11*(1), e0145096.

ARTICULATORY INTER-PAUSAL BEHAVIOR

Oxana Rasskazova[1,2], Christine Mooshammer[1], Susanne Fuchs[2]
[1]*Humboldt-Universität zu Berlin, Berlin, German*
[2]*Leibniz - Zentrum Allgemeine Sprachwissenschaft, Berlin, Germany*

During silent pauses speakers plan subsequent utterances on different levels. Phonetically, preparation involves breathing activity and the movement towards the post-pausal segments e.g. (Ramanarayanan et al., 2009). There are only a few studies investigating articulatory motion during pauses (Schaeffler et al., 2008; Ramanarayanan et al., 2009; Scobbie et al., 2011) and their results providing a rather inconsistent picture. It has been shown that some speakers press their tongue against the palate during pauses, others do not move their articulators until the next segment. Independently of this speaker-specific variation, the articulatory preparation for an upcoming segment always starts earlier than the acoustically onset. As for reaction time measures, this fact has been largely ignored for measuring of pause duration, which are generally based on the acoustical measurements.

The present study focuses on the articulatory behavior during silent pauses between two prosodic phrases (IP) in a reading situation. We investigated the differences between acoustically and articulatory onsets, i.e. how far in advance speakers start moving their articulators. Posture variation during the pauses between speakers as well as correlations between movement amplitude and pause duration were analyzed in order to gain further insights into motor planning process during silent pauses.

Eight native German speakers have been recorded in a sound proof cabin by means of EMA (AG 501, Carstens Electronics). The last segment of the first IP were alveolars (/t/, /n/) and bilabial /m/ and the first segment of the second IP was the vowel /a/. Tongue back and lower lip movement data were analyzed for post-pausal onsets, i.e. the opening gesture towards the vowel as well for acoustically onset and offsets of pauses. First results show that acoustically measured pause duration is consistently overestimating planning time because it includes the first post-boundary gesture. They also provide insights in inter-speaker variation of articulatory behavior during pauses.

References

Ramanarayanan, V., Bresch, E., Byrd, D., Goldstein, L., & Narayanan, S. S. (2009). Analysis of pausing behavior in spontaneous speech using real-time magnetic resonance imaging of articulation. *JASA, 126*(5), 160-165.

Schaeffler, S., Scobbie, J. M., & Mennen, I. (2008). An evaluation of inter-speech postures for the study of Language specific articulatory settings. *Proc. of the 8th ISSP Strasbourg*, 121-124.

Scobbie, J.M., Schaeffler, S. & Mennen, I. (2011). Audible aspects of speech preparation. In *Proc. of the ICPhS (Hong Kong)*, 1782-1785.

EFFECTS OF AUDITORY HABITUATION AND FEEDBACK DELAYS ON SPEECH MOTOR LEARNING

Douglas M. Shiller[1,2,3], Takashi Mitsuya[4], Ludo Max[4]
[1] *Université de Montréal, Canada*
[2] *CHU Sainte-Justine Research Centre, Montreal, Canada*
[3] *Centre for Research on Brain, Language and Music, Montreal, Canada*
[4] *University of Washington, Seattle, United States of America*

One potential problem for efficient motor learning arises when movement-related feedback is delayed. It is known that perceiving the consequences of ones own actions with a delay results in the sensory information being processed similar to externally-generated, rather than self-generated, sensory input (Blakemore et al., 2000). In the present study, we follow up on a recent study by Max & Maffett (2015) in which it was found that auditory-motor learning in speech is abolished with feedback delays of 100 ms or more. Here, we extend this work by testing whether, similar to visuomotor adaptation, such detrimental effects can be prevented by first letting subjects habituate to the delay as repeated exposure to two asynchronous events causes the point of subjective simultaneity to shift toward the time lag between those events (Fujisaki et al., 2004).

In the present study, subjects completed an auditory-motor learning task involving a real-time increase in the formant frequencies of vowels produced in simple, monosyllabic words (a manipulation known to induce a compensatory change in vowel production). Simultaneous with the formant shift, auditory feedback was delayed by 75 ms, thereby limiting the degree of adaptation. Prior to the motor learning task, one group of subjects habituated to the 75 ms auditory feedback delay during 30 minutes of speech tasks that also involved monosyllabic words (with no formant shift). Another group of subjects completed the 30 minutes of speaking tasks without exposure to any feedback delay. Contrary to our hypothesis that habituation to delayed feedback would facilitate speech motor adaptation relative to the no-habituation control group, subjects in the two groups were found to adapt similarly to the formant shift. Although further experiments with different forms or amounts of habituation and different feedback delays are warranted, this initial result suggests that the processing of delayed auditory feedback during speech production may be resistant to the mechanisms that underlie habituation to sensory delays in other effector systems.

References

Blakemore, S., Wolpert, D., Frith, C. (2000). Why can't you tickle yourself? *Neuroreport, 11*, R11-6.

Fujisaki, W., Shimojo, S., Kashino, M., Nishida S. (2004). Recalibration of audiovisual simultaneity. *Nature Neuroscience, 7*, 773-8.

Max, L. , and Maffett, D. G. (2015). Feedback delays eliminate auditory-motor learning in speech production. *Neuroscience Letters, 591*, 25-29.

AUDITORY-MOTOR IMPAIRMENT AS AN UNDERLYING BASIS OF HYPERFUNCTIONAL VOICE DISORDERS

Cara E. Stepp[1,2], Elizabeth S. Heller Murray[1], Rosemary A. Lester-Smith[1], Defne Abur[1], Ayoub Daliri[1], J. Pieter Noordzij[1,2], Ashling A. Lupiani[1]
[1] *Boston University, Boston, United States of America*
[2] *Boston University School of Medicine, Boston, United States of America*

Individuals with hyperfunctional voice disorders (HVDs) have dysphonic voices in the presumed absence of neurological dysfunction (Roy, 2008). In two studies, we examined auditory and auditory-motor function in individuals with HVDs. In the first study, we examined the auditory acuity to vocal loudness. In three conditions, 15 individuals with HVDs and 30 controls judged the differential loudness of tokens using a two-alternative forced choice (2AFC) procedure to calculate a just-noticeable-difference (JND; Macmillan & Creelman, 2004). In one condition, speech tokens were provided externally through headphones. In the two other conditions, subjects were asked to estimate their own vocal loudness, with or without the presence of auditory masking noise. No differences were found between the JNDs of control and HVD participants, whether stimuli were a) externally presented (reliance on auditory feedback only), b) autophonically produced without masking noise (reliance on both auditory and somatosensory feedback), or c) autophonically produced with masking noise (reliance on somatosensory feedback). In a second study, we examined auditory acuity to fundamental frequency (f_0) and auditory-motor responses to a f_0 sensorimotor adaptation paradigm. Participants produced sustained vowels over 160 trials in two separate conditions: a) while experiencing gradual increases in the fo of their auditory feedback ("shift-up") and b) under no auditory perturbation ("control"). The shift-up condition consisted of four ordered phases: baseline (no perturbation), ramp (gradual increases in heard f_0), hold (a consistently higher heard f_0), and after-effect (no perturbation). Adaptive responses were defined as the difference in produced fo during control and shift-up conditions and were collected in 9 individuals with HVDs and 9 controls. Using an adaptive 2AFC procedure to estimate JNDs, 6 individuals with HVDs and 5 controls judged the differential pitch of tokens presented externally. There were no differences in pitch JNDs between the two groups; however, adaptive responses were significantly different between groups. Controls showed compensatory adaptive responses, with decreased fo during the ramp and hold phases. Conversely, five participants with HVDs displayed the opposite effect, following the perturbation: when f_0 was experimentally increased, participants further increased their f_0. These results indicate that although auditory acuity to vocal parameters is typical in individuals with HVDs, some have disrupted auditory-motor control, suggesting a potential motor speech disorder.

References

Macmillan, N.A., & Creelman, C.D. (2004). Adaptive Methods for Estimating Empirical Thresholds. *Detection Theory: A User's Guide* (2nd ed., pp. 269-96): Psychology Press.

Roy, N. (2008). Assessment and treatment of musculoskeletal tension in hyperfunctional voice disorders. *International Journal of Speech-Language Pathology, 10*(4), 195- 209.

COORDINATION OF WORD ONSET ARTICULATORY GESTURES IN SWEDISH: ANTICIPATORY CUES TO WORD ACCENTS

Malin Svensson Lundmark
Lund University, Lund, Sweden

Different cues can be used to access the lexical meaning of words. For example, Swedish is known to make use of word stem tones to predict upcoming word endings (Roll et al., 2013). Accent 1 word stems (a high tone in South Swedish) cue less possible continuations (such as suffixes), than Accent 2 (a low tone), which cues more possible word endings, such as productive compounds. Here it will be argued that the stem tones that help listeners predict possible word stem continuations affect the coordination of the articulatory gestures, resulting in additional anticipatory cues being available to the listener even before the tone has been perceived.

Within the framework of Articulatory Phonology, tones have been proposed to constitute articulatory gestures, i.e. tone gestures, which are comparable to consonantal and vocalic gestures (Gao, 2008; Mücke et al., 2012). Lexical tone gestures in Mandarin have even been proposed to compete with consonantal gestures in the onset (Gao, 2008). Such an intercompetitive relationship with the vowel is the source of a phenomenon known as the *c-centre effect*. However, testing the c-centre effect on tone gestures is problematic since measuring its onset is not as easy as measuring the articulators movements in an ElectroMagnetic Articulograph (EMA). For the time being, we are left with measuring F0, which is a fairly representative account of the vocal folds movements, but has a lower degree of testability than measuring movements of the arytenoid cartilage.

Nonetheless, if vocalic, consonantal and tone gestures are coordinated in word onsets, we would expect to find cues of this coordination pattern in the consonantal and vocalic gestures. The present study on 20 South Swedish speakers investigates the c-centre effect in word accents by comparing a consonant-initial Accent 1 word (high stem tone) and an Accent 2 word (low stem tone). Results show that for some speakers the consonantal gesture starts earlier in Accent 2, displaying a greater c-centre effect for the tone gesture with a low tone target. However, not all speakers display this temporal pattern and instead indicate an effect on the spatial coordination pattern. Nonetheless, the tone gestures do appear to affect the articulatory coordination in a different manner in Accent 2 as compared to Accent 1, possibly allowing listeners to anticipate the nature of the upcoming tone before it has been perceived.

References

Gao, M. (2008). Tonal alignment in Mandarin Chinese: *An articulatory phonology account* (Unpublished doctoral dissertation). Yale University, New Haven.

Mücke, D., Nam, H., Hermes, A., & Goldstein, L. (2012). Coupling of tone and constriction gestures in pitch accents. In P. Hoole et al. (Eds.), *Consonant Clusters and Structural Complexity* (pp. 157-176). Munich, Mouton de Gruyter.

Roll, M., Söderström, P., & Horne, M. (2013). Word-stem tones cue suffixes in the brain. *Brain Research, 520*, 116-120

ARTICULATORY CONTROL PARAMETERS OF PHONOLOGICAL CONTRASTS: THE CASE OF CUEWEIGHING FOR DUTCH /ɑ/ - /a/

Hayo Terband, Manou van Montfort, Lydia Bax, Peter Hart, Sapna Sehgal, Laura Smorenburg, Fleur Versteeg, Tom Lentz
Utrecht University, Utrecht, The Netherlands

Speech-language acquisition involves learning the speech sounds of the language at hand as well as which acoustic cues are relevant to differentiate them. For example, the Dutch vowels /ɑ/ and /a/ in the words 'man' (man) and 'maan' (moon) differ both in their spectral properties (F1 and F2 are both higher for /a/) and in duration (longer for /a/). (Adank, van Hout, & Smits, 2004; Pols, Tromp & Plomp, 1973)

Perception experiments indicate that the way in which different perceptual cues are being combined and weighted is language specific, however, there are large individual differences in cue weighting within language groups (Escudero, Benders, & Lipsky, 2009). Furthermore, the different perceptual cues are not entirely independent from a speech acoustics point of view, meaning that one of the cues could play a more prominent role underneath.

The present study combines measurements of perceptual weighing of acoustic cues for the Dutch /ɑ/ - /a/ vowel contrast with measurements of perceptual acuity for spectral and durational differences and measurements of the production of the vowel contrast to unravel (1) what exactly influences how perceptual cues are being weighed, and (2) which cue is actively being manipulated by speakers. 45 young adult speakers of Dutch (age 19-29) participated in the study. The results show (1) a correlation between perceptual acuity for spectral properties and the perceptual weighing of spectral properties and duration and (2) a negative correlation between how participants weigh perceptual cues and how the participants use these cues in production.

References

Adank, P., Van Hout, R., & Smits, R. (2004). An acoustic description of the vowels of Northern and Southern Standard Dutch. *The Journal of the Acoustical society of America, 116*(3), 1729-1738.

Escudero, P., Benders, T., & Lipski, S. C. (2009). Native, non-native and L2 perceptual cue weighting for Dutch vowels: The case of Dutch, German, and Spanish listeners. *Journal of Phonetics, 37*(4), 452- 465.

Pols, L. C., Tromp, H. R., & Plomp, R. (1973). Frequency analysis of Dutch vowels from 50 male speakers. *The journal of the Acoustical Society of America, 53*(4), 1093-1101.

CONGRUENT AERO-TACTILE STIMULI BIAS PERCEPTION OF VOICING CONTINUA

Mark K. Tiede[1], Dolly Goldenberg[1,2], Douglas H. Whalen [1,2,3]
[1]*Haskins Laboratories, New Haven, United States of America*
[2]*Yale University, New Haven, United States of America*
[3]*City University of New York, New York, United States of America*

Previous work has established that orofacial somatosensation can affect speech perception (Ito et al., 2009). Gick & Derrick (2009) have shown that that puffs of air applied to the skin and timed with listening tasks can also bias the perception of voicing by naive listeners: when an air-puff is felt coincident with hearing /ba/ in noise they are more likely to respond “pa” than when no air-puff is presented, presumably because the tactile stimulus is associated with aspiration.

The current study has replicated and extended these findings in several ways. First, air-puff effects are tested on gradations of systematically modulated VOT continua (constructed from /pa/ and /ka/ by removing increasing amounts of aspiration) rather than endpoints masked by noise. Second, responses to bilabial and velar VOT continua are compared to an additional vowel quality control task (systematically shifting formants from /hɪd/ to /hɛd/). In addition, a pretest included with every testing session is used to confirm that participants are consistently able to detect a puff when presented, and (by moving the skin point of contact away from the airflow) are consistently unable to detect a presented puff through any uncontrolled sensory modality.

In these experiments air-puff stimuli are produced using an air compressor with a programmable solenoid valve. Custom software is used to synchronize the opening of the valve with playback of the auditory stimulus (presented through headphones), and to record the listeners forced choice response. Puff presentation (present vs. absent) is randomized. Puff durations are chosen to be the mean VOT of pre-recorded endpoint exemplars (90 ms), and are delivered through a tube placed 5 cm from the dorsal surface of the hand between the right thumb and forefinger. Within a session each participant performs 240 separate judgments per continuum type, with 15 per condition at each continuum step. Each of the three continua (bilabial, velar, vowel) has been tested with at least 20 participants.

GLMMs with random intercepts by listener predicting the voicing response for each continuum given fixed effects of continuum step and presence/absence of air-puffs show that puffs significantly shift responses towards voicelessness (bilabial $p < .002$, velar $p < .011$), with the largest effects in the middle (most ambiguous) part of each continuum. Conversely no effect of puff is observed on responses for the control (vowel quality) continuum. These results provide support for a tactile component of multimodal integration in speech perception, suggesting that such cues are weighted by the salience of aspiration in producing a given contrast.

References

Gick, B., & Derrick, D. (2009). Aero-tactile integration in speech perception. *Nature, 462*(7272), 502-504.

Ito, T., Tiede, M., & Ostry, D. J. (2009). Somatosensory function in speech perception. *Proceedings of the National Academy of Sciences, 106*(4), 1245-1248.

ARTICULATORY DIFFERENCES BETWEEN GLIDES AND VOWELS

Sam Tilsen, Dan Cameron Burgdorf
Cornell University, Ithaca, United States of America

Glides bear similarities to both consonants and vowels, but dont quite fit into traditional conceptions of either, and their phonetic properties are not well understood. Limited prior work has suggested that glides are realized with a greater degree of constriction and shorter duration than vowels, but articulatory studies of glides are rare. To determine how glides differ from vowels gesturally, both with regard to the parameters of the gesture (such as degree of constriction) and the nature of the difference (continuous or categorical), an electromagnetic articulography (EMA) study was conducted. Participants were asked to imitate stimuli that varied in duration and intensity from a vowel-like extreme to a glide-like extreme. Contrary to the common account, glides were not found to be more constricted than vowels. Reliable variation was found in duration, intergestural timing, and precision. The relationship between glides and vowels is much more nuanced than merely a difference in degree of constriction.

References

Browman, C.P., & Goldstein, L. (1992). Articulatory phonology: An overview. *Haskins Laboratories Status Report on Speech Research*.

Levi, S. V. (2008). Phonemic vs. derived glides. *Lingua, 118*(12), 1956-1978.

Nevins, A., & Chitoran, I. (2008). Phonological representations and the variable patterning of glides. *Lingua, 118*(12), 1979-1997.

Poster session II

INFLUENCE OF VOICE FOCUS ON TONGUE MOVEMENT IN SPEECH

Gillian de Boer[1], Tim Bressmann[1], Viviane Cristina Castro Marino[2],
Eliana Maria Gradim Fabron[2], Larissa Cristina Berti[2]
[1] *University of Toronto, Toronto, Canada*
[2] *Universidade Estadual Paulista, Marilïa, Brazil*

The present study evaluated global aspects of lingual movement during sentence production with backward and forward voice focus. Nine female participants read a sentence with a variety of consonants in a normal condition and with backward and forward voice focus. Midsagittal tongue movement was recorded with ultrasound. Tongue height over time at an anterior, a central and a posterior measurement angle was measured. The outcome measures were speech rate, cumulative distance travelled and average movement speed of the tongue. There were no differences in speech rate between the different conditions. The cumulative distance travelled by the tongue and the average speed indicated that the posterior tongue traveled a smaller cumulative distance and at a slower speed in the forward focus condition. The central tongue moved a larger cumulative distance and at a higher speed in the backward focus condition. The study offers first insights how tongue movement is affected by different voice focus settings and illustrates the plasticity of tongue movement in speech.

Further reading

De Boer, G., & Bressmann, T. (2016). Influence of voice focus on oral-nasal balance in speech. *Journal of Voice, 30*(6), 705-710.

De Boer, G., Marino, V., Berti, L., Fabron, E., & Bressmann, T. (2016). Influence of Voice Focus on Oral-Nasal Balance in Speakers of Brazilian Portuguese. *Folia Phoniatrica Logopaedica, 68*, 52-158.

Bressmann, T., de Boer, G., Marino, V., Fabron, E., & Berti, L. (2017). Influence of voice focus on tongue movement in speech. *Clinical Linguistics and Phonetics, 31*(3), 212-221.

L1 INTONATIONAL CATEGORIES AS "PERCEPTUAL ATTRACTORS" DURING L2 IMITATION

Leonardo Lancia[1], Caterina Petrone[2], Cristel Portes[2]
[1]*LPP, Paris, France*
[2]*LPL & AMU, Aix-en-Provence, France*

When adult speakers learn a foreign language (L2), native (L1) phonological representations can work as "perceptual attractors" constraining the linguistic interpretation of the acoustic signal as it builds up during the perception of L2 speech (Best, 1995). Such representations play an important role in tuning imitative behavior. For instance, English speakers are less capable to reproduce stops with reduced VOT, possibly to avoid phonological ambiguity with the voiced counterparts (Nielsen, 2011). Furthermore, with the passing of time, the auditory details of the speech sound fade in working memory and they are replaced by the properties defining the phonological category instantiated by that sound (e.g., Baddley, 2002).

Here, we investigate how L1 perceptual representations affect imitation of L2 intonation. We focus on the imitation of (Northern Standard) German yes/no questions by speakers of Neapolitan and Lecce varieties of Italian. In German, questions are marked by a terminal f0 rise with a steep slope. In Italian, the specific pattern varies across regional varieties. Lecce Italian questions are realized with a terminal rise but, differently from German, the terminal rise has a shallower slope. In Neapolitan, questions show a different phonological configuration (a f0 rise-fall). If perceptual representations behave as attractors, Lecce speakers will perceive German intonational contours as variants of their native dialect contours. Hence, they will be less accurate than Neapolitans in the imitation of German f0 rises. The reproduction of the phonetic details of the f0 rises will be worse when imitation is temporally delayed and/or when memory load is increased.

18 Lecce and 18 Neapolitan Italian speakers and with no knowledge of German were instructed to imitate the pronunciation of 20 questions produced in "a foreign language". Participants started to imitate (i) immediately after the end of the question; (ii) after a silent pause of 4 s. at the end of the question; (iii) after a filled pause of 4 s. at the end of the question (i.e., the answer to it). Degree of imitation was assessed through time-series analysis of the f0 contours, by integrating cross-recurrence-analysis and Wavelet-based Functional Mixed Models.

Preliminary results confirm that Lecce speakers are less accurate than Neapolitans in imitating German rises, suggesting that L2 terminal rises are "assimilated" to L1 rises. No effects have been found on memory, which might be explained by task-specific reasons (e.g., easy of the task).

References

Baddley, A. (2002). Is working memory still working? *European Psychologist, 7*(2), 85-97.

Best, C.T. (1995). A direct realist perspective on cross-language speech perception. In W. Strange (ed.), *Speech perception and Linguistic Experience: Theoretical and Methodological Issues*, 167-200, York. Timonium.

Nielsen, K. (2011). Specificity and abstractness of VOT imitation. *Journal of Phonetics, 39*(2), 132-142.

AUDITORY SENSORY PREDICTION DURING SPEAKING AND LISTENING

Douglas M. Shiller[1,2,3], Marc Sato[4,5]
[1]*Université de Montréal, Canada*
[2]*CHU Sainte-Justine Research Centre, Montreal, Canada*
[3]*Centre for Research on Brain, Language and Music, Montreal, Canada*
[4]*Aix-Marseille Université & CNRS, Aix-Marseille, France*
[5]*Brain and Language Research Institute, Aix-en-Provence, France.*

Auditory sensory prediction plays a central role in speech motor planning and control. One approach to investigating this predictive process is through the comparison of auditory cortical responses to auditory feedback during active speech vs. passive (offline) listening to the same acoustic speech signals. Neural responses during active speech production are typically suppressed, which is presumed to reflect a subtraction of the motor-sensory prediction from auditory feedback. In the present EEG study, we further explored the link between online (feedback) and offline (passive listening) speech processing by measuring auditory evoked responses in combination with two different manipulations that alter a talkers ability to accurately predict the sensory consequences of speech actions. One manipulation involved a real-time change in auditory feedback (F1 frequency) during vowel production, creating a mismatch between sensory prediction and auditory feedback, in order to reduce the accuracy of sensory prediction. The second manipulation involved presenting to subjects a stable auditory vowel target (corresponding to their own typical speech output) prior to each vowel production, which was predicted to enhance the accuracy of subjects sensory prediction. In response to the real-time auditory feedback manipulation, subjects showed a compensatory decrease in vowel F1 output relative to baseline, with no difference between the visual and auditory cueing conditions and no modulation of N1/P2 auditory evoked responses in relation to the compensatory changes. Differential effects of cueing modality (target word presented visually vs. auditorily prior to each utterance) on the auditory evoked responses were also observed. First, under the visually-cued condition, a speaking-induced suppression was found, consisting of a reduced N1/P2 amplitude during speaking compared to listening that likely reflects a motor-to-sensory prediction. Second, under the auditory-cued condition, a classical auditory adaptation effect (repetition suppression) was observed in the N1/P2 amplitude during the passive listening task but not during the active speaking task. This last result suggests that endogenous and exogenous auditory predictive processes during speaking and listening are functionally decoupled. Interestingly, however, the degree of N1/P2 amplitude reduction associated with speaking-induced suppression and auditory adaptation was found to be very similar, suggesting that these predictive processes may not be entirely unrelated.

References

Niziolek, C. A., Nagarajan, S. S., & Houde, J. F. (2013). What does motor efference copy represent? Evidence from speech production. *The Journal of Neuroscience, 33*, 16110-16116.

Wang, J., Mathalon, D.H., Roach, B.J., Reilly, J., Keedy, S.K., Sweeney, J.A., Ford, J.M. (2014) Action planning and predictive coding when speaking. *NeuroImage, 91*, 91-98.

AN EXPLORATORY STUDY INTO DIFFERENT TYPES OF ARTICULATION IN CLARINET PLAYING

Anneke Slis, Kornel Wolak, Aravind Namasivayam, Pascal van Lieshout
University of Toronto, Toronto, Canada

Spoken language evolution has been made possible by the specific shape of the human vocal tract. This shape likely also led to the emergence of certain musical instruments and its playing techniques. The purpose of this study is to provide a baseline for articulatory measures in clarinet playing, based on similar measures used for speech research. More specifically, we explored the kinematic features of three articulation types used in clarinet playing - single (ST), double (DT) and lateral tonguing (LT).

Movements of the tongue tip and tongue dorsum were recorded with Electromagnetic Articulography (AG501). A professional clarinetist played a diatonic scale from concert-d3 to a concert-Bb6 and back with single notes and 4-note repetitions. Three types of articulation were investigated: ST ("saying" /t/), DT ("saying" /tk/) and LT (tongue tip moves right to left and back). Each articulation was performed with three different vocal tract configurations resembling those used when producing the vowels /i/, /a/ and /y:/. Average values of vertical amplitude and cyclic Spatio-Temporal Index (cSTI) of the tongue tip and dorsum are reported for ST and DT. The same kinematic measures for LT were retrieved from left-to-right movements.

For ST and LT, the tongue tip showed the most stable performance as indexed by Csti; values for DT were much higher. Tongue tip values also showed higher values in DT than the tongue dorsum. It is speculated that ST is the most stable articulation as it also can be considered the most practiced articulation in clarinet playing. In addition, the tongue tip movement likely is less variable as the mouth piece stabilizes the jaw, resulting in fewer degrees of freedom. The higher tongue tip cSTI values in DT might have been caused by mechanical constraints between the tongue tip and dorsum; the latter being active in DT as well.

The amplitudes of the tongue dorsum were found to be larger than those of the tongue tip movements for DT. This likely results from the fact that in higher registers the tongue dorsum must move back and downwards to allow the air pressure to increase for initiating a tone on the clarinet (Wheeler, 2010). This amplitude increase depended on the type of vowel associated with a specific vocal tract configuration: /a/ and /i/ resulted in larger amplitude movements of the tongue dorsum than /y:/, resembling amplitude movement ranges found in speech (Slis and Van Lieshout, 2013). These data indicate that EMA is a promising tool to investigate articulatory movements in clarinet playing.

References

Wheeler, R. (2010). "*X-Ray Video of Dancing Tongue in Clarinetists Mouth*", http://blog.davidhthomas.net/2011/08/x-ray-video-of-dancing-tongue-in-clarinetists-mouth-not-what-you-expect/

Slis, A., & Van Lieshout, P. (2013). The effect of phonetic context on speech movements in repetitive speech. *The Journal of the Acoustical Society of America, 134*(6), 4496-4507.

EFFECTS OF SPEECH INTENSITY ON ACOUSTIC AND KINEMATIC ARTICULATORY WORKING SPACE

Christopher Dromey[1], Panika Palmer[1], Jason Whitfield[2]
[1]*Brigham Young University, Provo, United States of America*
[2]*Bowling Green State University, Bowling Green, United States of America*

Acoustic measures calculated from the formant frequencies of corner vowels, such as the vowel space areas of triangles or quadrilaterals, have often been used as a window into articulatory function. Other measures, such as the formant centralization ratio (Sapir, Ramig, Spielman, & Fox, 2010), have been suggested as more sensitive alternatives. The purpose of this study was to compare changes in acoustic and kinematic measures of articulation across soft, comfortable, and loud speech conditions using both corner vowel-based and sentencelevel metrics. There were 19 participants, 9 male and 10 female, with ages ranging from 20 to 34 years. Each participant had electromagnetic sensors attached to their tongue, jaw, and lips. It was anticipated that the acoustic measures would parallel the changes observed in the kinematic measures as articulation changed across intensity levels. Vowel space area (VSA) and vowel articulation index (VAI) were computed from the three corner vowels, /ɑ, i, u/. Articulatory-acoustic vowel space (AAVS), a sentence-level acoustic measure (Whitfield & Goberman, 2014), was computed from the continuous formant histories of all voiced segments in a sentence. Kinematic vowel space area (KVSA), kinematic vowel articulation index (KVAI), and articulatory kinematic vowel space (AKVS) were the kinematic equivalents of the acoustic measures, and were newly developed for the present study. Stroke metrics based on the speed history of the lingual movements were also used to reveal average kinematic features of the articulatory gestures in each participants speech. The data revealed that the isolated acoustic and kinematic measures that were based on the segmented corner vowels (VSA, VAI, KVSA, KVAI) did not change significantly across the three intensities. The sentence-level continuous measures of articulatory working space (AAVS and AKVS) increased with speech intensity. The sentence-level kinematic metrics based on the lingual speed history also changed significantly with intensity, including increased onset speed, peak speed, mean speed, and distance of the articulatory strokes. Stroke duration decreased as speech intensity increased, while hull area increased. These findings suggest that both acoustic and kinematic measures based on segmented corner vowels are not as sensitive to loudness-related changes in articulation as the more continuous sentence-level measures. Metrics that include the dynamic contributions of both consonant and vowel articulation may capture relevant changes that are not detected by measures based on vowel midpoints.

References

Sapir, S., Ramig, L. O., Spielman, J. L., & Fox, C. (2010). Formant centralization ratio: A proposal for a new acoustic measure of dysarthric speech. *Journal of Speech, Language, and Hearing Research, 53*(1), 114-125. doi:10.1044/1092-4388(2009/08- 0184)

Whitfield, J. A., & Goberman, A. M. (2014). Articulatory-acoustic vowel space: Application to clear speech in individuals with Parkinson's disease. *Journal of Communication Disorders, 51*, 19-28. doi:10.1016/j.jcomdis.2014.06.005

DYNAMICS AND KINEMATICS OF REPETITIVE SPEECH MOVEMENTS

Stephan R. Kuberski, Adamantios I. Gafos
University of Potsdam, Potsdam, Germany

standard model of isolated speech movements is thought to be a linear second order system (Saltzman and Munhall, 1989). Several variations to this model have been proposed which claim to render isolated as well as sequences of speech movements more accurately. There are also proposals from other areas of motor control which aim to render both non-repetitive and repetitive sequences of not specifically speech movements. We used the Harvard/Haskins database of regularly-timed speech (Patel et al., 1999) to extract kinematic relations of jaw and lower lip movements in opening and closing /b, m/ gestures. These relations are then compared with the via-simulation predicted relations from linear and nonlinear models of the speech gesture as well as autonomous and nonautonomous two-dimensional proposed models. In simulations, topological differences with regard to the kind and the number of the model-specific attractors (fixed points, limit cycles) are considered. It is shown that, for the experimental data in this work, none of these models are able to completely account for the data. Relative time to peak velocity (RTTP) values of closing movements tend to be higher than values predicted by any fixed point model. Those models which simulation results indicate are able to generate higher values of RTTP exhibit other incompatible kinematic relations (e.g., strongly nonlinear peak velocity vs. movement amplitude relation). We conclude with a discussion of the extent to which the task at hand uniquely determines the dynamical regime (repeated isolated gestures vs. limit cycle) underlying the observed performance. If time permits, we will review results from our own work with acquiring similar data, under more controlled conditions (specifically speech rate controlled), in an attempt to address the nature of the dynamical regime underlying the movements.

References

Saltzman, Elliot L. and Munhall, K. G. (1989). A dynamical approach to gestural patterning in speech production. *Ecological psychology, 1*(4), 333-382. doi: 10.1207/s15326969eco0104_2.

Patel, Aniruddh D., Löfqvist, A., & Naito, W. (1999). The acoustics and kinematics of regularly timed speech: a database and method for the study of the P-center problem. In: *Proceedings of the 14th International Congress of Phonetic Sciences.* (San Francisco). Vol. 1, pp. 405-408.

DUAL-TASK INTERFERENCE ON AUTOMATIC SPEECH AS A FUNCTION OF CONCURRENT TASK DEMAND

Marina Laganaro[1], Maryll Fournet[1,2], Sabina Catalano Chiuvé[2], Ursula Lopez[2], Michaela Pernon[1,3]
[1]*University of Geneva, Geneva, Switzerland*
[2]*Geneva University Hospitals, Geneva, Switzerland*
Hôpital Lariboisière, Paris, France

It is currently admitted that running a concurrent task interferes with speaking and in particular with lexical selection (Ferreira & Pashler, 2002) during language planning. Motor speech planning is thought to be much more automatic than linguistic planning and therefore relatively unaffected in dual tasks situations (Garrod & Pickering, 2006). It has however also been shown that speech rate is reduced under dual-task demand, especially in older speakers (Kemper, Herman & Nartowicz, 2005), but also in younger speakers when they give priority to the second (non speech) task (Kemper et al., 2011). It is possible however that speech rate is slowed down strategically to allow more time to linguistic planning in tasks where participants are requested to answer questions in dual-task conditions (Kemper et al, 2005; 2011), i.e. to plan a message from concept to speech. In the present study we investigate the effects of dual-task demands on automatic speech involving a reduced linguistic planning demand in adults. 66 French-speaking participants aged between 22 and 80 year-old performed looped days of the week recitation either in isolation or under dual-task conditions, while running a concurrent visuo-attentional task. We further manipulated the executive demand of the concurrent task, which was either a simple manual reaction to visual stimuli or a go no-go reaction task. Speech rate was computed on a 20 sec period and articulation rate was calculated on the "mercredi-jeudi" ("wednesday-thursday") sequences. Both speech rate and articulation rate were significantly slowed down relative to speaking in isolation when speakers run a concurrent go no-go task, without interaction with age. Speech was not significantly affected when the participants were engaged in a simple reaction task. Hence, the observation that speech is affected by a demanding dual-task suggests that motor speech planning is not completely automatic but involves some attentional resources; however, the amount of attentional control is probably low given that speech resists to a simple concurrent task.

References

Ferreira V.S., Pashler H. (2002) Central Bottleneck Influences on the Processing Stages of Word Production. *J Exp Psychol Learn Mem Cogn, 28*, 1187-1199.

Garrod S., Pickering M.J. (2006) Automaticity of language production in monologue and dialogue. *Automaticity and Control in Language Processing*. Psychology Press. Hove; . pp. 1-20.

Kemper, S., Herman, R. E., & Nartowicz, J. (2005). Different effects of dual task demands on the speech of young and older adults. *Neuropsychol Dev Cogn B Aging Neuropsychol Cogn, 2*, 12(4): 340-358.

Kemper, S., Schmalzried, R., Herman, R. & Mohankumar, D (2011) The Effects of Varying Task Priorities on Language Production by Young and Older Adults, *Experimental Aging Research, 37*(2), 198-219

TWO TYPES OF GESTURAL CONTROL IN MANDARIN NASAL CODAS

Yanyu Long
Cornell University, Ithaca, United States of America

This study used EMA to investigate the articulatory realization of two speech activities of Mandarin nasal codas [n] and [ŋ], i.e. oral constriction reduction/deletion before a vowel (/dæn/+/a:/ → [dæ̃:.a:]) and place coarticulation/assimilation before a stop (/dæn/+/pa:/ → [dæm.pa:]). Evidence shows two types of gestural control regimes: In the pre-vowel context, the gestures of the vowel and the first syllable are competitively selected. The degree of their gestural overlap is larger and more variant when the coda is [n] rather than [ŋ]. In the pre-stop context, the gestures of the stop and the first syllable are coordinatively selected.

Three native Mandarin speakers read disyllabic words [da:.a:] vs. [daN.a:] (N represents [n]/[ŋ]), [da:.ba:] vs. [daN.ba:], in a carrier sentence at three speeds. The results show that before [a:], [n] shows reduction in oral constriction, the degree of which increases as speech rate increases. [ŋ] shows small reduction across all speech rates. This indicates that the gestural overlap between [n] and the following [a:] increases in faster speech, while that between [ŋ] and [a:] is small and has minimal variation. In [daN.ba:], all speakers showed anticipation of the labial gesture of [b], with no difference between [dn.ba:] and [dan.ba:]. The relative timing of the labial gesture and the preceding syllable is relatively stable across speech rates. Speakers vary in whether the tongue gestures of the nasal codas are deleted or not during the anticipated labial gesture.

The findings suggest that the two processes have different mechanisms in motor control. Before [a:], [ŋ] shows a stably small overlap with the following vowel, suggesting a prototypical competitive relationship between the two, i.e., {a1ŋ}{a2} (Tilsen, 2016). It is similar for [n] in slower speech. [n] allows more overlap in faster speech due to time pressure. The difference between [n] and [ŋ] can be explained by the perceptual recoverability of the nasal identity in [æn.a:] vs [aŋ.a:]: [n] can be easily identified using the preceding (front) vowel even if [n] is largely distorted/reduced (Mou, 2006), while [aŋ] may be perceptual confusable with [a:] without enough oral constriction. Therefore, perceptual pressure impedes more gestural overlap in [ŋ] but not [n]. In [daN.ba:], the labial gesture and the first syllable form a close relationship with large overlap and stable relative timing. This can be modeled by coordinative control, in which gestures are selected together, and their relative timing of activation is precisely controlled, i.e. {a1(N)b}.

References

Mou, X. (2006). *Nasal codas in Standard Chinese: a study in the framework of the distinctive feature theory*. Doctoral dissertation, MIT.

Tilsen, S. (2016). Selection and coordination: the articulatory basis for the emergence of phonological structure. *Journal of Phonetics, 55*, 53-77.

A HIERARCHICAL STATE FEEDBACK CONTROL MODEL FOR SPEECH SIMULATES TASK-SPECIFIC RESPONSES TO AUDITORY AND ARTICULATORY PERTURBATIONS

Benjamin Parrell[1], Vikram Ramanarayanan[2], Sri Nagarajan[3], John Houde[3]
[1] *University of Delaware, Newark, United States of America*
[2] *Educational Testing Service R&D, San Francisco, United States of America*
[3] *University of California, San Francisco, United States of America*

We present a multi-level state feedback based model of speech motor control. As in other state feedback models, the controller governing behavior operates on an internal estimate of the current state of the vocal tract that combines sensory feedback with an internal prediction of the vocal tract state. In our model, this estimate is generated using an extended Kalman filter (Ramanarayanan et al., 2016). In speech, at least some the sensory feedback received by the system is not in the same domain as the output of the controller-feedback is received about speech acoustics, but motor commands must move the speech articulators. We resolve this issue by learning the non-linear mapping between sensory errors and state corrections from a set of simulated vocal tract movements using locally weighted projection regression.

In the current model, a hierarchical structure separates the control of high-level tasks from low-level control of the speech articulators. Currently, we assume tasks to be articulatory gestures, but other tasks such as desired vowel formants can be implemented in the same framework. Importantly, a hierarchical feedback controller allows for redundant control (Todorov et al., 2005), as each task may be completed with multiple configurations of the speech articulators. Such motor equivalence has been demonstrated for speech - when a downward external force is applied to the jaw during production of a /b/, the upper and lower lips compensate for the lower jaw position (Kelso et al., 1984). This response is task-specific and complete, such that bilabial closure is achieved. This contrasts with acoustic perturbations of vowels, where the response is typically a small fraction of the perturbation. The reasons for this incomplete compensation are currently unknown.

Simulations with our model are able to qualitatively reproduce task-specific response to jaw perturbations, with lower and upper lip movements compensating for a lower jaw position only for /b/, but not /z/. Similarly, we are able to reproduce the partial response to acoustic perturbations seen experimentally. However, our results indicate that while compensation is incomplete in the *acoustic* domain, we see complete compensation in the *task* domain. Interestingly, perturbation-induced differences in the estimated position of the *non-relevant* tasks is not corrected for, suggesting that our control scheme agrees with predictions of the minimal interventional principle.

References

Ramanarayanan, V., Parrell, B., et al. (2016). A new model of speech motor control based on task dynamics and state feedback. In I*nterspeech 2016*.

Todorov, E., et al. (2005). From task parameters to motor synergies: A hierarchical framework for approximately-optimal control of redundant manipulators. *Journal of Robotic Systems, 22*(11), 691-710

Kelso, J., et al. (1984). Functionally specific articulatory cooperation following jaw perturbations during speech: Evidence for coordinative structures. *Journal of Experimental Psychology. Human Perception and Performance, 10*(6), 812-32.

TOWARD A DIAGNOSTIC TOOL FOR ASSESSING ARTICULATORY INVOLVEMENT IN ALS

Panying Rong[1], Brian Richburg[2], Jordan R. Green[2]
[1]*University of Kansas, Lawrence, United States of America*
[2]*MGH Institute of Health Professions, Boston, United States of Americva*

The goal of this research is to develop a clinically feasible assessment tool for detecting and monitoring speech motor involvement in Amyotrophic lateral sclerosis (ALS) and other neurologic disorders affecting speech. Recent findings from research on bulbar motor involvement in ALS suggest that across all speech subsystems, articulatory changes occur relatively early and have a significant impact on speech. Among various articulatory features, lip kinematics is a sensitive articulatory indicator of early bulbar motor involvement (Rong, Yunusova, Wang, & Green, 2015; Rong et al., 2016; Yunusova et al., 2010). These findings have motivated our current work on the development of low-cost and automated approaches for analyzing impaired lip movements.

We examined the accuracy of an automated analysis of prerecorded lip movements during a diadochokinetic (DDK) task assessing the speech of healthy controls and persons with bulbar motor involvement in ALS. The algorithm extracts 23 lip movement features including traditional measures of articulatory kinematics (e.g., displacement, peak velocity) as well as novel measures that captured systematic (time-dependent) and time-independent across-trial variation. The automatically extracted features were screened for their sensitivity to differences between the healthy controls and participants with ALS. We also examined the differences in the strength of associations between these features and speech intelligibility.

One hundred and nineteen DDK lip movement recordings from 33 subjects with ALS and 17 healthy subjects were analyzed. Our preliminary findings suggested that the accuracy of the algorithmic approach was acceptable for most measures within an error range of 0.57%. The across group comparisons revealed that among the temporal features, frequency and temporal variability were most affected (Cohens D = 1.54, 1.18, respectively), which predicted the impending speech intelligibility loss. Several spatial features, including spatial variability and peak velocity, showed non-monotonic changes with bulbar disease progression, which is consistent with prior findings (Yunusova et al., 2010).

The current findings provide preliminary support for the algorithmic approach to quantifying DDK features predictive of bulbar motor and speech decline. Because of the clinical feasibility of DDK task, we will discuss the potential clinical applications of this algorithmic approach.

References

Rong, P., Yunusova, Y., Wang, J., & Green, J. R. (2015). Predicting Early Bulbar Decline in Amyotrophic Lateral Sclerosis: A Speech Subsystem Approach. *Behav Neurol, 2015*, 183027. doi:10.1155/2015/183027

Rong, P. Y., Yunusova, Y., Wang, J., Zinman, L., Pattee, G. L., Berry, J. D., . . . Green, J. R. (2016). Predicting Speech Intelligibility Decline in Amyotrophic Lateral Sclerosis Based on the Deterioration of Individual Speech Subsystems. *Plos One, 11*(5). doi:ARTN e0154971/10.1371/journal.pone.0154971

Yunusova, Y., Green, J. R., Lindstrom, M. J., Ball, L. J., Pattee, G. L., & Zinman, L. (2010). Kinematics of disease progression in bulbar ALS. *J Commun Disord, 43*(1), 6-20. doi:10.1016/j.jcomdis.2009.07.003

ASSESSING MECHANISMS OF POSTURAL CONTROL OF SPEECH WITH A BIOMECHANICAL TONGUE MODEL

Andrew Szabados[1], Pascal Perrier[2,3], Takayuki Ito[3,4]
[1] *University of Grenoble Alpes, Grenoble France*
[2] *GIPSA-lab, Grenoble, France*
[3] *Grenoble INP, Grenoble, France*
[4] *CNRS, Grenoble. France*

An open question in speech motor control is the degree to which postural control of the tongue is accomplished through passive mechanisms - such as the mechanical and elastic properties of the tongue itself - reflexes - such as the stretch reflex - or active control mechanisms. Previously, Perrier et al. (2003) have developed a biomechanical simulation of the tongue which models the movement and configurations of the tongue resulting from muscle forces calculated using Feldmans lambda model (Feldman and Levin, 1995) which implements the stretch reflex. In order to make progress in answering the question of the mechanisms of postural control, we modify this biomechanical simulation of the tongue to accurately model a specific external perturbation, namely one in which the tongue is pulled anteriorly using a specific force profile exerted on the tongue body using a force effector attached to the superior part of the tongue blade. This type of perturbation was chosen as the target of our modeling efforts since it is performed on normal subjects in a separate study which provides us with physical parameters against which we can validate our models performance and ensure biologically realistic results (in (Ito et al., 2009) a similar methodology to that used here is described, although in our study the perturbation is applied directly to the tongue). Once this model is tuned to accurately represent this perturbation, we present an analysis of a preliminary set of modeled perturbations and then describe an experimental paradigm for comparing the models predictions with future real data in order to evaluate different hypotheses regarding the importance of passive, reflex, and active strategies in postural control of the tongue.

References

Perrier, P., Payan, Y., Zandipour, M., & Perkell, J. (2003). Influences of tongue biomechanics on speech movements during the production of velar stop consonants: A modeling study. *The Journal of the Acoustical Society of America, 114*(3), 1582-1599.

Feldman, A. G., & Levin, M. F. (1995). The origin and use of positional frames of reference in motor control. *Behavioral and Brain Sciences, 18*(04), 723-744.

Ito, T., Tiede, M., & Ostry, D. J. (2009). Somatosensory function in speech perception. *Proceedings of the National Academy of Sciences, 106*(4), 1245-1248.

AUTOMATIC SPEECH RECOGNITION (ASR) SOFTWARE EFFICACY FOR CONGENITALLY BLIND VS. SIGHTED SPEECH

Pauline Veenstra[1], Martijn Wieling[1], Chris Horst[2]
[1]*University of Groningen, Groningen, The Netherlands*
[2]*Koninklijke Visio, Huizen, The Netherlands*

Visual and acoustic input is important in language acquisition (Kuhl & Meltzoff, 1992). Recent studies of Ménard et al. (2013) have shown that the articulation of Canadian French speakers who are congenitally blind (CB) differs from articulation of non-blind (NB) speakers. Lip movements of CB speakers are less strong than those of NB speakers, which they compensate by using stronger tongue movements (Menard, et al., 2009, 2013). However, this compensation strategy is not sufficient, as the acoustic difference between certain vowel pairs is more pronounced in NB speech. This suggests a clear link between visual input during language acquisition and the resulting pronunciation. The effect of this lack of vowel distinction is minimal for the human ear. However, the performance of ASR systems relies on the acoustic model, which is based on the relation between the acoustic speech signal and individual sounds, which in turn is based on speech of NB speakers. This means that if Dutch CB speakers indeed have a tighter formant space in their running speech, ASR systems are expected to have more trouble with recognition of their speech.

We investigated acoustic (ASR, formant measurements) properties of CB and NB Dutch speakers for read and spontaneous speech, and present the results on the efficacy of ASR. Audio samples of read speech and spontaneous speech were fed through various ASR systems. Levenshtein-distances were computed and the data was analysed using mixed-effects regression. For read speech, ASR systems performed significantly better for CB (n=10) speakers ($p<0.001$). For spontaneous speech no significant difference between CB and NB (n=11) was found ($p=0.84$).

Preliminary analysis of the auditory data indicates that ASR software performs better for read CB speech than for read NB speech. However, braille reading is slower than print reading; this is reflected in the results. As CB speakers read more slowly they have more time to pronounce the words carefully, which results in clearer speech which is better recognized by ASR systems. CB speakers however spoke at least as fast as NB speakers in the spontaneous speech condition. Analysis shows no ASR performance difference between CB and NB spontaneous speech. Further testing with more comparable semi-spontaneous speech is necessary to fully answer our research question.

References

Ménard, L., Dupont, S., Baum, S.R., & Aubin, J. (2009). Production and perception of French vowels by congenitally blind adults and sighted adults. *The Journal of the Acoustical Society of America, 126*(3), 1406-14.

Ménard, L., Toupin, C., Baum, S.R., Drouin, S., Aubin, J., & Tiede, M. (2013). Acoustic and articulatory analysis of French vowels produced by congenitally blind adults and sighted adults. *The Journal of the Acoustical Society of America, 134*(4), 2975-87.

Kuhl, P.K., & Meltzoff, A.N. (1982). The bimodal perception of speech in infancy. *Science, 218*(4577), 1138-1141.

SPEECH MOTOR CONTROL IN 4YEAROLD CHILDREN VS. ADULTS: ANTICIPATION AS INDEX OF SPEECH MOTOR CONTROL MATURITY

Guillaume Barbier[1], Pascal Perrier[2,3], Lucie Ménard[3,4], Yohan Payan[3], Mark K. Tiede[4], Joseph S. Perkell[5]

[1]*GIPSA-lab & Grenoble University, Grenoble, France*

[2]*Université du Québec à Montréal, Montréal, Québec, Canada*

[3]*Univ. Grenoble Alpes & CNRS, TIMC-IMAG laboratory, Grenoble, France*

[4]*Haskins Laboratories, New Haven, United States of America*

[5]*Boston University, Boston, United States of America*

Introduction Learning to speak involves control of the oro-facial articulators, as well as the construction of relationships between motor commands and auditory and somatosensory sensations. The main goal of this study is to further investigate the hypothesis that differences in speech production between children and adults can be explained on the basis of speech motor control maturity. With this aim, we have designed a speech production study focused on two indices: token-to-token variability in the production of isolated vowels, and extra-syllabic anticipatory coarticulation within V1-C-V2 sequences. Token-to-token variability reflects the maturation of speech motor control in terms of motor patterns for the production of a given target speech sound. In line with theories of optimal motor control, anticipatory coarticulation is assumed to be based on the use of internal models, i.e. sensorimotor representations of speech sounds, and the amplitude of anticipatory coarticulation is presumed to reflect the increasing maturation of these sensorimotor representations as speech develops. Our hypothesis is that the neural representations of the speech motor systems of four-year-old children are immature, particularly in their inability to account for the appropriate variability compatible with correct perception of the target sound, leading to a lack of effectiveness in anticipating forthcoming gestures.

Methods & Results Acoustic and articulatory data have been recorded using ultrasound tongue imaging, corrected for head movement with the HOCUS system. Acoustic data from 20 children and 10 adults have been analyzed. Ultrasound data have been analyzed from a subset of these participants: six children and two adults. In agreement with previous studies, token-to-token variability was greater in children than in adults. Strong anticipation of V2 in V1 realization has been found for all adults. In most cases, anticipation has not been observed in children, and when observed, it was of smaller amplitude than in adults. More specifically, only five children among the 20 studied showed a small amount of anticipation, mainly along the antero-posterior dimension.

Discussion These results are interpreted as evidence for the immaturity of the speech motor control of children observed from two perspectives: insufficiently stable motor control patterns for vowel production, and a lack of effectiveness in anticipating forthcoming gestures. According to our theoretical framework, this lack of effective anticipation of forthcoming gestures reflects the immaturity of the neural representations of four-year-old children's speech motor systems, particularly in their inability to account for the appropriate variability compatible with correct perception of the target sound.

Further reading

Perkell, J. S. (2012). Movement Goals and Feedback and Feedforward Control Mechanisms in Speech Production. *J. Neurolinguistics, 25*(5), 382-407.

Whalen, D. H., Iskarous, K., Tiede, M. K, et al. (2005). The Haskins Optically Corrected Ultrasound System (HOCUS). *Journal of Speech, Language, and Hearing Research, 48*(3), 543-55.

LEARNING TO READ AND THE SENSORIMOTOR CONTROL OF SPEECH

Margriet A. Groen[1], Mark R. Van den Bunt[1], Stephen J. Frost[2], Airey Lau[2], Jonathan L. Preston[2,3], Vincent L. Gracco[2], Kenneth R. Pugh[2], Ludo Verhoeven[1]
[1]*Radboud University, Nijmegen, The Netherlands*
[2]*Haskins Laboratories, New Haven, United States of America*
[3]*Syracuse University, Syracuse, United States of America*

Reading and reading-development are often associated with the quality of phonological representations. Although phonological representations have both a perceptual and an articulatory target, studies on the relation between phonological representations and reading have almost exclusively focused on speech perception. Neurocomputational models of speech productions suggest that the quality of these representations is crucially dependent on the interaction between speech feedforward and feedback mechanisms (Guenther, Ghosh, & Tourville, 2006). Administering online alterations in auditory feedback, while participants produce speech, artificially induces mismatches between speech perception and production. Measuring the subsequent speech productions allow to determine how speech perception interacts with production, and linking this to reading and reading-related skills could shed new light on what aspect of phonological representations is most relevant in learning to read (Van den Bunt et al., 2017). Three research questions were examined: (1) Is the sensorimotor control in preliterate children different compared to literate children? (2) are individual differences in reading and reading-related skills associated with sensorimotor control? (3) are these relations different for languages that differ in orthographic transparency?

Eighty American children and 140 Dutch children in the age range from 4-8 years old were recruited and measures of reading ability, phonological awareness, letter knowledge and rapid naming were obtained. In addition, all participants underwent an altered auditory feedback task in which the first formant of the /ε/ vowel was not altered in a baseline phase, gradually increased in the ramp phase, hold at maximal perturbation (25%) in the hold phase and unaltered in the after-effect phase.

The results showed that being able to read was associated with a stronger response to altered auditory feedback. In addition, in the preliterate children, measures of phonological awareness, rapid naming and letter knowledge were positively correlated with response to altered feedback. No differences were found between languages. We propose that being able to read is associated with more precise articulatory and/or perceptual targets. Readers are possibly better in detecting the mismatch between their speech production and the fed back percept, and/or better able to modify their commands to adapt to the new circumstances. In addition, the correlations in preliterate children, of the response to altered feedback with precursors of reading, suggest a role for the sensorimotor control of speech in reading development.

References

Guenther, F. H., Ghosh, S. S., & Tourville, J. A. (2006). Neural modeling and imaging of the cortical interactions underlying syllable production. *Brain and Language, 96*, 280-301. doi:10.1016/j.bandl.2005.06.001

Van den Bunt, M. R., Groen, M. A., Ito, T., Francisco, A. A., Gracco, V. L., Pugh, K. R., & Verhoeven, L. (2017). Increased response to altered auditory feedback in dyslexia: A weaker sensorimotor magnet implied in the phonological deficit. *Journal of Speech, Language, and Hearing Research, 60*(3), 654-667. doi:10.1044/2016_JSLHR-L-16-0201

NORMATIVE DATA OF MAXIMUM PERFORMANCE TESTS OF SPEECH PRODUCTION ACROSS LIFE SPAN

Simone Knuijt, Leenke van Haaften, Lenie van den Engel-Hoek
Radboud University Medical Center, Nijmegen, The Netherlands

Background The Radboud Dysarthria Assessments consists of maximum performance tests of speech production (Knuijt et al., 2014) to examine the upper limits of speech motor performance. However, normative data are needed in order to distinguish pathological from typical speech motor performance. The aim of this study was to gather normative data in children and adults for maximum repetition rate (MRR) of /pataka/, maximum phonation time (MPT) of /a:/, fundamental frequency range (FFR) and maximum phonation volume (MPV).

Method 522 healthy Dutch speakers (5-80 years) performed the maximum performance tests three times. The best performance was used in the analysis. MRR, MPT and FFR were recorded with a linear PCM recorder (Tascam DR-05) and analyzed with Praat (Boersma en Weening, 1995). MPV was measured with a dB-meter (Voltcraft SL-100). Regression analysis were used to assess the effect of sex, age and body height. For MRR, we only assessed the effect of sex and age.

Results In MRR (n=424), we found a significant effect of age (p=.002) with a decreasing MRR with age. In MPT (n=493), we found a significant effect of sex (p=.001), with men having a longer MPT than women, and positive effects of age ($p<.001$) and body height ($p<.001$). In FFR (n=511), we found positive effects of age ($p<.001$) and body height (p=.001). In MPV (n=494), we found a significant effect of age ($p<.001$), with decreasing MPV with age.

Conclusion Age was the most important factor: MPT and FFR increased with age, while MRR and MPV decreased with age, especially after 70 years. Sex and body height were also related to MPT and sex was also related to FFR. Men having a longer MPT than women is plausible because of the effect of body height on pulmonary function (Quanjer et al., 2012). We found no effect of sex regarding MRR, which is in accordance with the literature (Tsao and Weismer, 1997). Despite the ongoing debate about motor control in speech versus motor control in nonspeech tasks, maximum performance tasks are important in clinical dysarthria assessment, not only for quantitative reasons, but the quality of performance may identify underlying pathology (weakness, coordination problems).

References

Knuijt, S., Kalf, J. G., van Gerven, M. H. J. C., Kocken, J. E. M., Kromhout, L. M., Goos, H. M. M., & De Swart, B. J. M. (2014). Nederlandstalig *Dysartrieonderzoek - volwassenen* (NDO-V). Houten: Bohn Stafleu van Loghum.

Quanjer, P. H., Hall, G. L., Stanojevic, S., Cole, T. J., Stocks, J., & Global Lungs, I. (2012). Age- and height-based prediction bias in spirometry reference equations. *Eur Respir J, 40*(1), 190-197.

Tsao, Y. C., & Weismer, G. (1997). Interspeaker variation in habitual speaking rate: evidence for a neuromuscular component. *J Speech Lang Hear Res, 40*(4), 858-866.

VARIABILITY IN LINGUAL CONTROL FOR VOWEL PRODUCTION IN GERMAN PRESCHOOLERS AND SCHOOLCHILDREN

Stella Krüger[1], Jan Ries[1], Aude Noiray[1,2]
[1] *University of Potsdam, Potsdam, Germany*
[2] *Haskins Laboratories, New Haven, United States of America*

Vowel acquisition involves learning articulatory configurations for their isolated productions (para-digmatic aspect) and their integration into coarticulated speech (syntagmatic aspect). While the former is acquired by the age of three, refinement of the latter takes longer (James et al., 2001). Such protract-ed development arises from anatomical growth as well as from maturation of speech motor control. This is reflected by age-dependent variability in vowel production affecting the shape and size of the vowel space. Therefore, variability indicates maturational processes and decreases as a function of age (Vorperian & Kent, 2007).

While there has been some attempts to integrate anatomical and acoustic data to investigate age-related variability in vowel production only a few have provided direct articulatory accounts. This study pro-vides a synthesis of acoustic and articulatory data to account for age effects on variability in vowel pro-duction investigating 3-, 4-, 5- and 7-year old German children. Similar data from adults were collected as a baseline. Rounded and unrounded tense long vowels (/i/, /y/, /u/, /a/, /e/, /o/) were recorded in non-words of the structure C1VC2ə. The acoustic speech signal was recorded along with movement of the tongue using ultrasound imaging. Hence, tongue position for target vowels could be tracked directly from the ultrasound images and estimated from the acoustic signal increasing the amount of available data. Our sample included both females and males to also account for sexual dimorphism.

The midpoint of the vowel was used as reference for data analysis. For each target vowel, the x- and y-coordinates of the highest point of the tongue were extracted using SOLLAR (Noiray et al., 2015). From the acoustic signal the first three formants were obtained semi-automatically in MATLAB (Story & Bunton, 2015).

Data are currently analyzed. Overall, we expect the 3-year olds to show the greatest within and be-tween-speaker variability in both articulation and acoustics. With increasing age and speech motor control development variability should decrease. This decline may not proceed in a linear fashion but rather be gradual and non-uniform across child. We expect to observe effects from sexual dimorphism by the age of 4 years.

References

James, D., van Dorn, J. & McLeod, S. (2001). Vowel production in mono-, di- and poly-syllabic words in children 3;0 - 7;11 years. In L. Wilson & S. Hewat (Eds.), *Proceedings of the Speech Pathology Australia Conference*, 127-136.

Story, B. H., & Bunton, K. (2015). A spectral filtering method for tracking formants in children's speech. In *Proceedings of Meetings on Acoustics 169ASA* (Vol. 23, No. 1, p. 060002). ASA.

Vorperian, H. K., & Kent, R. D. (2007). Vowel acoustic space development in children: A synthesis of acoustic and anatomic data. *Journal of Speech, Language, and Hearing Research, 50*(6), 1510-1545.

DOES REPETITION ALTER ACOUSTIC INDICES OF MOTOR PLANNING AND CONTROL?

Melissa A. Redford, Jeffrey E. Kallay
University of Oregon, Eugene, United States of America

Experimental investigations of speech production are typically based on multiple elicitations of the same items. These items are themselves controlled for numerous phonetic and phonological factors, rendering the speech task even more repetitive. The procedure minimizes the effect of noise on measures of interest, but it may also introduce motor learning into the experiment. Learning could undermine the validity of the conclusions drawn, especially about motor planning and control. For example, it could minimize age-related effects on production if repetition-induced increases in articulation rate and motor entrenchment are more pronounced in the (presumably) less optimized speech of younger children compared to the well-optimized speech of adults. This possibility was tested in the current study. Ten American-English speaking 5-year-old children and 10 adults were recorded producing 27 different simple SVO sentences that had a proper noun in S position, one of 3 verbs in V position, and 9 different noun phrases in O position. A determiner modified the real-word, monosyllabic nouns, which crossed 3 different onset consonants with 3 different vowel nuclei. Sentence elicitation was blocked by verb; object nouns were randomized. The protocol called for 6 repetitions of each sentence. Additional repetitions were elicited when a production was deemed incorrect or disfluent by the experimenter. In this way, the speakers analyzed so far (8 children, 2 adults) produced an average of 230 similar sentences during the study session. These sentences were divided into terciles based on order of elicitation to test whether early repetitions (1^{st} tercile) differed from later ones (2^{nd} and 3^{rd} terciles). All sentences with disfluencies or prosodic breaks in the predicate were identified and then excluded from further analysis. The determiner and object noun vowels were segmented in the remaining sentences. The segmented vowel durations and full formant transitions were extracted so that repetition effects on rate and coarticulation could be evaluated. Temporal variability (standard deviation duration/mean duration) was calculated within speaker and tercile across sentences with the same verb. Results from the subset of data analyzed so far indicated the expected significant age-related differences in duration and temporal variability. Vowel formants varied with the adjacent noun onset and/or nucleus, as expected. Elicitation tercile had no effect on adult speech, but it did effect childrens vowel durations. Conclusions are pending analyses on the entire data set.

Further reading

Redford, M.A. & Oh, G. (2017). Representation and execution of articulatory timing in first and second language acquisition. *Journal of Phonetics.* doi: 0.1016/j.wocn.2017.01.004.

Smith, A., & Goffman, L. (1998). Stability and patterning of speech movement sequences in children and adults. *Journal of Speech, Language, and Hearing Research, 41*, 18-30.

DEVELOPMENTAL FEATURES OF SPEECH IN TYPICALLY DEVELOPING CHILDREN - IMPLICATIONS FOR THE ASSESSMENT OF CHILDHOOD DYSARTHRIA

Theresa Schölderle, Elisabet Haas, Wolfram Ziegler
Ludwig-Maximilians-University, Munich, Germany

Background & Aim Childrens speech strongly differs from adult speech due to developmental influences. The anatomical immaturity of the speech apparatus, for instance, is mirrored in shorter expirations (Boliek, Hixon, Watson, & Jones, 2009), and developmental interactions with cognitive-linguistic abilities have been associated with slower articulation and higher pause rates. Remarkably, these features represent characteristics of typical speech motor development in children and, at the same time, are indicators of disturbed speech in adults. Therefore, the common practice to apply adult norms to children with dysarthria seems questionable (Cornwell, Murdoch, Ward, & Kellie, 2003). The objective of this study was to analyze developmental speech features in typically developing children following a neurophonetic taxonomy. We thereby address the long-term goal to establish age-specific norms for the assessment of childhood dysarthria.

Methods Twenty typically developing children between 3;0 and 6;11 years;months participated. Speech samples were elicited through a computer-based game prompting standard repetition sentences [see submission Haas et al.]. We analyzed childrens speech using the evaluation criteria of the Bogenhausen Dysarthria Scales (BoDyS), a standardized German assessment tool for adult speakers (Ziegler, Staiger, Schölderle, & Vogel, in press), which allows for a systematic neurophonetic analysis of respiration, phonation, articulation, and prosody. The auditory BoDyS judgements were validated by acoustic analyses.

Results The analyses are still underway. First results indicate that children between 3 and 6 years show a bundle of developmental speech characteristics that may be misinterpreted as dysarthric symptoms in neurologically impaired children, if an adult norm is applied (e.g., high frequency of inspirations, breathy voice, slow articulation rate).

Discussion Our study for the first time provides normative data of childrens speech corresponding to relevant neurophonetic categories. We thereby make a first step towards an age-standardized, valid assessment of childhood dysarthria.

References

Boliek, C. A., Hixon, T. J., Watson, P. J., & Jones, P. B. (2009). Refinement of speech breathing in healthy 4-to 6-year-old children. *Journal of speech, language, and hearing research, 52*(4), 990-1007.

Cornwell, P. L., Murdoch, B. E., Ward, E. C., & Kellie, S. (2003). Perceptual evaluation of motor speech following treatment for childhood cerebellar tumour. *Clinical linguistics & phonetics, 17*(8), 597-615.

Ziegler, W., Staiger, A., Schölderle, T., & Vogel, M. (in press). Gauging the Auditory Dimensions of Dysarthric Impairment: Reliability and Construct Validity of the Bogenhausen Dysarthria Scales (BoDyS). *Journal of Speech, Language and Hearing Research.*

QUANTIFYING THE DORSAL CONSTRICTION IN TYPICAL /k/-/t/ CONTRASTS IN CHILDREN AND ADULTS

James M. Scobbie[1], Joanne Cleland[2]
[1]*Queen Margaret University, Edinburgh, United Kingdom*
[2]*University of Strathclyde, Glasgow, United Kingdom*

We present a comparative quantification of the primary dorsal gesture in velar stops, comparing the tongue dorsum for /k/ against the tongue shape and location in /t/ - based on mid-sagittal ultrasound tongue images which have been stabilised using a headset. For a given vowel environment, /t/ provides the baseline against which the active, primary constriction undertaken by the tongue dorsum position to achieve /k/ is compared. Here, /k/ and /t/ are measured at a single target, the acoustic burst, though a comparison of the unfolding difference in normalised time differentiation is also possible.

This quantification is useful for a number of reasons, including cross-linguistic analysis, the study of co-articulation, or the longitudinal tracking of the acquisition of the /k/-/t/ contrast in a speech and language therapy clinic. It is this latter use which is our main interest (Cleland et al. 2016, 2017). It is relatively common in developmental disorder or delay (and in typical development, in younger children) for /k/ to undergo a "velar fronting" process in which /k/ appears perceptually indistinguishable from /t/, and our aim is to track the dorsal component of the contrast before, during, and after treatment, using an efficient measure related to the functional contrast between /k/ and /t/. Comparison to child (and adult) norms is important, and we report such norms here.

Such a measure does not directly examine the planning and execution of the gestural formation of the dorsal constriction directly, but still reveals a great deal about speech production. We discuss both an area measure of the "dorsal crescent" distinguishing /k/ and /t/, which is robust under mid-sagittal rotation and translation, and a radial linear measure, which is quick to measure and robust in cases where a dorsal crescent is incomplete. The two measures are highly correlated.

Our main results are from a sample of 30 typically-developing English-speaking children (5;7 to 12;8) each pronouncing one token of /ata/, /iti/, /oto/ and /aka/, /iki/, /oko/, providing three measures of the /k/-/t/ contrast per child. As well as area and radial linear differences, we will report other descriptive measures, like the visible length of tongue surface and the distance from probe to tongue surface, with reference to co-articulation and cross-sectional age-based developmental trends.

References

Cleland, J., Scobbie, J.M. and Wrench, A.A. (2015). Using ultrasound visual biofeedback to treat persistent primary speech sound disorders. *Clinical Linguistics & Phonetics, 29*(8-10), 575-559.

Cleland, J., Scobbie, J.M., Heyde, C.J., Roxburgh, Z. and Wrench, A.A. (2017). Covert contrast and covert errors in persistent velar fronting. *Clinical Linguistics & Phonetics, 31*(1), 35-55.

TYPICAL SPEECH RATE DEVELOPMENT AT A YOUNG AGE: A LONGITUDINAL STUDY

Anna Tendera, Matthew Rispoli, Torrey M. Loucks
University of Illinois at Urbana-Champaign, Champaign, United States of America

In typical development, one might hypothesize that speech rate increases gradually from when speech first emerges until it reaches a peak during adolescence. However, several studies have shown an absence of change in speech rate during the preschool years (e.g., Hall et al., 1999) leading several investigators to suggest that speech rate development does not follow a linear trajectory. Furthermore, the extant data on speech rate development does not support conclusions about linearity or any other pattern since previous longitudinal studies did not include enough time points. At least 4 time points are required to infer the shape of a growth function. The current study extends our first longitudinal report in which a considerable increase in speech rate from 24 to 36 months was observed. However, speech rate was only measured at two time points constraining the evaluation of the underlying pattern (Tendera et al., under review). In the present project speech rate growth functions were estimated at 4 time points (24, 30, 33 and 36 months) in 12 typically developing children to test whether early rate development followed a linear trajectory. Each child who participated in study produced at least ten active declarative sentences or ADS per session (Rispoli & Hadley, 2008). An ADS is a spontaneous production minimally consisting of a subject and a verb. Any repetitions, unintelligible speech, interrupted speech, or disfluencies longer than 250 msec were excluded. Each session was transcribed using the Systematic Analysis of Language Transcripts (SALT). Each audio recording of the sentences was processed with a noise reduction algorithm. The duration of each ADS was determined independently by two native English speakers who marked the onset and the offset of each ADS by examining an oscillogram and wide-band spectrogram concurrently. Speech rate in syllables per second was calculated as the number of syllables per ADS divided by duration. Growth curve analysis was used to model the shape of rate change across the 4 time points for each child. The results indicate that the speech rate growth function for each child during the critical 12-month period was linearly increasing, although the slope of the functions varied considerably across children. Assessing the speech rate growth is imperative for understanding typical speech development and speech disorders, but further longitudinal studies that include multiple time points are still needed to determine how speech rate develops after 3 years of age.

References

Hall, K. D., Amir, O., & Yairi, E. (1999). A longitudinal investigation of speaking rate in preschool children who stutter. *Journal of Speech, Language, and Hearing Research, 42*(6), 1367-1377.

Rispoli, M. & Hadley, P. A. (2008). *The growth of tense and agreement: final report.* National Science Foundation, BCS-0822513.

Tendera, A., Rispoli, M., & Loucks, T. (2017). *Speech rate in normal developing children between 24 and 36 months of life.* Manuscript submitted for publication.

THE DEVELOPMENT AND CLINIMETRIC EVALUATION OF A SPEECH PRODUCTION TEST FOR CHILDREN: RELIABILITY AND VALIDITY OF THE COMPUTER ARTICULATION INSTRUMENT (CAI)

Leenke van Haaften[1], Sanne Diepeveen[2], Lenie van den Engel-Hoek[1], Bert de Swart[1,2], Ben Maassen[3]
[1]*Radboud University Medical Center, Nijmegen, The Netherlands*
[2]*HAN University of Applied Sciences, Nijmegen, The Netherlands*
[3]*University of Groningen, Groningen, The Netherlands*

Introduction Children with speech sound disorders form a heterogeneous group. Classification systems are based on etiological, descriptive-linguistic or processing approaches (Waring & Knight, 2013), but there are no generally accepted criteria for differential diagnosis. Therefore, , the Computer Articulation Instrument (CAI) is designed and developed (Maassen et al., in press). The CAI has a modular structure and provides interactive administration and scoring of four tasks: (1) Picture naming, (2) Nonword imitation, (3) Word and Nonword repetition and (4) Maximum Repetition Rate (MRR), thereby covering phonological and speech motor skills. It is completely computer implemented and involves two steps: (1) assessing objective and quantitative symptoms, and (2) contrasting tasks varying in production processes. Normative data were collected. The aims of this study were to assess interrater and test-retest reliability and the construct validity of the CAI.

Method In total, 1524 typically developing children aged between 2;0 and 7;0 years participated in the normative study. The sample was representative for gender, geographic region and degree of urbanization in the Netherlands. Parameters such as percentage of consonants correct (PCC), percentage of vowels correct (PVC) and percentage of correct syllable structures were calculated for Picture naming and Nonword imitation. A proportion of whole word variability was calculated for Word and Nonword repetition and number of syllables per second was calculated for MRR. Interrater reliability and test-retest reliability were analyzed using subgroups of the total sample. Interrater reliability ($n = 99$) was studied by estimating Intraclass Correlation Coefficients (ICCs), and test-retest ($n = 124$) by correlation coefficients. Construct validity was assessed by determining age-related changes of test-results.

Results For interrater reliability ICCs ranged from sufficient to good, except for PVC for picture naming and nonword imitation and for the bi- and trisyllabic items of MRR. The correlation coefficients of the test-retest reliability ranged from good (picture naming) to insufficient (word and nonword repetition, MRR). Continuous norms show development patterns for all the parameters of the CAI.

Discussion First results indicate that most of the parameters of the CAI are reliable and valid. Analyses of phonological inventory, phonological structure, syllabic structure and speech motor skills yield indices of speech development and from there norm-referenced estimates of speech delay. The assessment tasks tap into different levels of processing, and therefore the CAI is a valuable tool for the assessment of speech sound disorders in a psycholinguistic framework.

References

Maassen, B., van Haaften, L., Diepeveen, S., Terband, H., van den Engel-Hoek, L., Veenker, T., de Swart, B. (in press). *Computer Articulation Instrument.* Amsterdam: Boom Uitgevers.

Waring, R., & Knight, R. (2013). How should children with speech sound disorders be classified? A review and critical evaluation of current classification systems. *Int J Lang & Comm Dis, 48*(1), 25-40.

ARTICULATORY DIFFERENCES BETWEEN L1 AND L2 SPEAKERS OF ENGLISH

Martijn Wieling[1], Pauline Veenstra[1], Patti Adank[2] and Mark Tiede[3]
[1] *University of Groningen, Groningen, The Netherlands*
[2] *University College London, London, United Kingdom*
[3] *Haskins Laboratories, New Haven, United States of America*

Second language learners (L2) learners typically have a noticeable accent, especially when L2 learning begins at a later age. Speech learning models, such as Fleges Speech Learning Model (SLM; Flege, 1995) explain these difficulties in L2 pronunciation by considering the phonetic similarity of the speakers L1 and L2. L2 sound segments that are very similar to those in the L1 are predicted to be harder to learn than those which are not. The SLM predicts that the former map onto an existing sound category, while the latter would generate a new sound category.

In this study we focus on the dental fricative /θ/ in English and assess if Dutch and German L2 speakers of English distinguish this (novel) sound from the similar sound /t/ in their respective native languages. For this purpose, we collected articulatory data (using a portable 16-channel 100 Hz NDI Wave device) for 22 native English speakers, 20 native Dutch speakers and 27 native German speakers. Our material consisted of 10 sets of minimal pairs in English involving /t/ vs. /θ/ (e.g., "fate"-"faith"). Each word was preceded and followed by a /ə/ in order to ensure a neutral articulatory context during the word pronunciation.

After the articulatory segmentation, we analyzed the word-based tongue sensor trajectories using generalized additive modeling (see Wieling et al., 2016). Our findings show that English speakers clearly distinguish /t/ from /θ/, with a significantly more anterior tongue tip position for the /θ/-words than for the /t/-words ($p < .001$). While the German speakers showed a similar pattern ($p < .001$), the Dutch speakers did not show a significant anterior position difference ($p > .1$). Further analyses showed that the Dutch speakers English /t/-words and /θ/-words did not differ significantly from similar sounding /t/-words pronounced in Dutch (e.g., "tiem" [tim]).

In the context of Fleges SLM, our results suggest an articulatory merger between /t/ and /θ/ for Dutch L2 speakers of English, and this finding adds to earlier acoustic results (Hanulikova & Weber, 2010) which focused on similar speaker groups.

References

Flege, J. E. (1995). Second language speech learning: Theory, findings, and problems. *Speech perception and linguistic experience: Issues in cross-language research*, 233-277.

Hanulikova, A., & Weber, A. (2010). Production of English interdental fricatives by Dutch, German, and English speakers. In *New Sounds 2010: Sixth International Symposium on the Acquisition of Second Language Speech* (pp. 173-178). Adam Mickiewicz University.

Wieling, M., Tomaschek, F., Arnold, D., Tiede, M., Bröker, F., Thiele, S., Wood, S.N. & Baayen, R. H. (2016). Investigating dialectal differences using articulography. *Journal of Phonetics, 59*, 122-143.

THE EFFECT OF LEXICAL STRESS ON SEGMENTAL PRODUCTION ACCURACY IN PERSONS WITH APRAXIA OF SPEECH

Dallin J. Bailey[1], Lisa Bunker[2,3], Shannon Mauszycki[2,3], Julie L. Wambaugh[2,3]
[1]*Auburn University, Auburn, United States of America*
[2]*University of Utah, Salt Lake City, United States of America*
[3]*Salt Lake City VA Healthcare System, Salt Lake City, United States of America*

Introduction Acquired apraxia of speech (AOS) includes deficits in prosody and segmental accuracy. Recently, Aichert, Spath, and Ziegler (2016) identified an interaction between the metrical structure of bisyllabic words and word production accuracy in a group of German-speaking adults with AOS, with trochaic words (strongweak) being more resistant to production errors than iambic words (weak-strong). The purpose of this study was to identify whether similar patterns exist in English-speaking adults with AOS, and to measure the reliability of any pattern over repeated sampling occasions.

Methods
Participants Data for the current study came from 27 with AOS and were collected as part of a larger study investigating AOS treatment outcome measures. Participants were at least 12 months post stroke or penetrating brain injury. All were diagnosed with aphasia and AOS, with a range of AOS severity.
Procedure Productions were elicited via verbal model. Sampling was conducted on three separate occasions: the initial data collection session, and subsequent sampling completed one and four weeks later.
Stimuli Stimuli consisted of 79 bisyllabic words with a CVCVC segmental structure. Metrical structures included 42 trochees and 37 iambs.
Transcription and Scoring All speech samples were phonetically transcribed into machine-readable characters. These transcriptions were aligned with canonical transcriptions via an edit distance algorithm (Levenshtein, 1966) that followed transcription alignment principles (Preston, Ramsdell, Oller, Edwards, & Tobin, 2011). Phonetic-level errors (distortions) were penalized less severely than phonemic-level errors. Trochee and iamb lists were analyzed separately.

Results Paired samples t-tests indicated that the modified edit distance was significantly lower for the trochee lists than for the iamb lists. There was a lack of a significant effect of time on the absolute difference between modified edit distance for the two lists.

Discussion The results strongly suggest that in English, as in German, the trochaic structure is more resistant to segmental errors in persons with AOS. Further, this effect is robust over time. Implications for selecting treatment stimuli will be discussed.

References
Aichert, I., Spath, M., & Ziegler, W. (2016). The role of metrical information in apraxia of speech. Perceptual and Acoustic Analyses of Word Stress. *Neuropsychologia*.

Levenshtein, V. I. (1966). Binary codes capable of correcting deletions, insertions, and reversals. *Soviet physics doklady, 10*(8), 707-710.

Preston, J. L., Ramsdell, H. L., Oller, D. K., Edwards, M. L., & Tobin, S. J. (2011). Developing a Weighted Measure of Speech Sound Accuracy. *Journal of Speech Language and Hearing Research, 54*(1), 1.

MOTOR SPEECH DISORDERS ACCOMPANYING RS AND SL VARIANTS OF PROGRESSIVE SUPRANUCLEAR PALSY

Heather Clark, Joseph Duffy, Edythe Strand, Jennifer Whitwell, Keith Josephs
Mayo Clinic, Rochester, United States of America

Progressive supranuclear palsy (PSP) is a tauopathy affecting gait, eye movements, swallowing, and communication. Typical PSP (Richardson Syndrome; PSP-RS) presents as balance impairment and vertical gaze palsy, with dysarthria appearing early or as the disease progresses. Hypokinetic features are most common, but spastic and ataxic features may also be present. Josephs and colleagues (2005) described a cohort of patients with atypical PSP (Speech and Language; PSP-SL), in which progressive apraxia of speech (AOS) was the presenting symptom (Duffy et al., 2015). Whereas previous studies compared the imaging characteristics of PSP-RS and PSP-SL, the current study is the first to describe performance of these groups on a standard assessment of motor speech function.

Participants included 15 patients with PSP-RS and 15 with PSP-SL. Group designation was determined by presenting complaints and neurologic assessment. All participants underwent a standard motor speech assessment as one component of larger studies examining imaging findings in neurodegenerative disease. The speech sample included picture description, word and sentence repetition, and alternate and sequential motion rate (diadochokinetic) tasks. Nature and severity of dysarthria and/or apraxia of speech were determined by speech-language pathologists experienced in differential diagnosis of motor speech disorders (authors 1-3). Functional severity of overall motor speech function was informed by patient and family report.

Age of onset, sex, handedness, and overall disease severity did not vary across groups. All participants with PSP-RS demonstrated dysarthria. Isolated hypokinetic dysarthria was observed in 9 participants; the remaining participants demonstrated mixed dysarthria with predominant hypokinetic features and less prominent spastic and/or ataxic features. Of the nine PSP-SL participants who exhibited dysarthria, five demonstrated spastic dysarthria, two hypokinetic dysarthria, and two mixed dysarthria characterized by hypokinetic, spastic, and/or ataxic features. AOS was observed for all participants with PSP-SL. In contrast, no participants with PSP-RS had unequivocal evidence of AOS; a single participant exhibited subtle segmentation of syllables. Functional severity of motor speech impairment was slightly greater in the PSP-SL group, with four participants using some form of AAC to supplement speech. Most participants in both groups reported that they were asked to repeat themselves frequently.

In summary, dysarthria was a common finding in PSP-RS, with hypokinetic features predominating. Speakers with PSP-SL less commonly demonstrated dysarthria and features were more commonly spastic. Although AOS was a defining feature of PSP-SL, it was not observed in PSP-RS. These findings suggest subcortical dysfunction is more severe in PSP-RS, while cortical dysfunction may be more severe in PSP-SL.

References

Duffy, J. R., Strand, E. A., Clark, H., Machulda, M., Whitwell, J. L., & Josephs, K. A. (2015). Primary progressive apraxia of speech: clinical features and acoustic and neurologic correlates. *Am J Speech Lang Pathol, 24*(2), 88-100. doi:10.1044/2015_AJSLP-14-0174

Josephs, K. A., Boeve, B. F., Duffy, J. R., Smith, G. E., Knopman, D. S., Parisi, J. E., ... Dickson, D. W. (2005). Atypical progressive supranuclear palsy underlying progressive apraxia of speech and nonfluent aphasia. *Neurocase, 11*(4), 283-296. doi:10.1080/13554790590963004

PREDICTIVE TIMING IN STUTTERING

Simone Falk[1,2,3], Marie-Charlotte Cuartero[2], Ramona Schreier[3], Simone Dalla Bella[4,5]
[1] *University Sorbonne Nouvelle Paris-3, Paris, France*
[2] *Aix-Marseille University, Marseille, France*
[3] *Ludwig-Maximilians-University, Munich, Germany*
[4] *University of Montpellier, Montpellier, France*
[5] *Institut Universitaire de France, France*

Stuttering is a developmental fluency disorder characterized by severe disruptions of the rhythmic flow of speech. Approximately one percent of the adult population is affected by stuttering. Recent studies suggest that stuttering may arise from deficits in speech motor control linked to alterations in the cortico-subcortical timing and rhythm network (e.g., Etchell et al., 2014). In the present study, we tested whether temporal predictions guiding the coupling of perception and action ("predictive timing") are altered in verbal as well as non-verbal rhythmic tasks in adults who stutter. We tested predictive timing using verbal and non-verbal synchronization. In the tasks, participants synchronized words or finger taps with a metronome beep (i.e., paced speech / paced finger tapping) at different rates. Thirty-four French-speaking adults (17 participants who stutter, 17 age-matched controls) were tested on these tasks. Our findings revealed predictive timing deficits in both verbal and non-verbal production in adults who stutter. Adults who stutter were less consistent and less accurate in paced tapping and also showed altered timing in paced speech. In particular, we found alterations in vowel timing in syllable production. Overall, these results corroborate and extend recent findings on non-verbal predictive timing deficits we obtained with children and adolescents who stutter (Falk et al., 2015). More generally, the results are discussed in light of recent theories of predictive timing and the coupling of perception and action in fluent speech production as well as potential mechanisms underlying the motor basis of stuttering.

References

Etchell, A. C., Johnson, B. W., & Sowman, P. F. (2014). Behavioral and multimodal neuroimaging evidence for a deficit in brain timing networks in stuttering: a hypothesis and theory. *Frontiers in Human Neuroscience, 25*(8), 467. http://dx.doi.org/10.3389/fnhum.2014.00467

Falk, S., Müller, T., & Dalla Bella, S. (2015). Non-verbal sensorimotor timing deficits in children and adolescents who stutter. *Frontiers in Psychology, 6*, 847. http://dx.doi.org/10.3389/fpsyg.2015.00847.

FOREIGN ACCENT SYNDROME - A TYPE OF KINETIC AOS

Judith Feiken[1], Melanie Kirmess[2,3]
[1]*University Medical Center Groningen, Groningen, The Netherlands*
[2]*University of Oslo, Oslo, Norway*
[3]*Sunnaas Rehabilitation Hospital, Norway*

Foreign Accent Syndrome (FAS) is a speech disorder following brain injury leading to a perceived presence of a new accent in a speaker's speech (Jonkers et al., 2016). Most researchers agree on FAS to be due to a deficit in motor speech. Jonkers et al. (2016) argue that when persons with apraxia of speech (AoS) adapt to their underlying motor programming disorder using strategies as segmentation and scanned speech, speakers will put more force of articulation. This leads to symptoms like a lack of assimilation and stress on every single syllable, which may cause the different perceived accent. Further, they describe that FAS is also observed in speakers with dysarthria. However, the typical AoS symptoms, like halting, groping and false starts, are rarely reported for persons with FAS. In this presentation we therefore propose another explanation: FAS is related to a specific type of AoS - the kinetic subtype.

AoS can be divided into subtypes (Feiken and Jonkers, 2012). The most regular subtype can be referred to as ideomotor AoS, in parallel to ideomotor apraxia as one of the subtypes of limb apraxia. These persons have difficulties performing purposeful movements, while these movements can be facilitated correctly as an automatic or emotional reaction. The cause of this deficit is an impairment in the access to motor programs of speech that are automate during speech development. The second subtype can be referred to as kinetic AoS in parallel to kinetic limb apraxia. In kinetic AoS, there are no problems in initiating speech movements, but the emphasis is on the improperly executed movements. This means that speakers express mainly phonetically poorly formed phonemes, causing distortions, substitutions, reduced coarticulation and thereby decreased intelligibility. These symptoms resemble those of dysarthria, which can easily lead to misdiagnosis. Furthermore, these symptoms together with disturbed dysprosody caused by adaptations can also be perceived as another accent. We therefore propose that the underlying deficit of FAS is a kinetic subtype of AoS, and suggest reconsidering dysarthria as an underlying deficit of FAS.

This specification of FAS is of great importance for therapy, because speech therapy can be targeted more accurate. Instead of getting dysarthria therapy or learning to regain access to the motor programs as with ideomotor AoS, FAS patients have to relearn and automate the correct articulatory motor programs and learn how to use preventive adaptations as efficiently as possible.

References

Jonkers, R., van der Scheer, F., & Gilbers, D. (2016). The common denominator in the perception of accents in cases with foreign accent syndrome. *Aphasiology*, 1-23.

Feiken, J.F., & Jonkers, R. (2012). *A division into three subtypes of apraxia of speech: ideomotor, kinetic and ideational.* Paper presented at the CAC 2012, CA, 20-24, May 2012.

MODULATING NEURAL ACTIVITY WITH NON-INVASIVE BRAIN STIMULATION IN ADULTS WHO STUTTER

Emily O. Garnett[1], Ai Leen Choo[2], Ho Ming Chow[1], Soo-Eun Chang[1]
[1]*University of Michigan, Ann Arbor, United States of America*
[2]*California State University East Bay, Hayward, United States of America*

There is a fundamental and unmet need for finding lasting therapeutic approaches to stuttering, a speech disorder that affects 1% of the population. Stuttering that persists into adulthood often results in negative psychosocial consequences. Although speech fluency can be temporarily induced via externally-paced conditions like choral speech, interventions that result in long-term alleviation of stuttering symptoms are limited. There is accumulating evidence of anomalous brain connectivity in stuttering, particularly in the basal ganglia thalamocortical (BGTC) network that includes supplementary motor area (SMA) and ventral premotor cortex (vPMC). Transcranial direct current stimulation (tDCS) offers the possibility of short-term and, in some cases, long-term functional improvement in language and motor skills. The objective of this study was to test potential augmentative effects of pairing high definition (HD)-tDCS with fluency-inducing conditions (choral speech). We hypothesized that anodal stimulation targeting left SMA would modulate activity in SMA and functionally connected regions within the BGTC network.

Ten adults who stutter (AWS) and 10 fluent speakers completed the study, a sham-controlled, counterbalanced, double-blind, within-subjects design. Participants completed two sessions, one week apart, consisting of fMRI before and after 20 minutes of HD-tDCS (active/sham) targeting the left SMA. During stimulation, participants read aloud synchronized with a metronome. Participants performed two speech tasks during scanning (solo and choral reading) and completed a resting state scan. Data preprocessing and statistical analyses were completed using the GLM in SPM12. T-tests were used to compare brain activity before and after sham and active stimulation separately in the solo and choral reading conditions.

Increased neural activity was found in left inferior frontal gyrus (IFG) after active but not sham stimulation during solo reading, in one cluster spanning left BA44 and BA6 (vPMC). No significant differences were found in the choral reading condition. Preliminary results suggest that neural activity among areas relevant to speech timing may be enhanced with HD-tDCS in AWS. Stimulating an area involved in speech initiation/timing (SMA) specifically increased activity in left vPMC, shown in previous studies to exhibit anomalous function and structure in AWS. These results will be discussed in the context of the frontal aslant tract that connects SMA and vPMC, found to have lower white matter organization in AWS.

References

Beal, D. S., Gracco, V. L., Brettschneider, J., Kroll, R. M., & De Nil, L. F. (2013). A voxel-based morphometry (VBM) analysis of regional grey and white matter volume abnormalities within the speech production network of children who stutter. *Cortex, 49*(8), 2151-2161.

Chang, S.-E., & Zhu, D. C. (2013). Neural network connectivity differences in children who stutter. *Brain, 136*(12), 3709-3726.

Kell C.A., Neumann K., von Kriegstein K., Posenenske C., von Gudenberg A.W., Euler H., & Giraud A.L. (2009). How the brain repairs stuttering. *Brain, 132*, 2747-2760.

COMPARISON OF BIOFEEDBACK APPROACHES TO INTERVENTION FOR RESIDUAL SPEECH ERRORS

Elaine R. Hitchcock[1], Tara McAllister Byun[2]
[1]*Montclair State University, Montclair, United States of America*
[2]*New York University, New York, United States of America*

Speech sound disorders affect up to 10% of children and have been shown to negatively impact long-term health outcomes. Despite considerable research attention, best-practice guidelines for intervention remain elusive. Evidence has recently accumulated that treatment using technology to provide enhanced sensory feedback can improve production in individuals who have not responded to previous forms of intervention (McAllister Byun & Hitchcock, 2012). The present study was undertaken with two goals. (1) For three types of biofeedback (ultrasound, electropalatography, real-time LPC spectra), conduct a rigorous assessment of efficacy using an innovative single-case methodology, Masked Visual Analysis (MVA, Ferron & Jones, 2006); (2) compare outcomes when equal doses of treatment are provided across these three types of biofeedback.

MVA aims to reduce the influence of experimenter bias on single-case design while still using a response-guided approach to ensure that each phase of the study exhibits a consistent and interpretable response pattern. The study team is divided into the intervention team, which provides treatment and collects data for plotting, and an analysis team, which inspects the data and determines when all data appear sufficiently stable to initiate treatment for one randomly selected subject. Crucially, the data reviewed by the analysis team are masked for participants identity/treatment status. After study completion, the analysis team guesses the order in which treatment was applied across subjects. A p-value can be derived by dividing the number of guesses needed to identify the correct order by the total number of possible orderings of the data.

This study consisted of three investigative arms (ultrasound, LPC, EPG), each a multiple-baseline across-subjects study of four participants with /r/ misarticulation (ages 7;6-13;0). The protocol for intervention was kept as consistent as possible across all three arms, and a standard dosage of 72 trials was elicited per session over twenty 30-minute sessions.

In each study arm, the analysis team correctly guessed the order of treatment application on the first try, yielding $p<.05$. Thus, the hypothesis that biofeedback can be an effective intervention for residual /r/ errors was supported for the technologies studied. However, mean effect sizes varied considerably across biofeedback types (mean d2: ultrasound, 1.8; LPC, 7.1; EPG, 11.5). Although the samples used here are too small to draw conclusions regarding the relative efficacy of the three types of biofeedback, we discuss factors that may have contributed to the observed differences, and how these results can guide future larger-scale comparison studies.

References

Ferron, J., and Jones, P. K. (2006). Tests for the visual analysis of response-guided multiple-baseline data. *The Journal of Experimental Education, 75*(1), 66-81.

McAllister Byun, T., and Hitchcock, E. R. (2012). Investigating the use of traditional and spectral biofeedback approaches to intervention for /r/ misarticulation. *American Journal of Speech-Language Pathology, 21*(3), 207-221.

CONSONANT CLUSTERS IN PATIENTS WITH PHONOLOGICAL IMPAIRMENT VS. APRAXIA OF SPEECH

Hanna Jakob, Wolfram Ziegler
Ludwig-Maximilians-Universität, Munich, Germany

Background Patients with phonological impairment as well as patients with apraxia of speech produce errors like phoneme substitutions or -elisions. It is still unclear if these two patient groups, which are deemed to suffer from different underlying pathologies, can be distinguished on the basis of their error patterns. Several studies have used phonologically complex stimuli, e.g., words containing consonant clusters, to evoke phonological errors in these patients. A common finding is that patients from both groups tend to simplify syllable structure, e.g., by omitting one of the consonants of a cluster or by inserting a schwa between cluster consonants (Buchwald & Miozzo, 2011; 2012). However, several studies found this tendency only in patients with apraxia of speech, whereas patients with phonological impairment produced similar amounts of simplification and complication errors (Galluzzi et al., 2015).

Methods Previous studies used paradigms involving elicitation of single words, e.g., through word repetition or object naming tasks. Yet, it is known that speech production in patients with aphasia or apraxia of speech can be particularly vulnerable to context influences. In the present study we therefore elicited word pairs to examine potential context effects on speech accuracy in clusters vs. singletons. Single word production served as a control condition. Within word pairs we controlled for phonological similarity effects. Clusters were varied with regard to their conceived amount of gestural overlap. In total, 20 patients with apraxia of speech and phonological impairment were included.

Results & Discussion The analyses are still under way. First results indicate that both groups simplified consonant clusters. Consistent with previous studies, there were more simplification errors in the patients with apraxia of speech than in the phonological impairment group. However, we also found many errors creating more complex syllable structures in both groups, which contradicts optimality constraints.

The findings will be discussed on the background of phonetic and phonological theories.

References

Buchwald, A. & Miozzo, M. (2011). Finding levels of abstraction in speech production evidence from sound-production impairment. *Psychological Science, 22*(9), 1113-1119.

Buchwald, A. & Miozzo, M. (2012). Phonological and motor errors in individuals with acquired sound production impairment. *Journal of Speech, Language, and Hearing Research, 55*(5), S1573-S1586.

Galluzzi, C., Bureca, I., Guariglia, C. & Romani, C. (2015). Phonological simplifications, apraxia of speech and the interaction between phonological and phonetic processing. *Neuropsychologia, 71*, 64-83.

THE EFFECTS OF F0, INTENSITY AND DURATIONAL MANIPULATIONS ON THE PERCEPTION OF STRESS IN DYSARTHRIC SPEECH

Anja Lowit, Tolulope Ijitona, John Soraghan
University of Strathclyde, Glasgow, United Kingdom

Marking stress plays an important role in conveying meaning and directing the listeners attention to the important parts of a message. Extensive research has been conducted into how healthy speakers produce stress, with the key phonetic cues acknowledged as F0, intensity and duration (Bolinger 1961). We also know that speakers with dysarthria experience problems in marking stress successfully (Lowit et al., 2012, Patel & Campellone, 2009). However, we currently lack sufficiently specific information on these deficits and potential compensatory techniques to allow Speech and Language Therapists (SLTs) to provide effective treatment methods to address stress production problems.

In order to build an evidence base for intervention, it is essential to first establish the relationship between features of disordered stress production and their perceptual outcomes. In particular, we need to know which phonetic cues (or combinations thereof) are most salient and what degree of change needs to be achieved in order to signal stress effectively to listeners.

This project aims to explore these questions in detail by performing perceptual experiments on data from disordered speakers that have been acoustically manipulated, in order to produce guidance to clinicians on how to improve their patients ability to signal stress successfully.

We used contrastive stress sentences from 10 speakers with ataxic dysarthria. Each speaker produced 30 sentences, i.e. 10 sentences (SVOA structures) across 3 conditions (stress on initial (S), medial (O), or final (A) target words). Sentences were perceptually scored by 5 listeners regarding the location of the stress target. We then chose 15 utterances where listeners had been unable to identify the target, five for each of the sentence positions. These utterances were subsequently manipulated acoustically by incrementally increasing the F0, intensity and duration of the target words, in accordance with the degree of change observed in the healthy control group. In addition, pausing patterns as well intonation contours were altered. The manipulated utterances were played to 50 listeners to evaluate what degree and combination of alteration resulted in correct identification of the stress target.

We will report on the patterns of impairment observed in the disordered speech samples, as well as the impact of the above manipulations on listener accuracy. This will provide information for future studies on stress production regarding focus of analysis, as well as guide clinicians on how best to address deficits in this area in their patients.

References

Bolinger, D. (1961). Contrastive Accent and Contrastive Stress. *Language, 37*, 83-96.

Lowit, A., Kuschmann, A., MacLeod, J. M., Schaeffler, F., & Mennen, I. (2010). Sentence Stress in Ataxic Dysarthria - A Perceptual and Acoustic Study. *Journal of Medical Speech - Language Pathology, 18*, 77-82.

Patel, R., & Campellone, P. (2009). Acoustic and Perceptual Cues to Contrastive Stress in Dysarthria. *Journal of Speech, Language, and Hearing Research, 52*, 206-222.

EXAMINING THE ARTICULATORY BUFFER IN APRAXIA OF SPEECH

Marja-Liisa Mailend[1], Edwin Maas[2], Pelagie Beeson[1]
[1] *University of Arizona, Tucson, United States of America*
[2] *Temple University, Philadelphia, United States of America*

In this study, we investigated the underlying nature of apraxia of speech (AOS) in the context of the DIVA model (Guenther, 2016) by testing two competing hypotheses focused on the function of the articulatory buffer. The *Reduced Buffer Capacity Hypothesis* argues that people with AOS can hold only one syllable at a time in the speech motor planning buffer (Rogers & Storkel, 1999). The *Program Retrieval Deficit Hypothesis* states that people with AOS have difficulty retrieving the intended motor program from the buffer when it contains competing motor programs (Mailend & Maas, 2013).

The experimental paradigm required participants to plan and hold a two-word utterance until a go-signal prompted a spoken response. Target words shared the rhymes while the phonetic similarity of the onset consonants was manipulated to create four experimental conditions. In the *Identical* condition the onset consonants were the same (e.g., *bill bill*), in *Similar Place and Manner (PM)* condition they differed only in the voicing feature (*bill pill*), in *Similar Voicing and Manner (VM)* condition they differed only in the place of articulation (*bill dill*), and in the *Different* condition, the onsets did not share any phonetic features (*bill fill*). The primary dependent measure was reaction time of correct responses.

The Reduced Buffer Capacity Hypothesis predicted that speakers with AOS will not show RT differences across conditions because they plan the utterance word-by-word and the first word is the same in all conditions. In contrast, the Program Retrieval Hypothesis predicted a Group x Condition interaction (greater difference between the Identical and other conditions for speakers with AOS than unimpaired speakers) because retrieval of the intended program is difficult when several programs are activated and compete for selection.

Data from six speakers with AOS and concomitant aphasia, ten speakers with aphasia without AOS, and 24 age-matched control speakers revealed a group by condition interaction for reaction times: the difference between the Identical Condition and all other conditions was larger in the AOS group compared to the other two groups.

These results are consistent with the Program Retrieval Deficit Hypothesis suggesting that at least one component of the speech motor planning impairment in AOS reflects difficulty resolving the competition between programs that are simultaneously activated in the speech motor planning buffer.

References

Guenther, F. H. (2016). *Neural control of speech.* Cambridge, MA: MIT Press.

Mailend, M.-L., & Maas, E. (2013). Speech motor programming in apraxia of speech: Evidence from a delayed picture-word interference task. *American Journal of Speech-Language Pathology, 22*(2), S380-S396.

Rogers, M. A., & Storkel, H. L. (1999). Planning speech one syllable at a time: the reduced buffer capacity hypothesis in apraxia of speech. *Aphasiology, 13*(9-11), 793-805.

THE LOSS OF COMPLEX COUPLING MODES IN SPEECH DYNAMICS: AN EMA-STUDY ON ESSENTIAL TREMOR PATIENTS TREATED WITH DEEP BRAIN STIMULATION

Doris Mücke[1], Anne Hermes[1], Tabea Thies[1], Johannes Becker[2], Michael Barbe[2]
[1]*University of Cologne, Cologne. Germany*
[2]*University Hospital Cologne, Cologne, Germany*

Chronic deep brain stimulation (DBS) of the nucleus ventralis intermedius (VIM) is an effective treatment for patients with medication-resistant Essential Tremor (ET). However, these patients report that stimulation has a deleterious effect on their speech. Previous studies on VIM-DBS in patients report a deterioration of the oral speech motor system, but those studies were restricted to the acoustic domain (Mücke et al, 2014; Pützer, Barry & Moringlane, 2007). The present study investigates gestural coordination patterns in ET patients treated with VIM-DBS in the articulatory domain. We focus on consonant and vowel timing, and relate the results to the coupling hypothesis syllable structure (Browman & Goldstein, 2000).

We recorded 12 ET patients, German native speakers treated with DBS in the ON and OFF condition, and 12 age-matched control speakers with Electromagnetic Articulography (AG501). Sensors were placed on the lips, the tongue tip and the tongue dorsum. We recorded a fast syllable repetition task (DDK) as well as a sentence production task, where target words containing syllable onsets with a low or high complexity, such as in <Lima> vs. <Klima> or <Plina>, were embedded in carrier sentences. Gestural landmarks such as onset, peak velocity and target of consonantal and vocalic movements were labelled in the vertical plane.

(1) In the DDK-task, different variables used in a mass-spring model (displacement, duration, peak velocity, stiffness) were computed. We found a decrease in stiffness for patients in the stimulation-OFF condition compared to healthy controls ($p<0.0057$, linear mixed model). This effect intensified in the patients production in the stimulation-ON condition ($p<0.0001$), revealing an overall slowing down of the speech motor system under stimulation. However, we also found a high degree of performance variability: comparing patients to healthy controls, patients had problems adapting to the requirements of this novel motor task [6].

(2) In the sentence production task, we measured the degree of pattern stability in syllable onsets with low and high complexity (shifts of rightmost-C, leftmost-C) proposed within the coupled oscillator model framework. This paradigm follows the assumption that the prevocalic consonant in /*lima*/ (in-phase) shifts towards the following vowel to make room when a consonant is added to the syllable, as in /*klima*/ (competition of in-phase and anti-phase). Results for the patients show that the prevocalic, i.e. rightmost, C does not shift towards the following V, but away from it (-27 ms). This mistiming worsens with stimulation-ON (-48 ms), revealing a gradient loss of complex, learnt coupling patterns in motor control while simple, innate patterns remain stable.

References

Browman, C. P., & Goldstein, L. (2000). Competing constraints on intergestural coordination and selforganization of phonological structures. *Bulletin De La Communication Parlée, 5*, 25-34.

Mücke, D., Becker, J. Barbe, M.T., Roettger, T.B., Meister, I., Liebhart, L., Timmermann L., & Grice, M. (2014). The effect of Deep Brain Stimulation on the speech motor system in Essential Tremor Patients. *Journal of Speech, Language, and Hearing Research, 57*(4), 1206-1218.

P¨utzer, M., Barry, W. J. & Moringlane, J. R. (2007). Effect of deep brain stimulation on diff er-ent speech subsystems in patients with multiple sclerosis. *Journal of Voice 21*(6), 741-753.

EXPLORING QUANTIFIABLE MEASURES FOR THE EVALUATION OF SLP INTERVENTION FIDELITY

Aravind K. Namasivayam[1,2], Roslyn Ward[3,4,5], Rohan Bali[1], Paul Davey[5], Geoff Strauss[5], Mary Claessen[5], Deborah Hayden[6], Pascal van Lieshout[1,2]
[1] *University of Toronto, Toronto, Canada*
[2] *Toronto Rehabilitation Institute, Toronto, Canada*
[3] *Princess Margaret Hospital for Children, Subiaco, Australia*
[4] *The University of Western Australia, Perth, Australia*
[5] *Curtin University, Bentley, Australia*
[6] *The PROMPT Institute, Santa Fe, United States of America*

Aim An important aspect of treatment fidelity is the quality or the skillfulness with which a Speech-Language Pathologist (S-LP) delivers a given treatment. We propose to explore new methods to quantitatively assess perceptual and fine motor skills of clinicians working in the area of speech motor disorders, as a part of a larger fidelity assessment protocol.

Methods First, we explore the novel application of a visual categorical perception methodology to assess a clinicians perceptual sensitivity to detect speech motor control issues. Specifically, we evaluated their ability to categorize lateral jaw deviations in children. We tested 32 S-LPs with two different levels of clinical experience (novice N =19; experienced N =13). The stimuli comprised of a linearly spaced continuum of 11 images of a 7-year old child producing a target word with a lateral jaw slide. Frame 1 had no lateral jaw deviation (0 radians) while frame 11 showed a maximum jaw deviation (0.26 radians). Participants performed a standard alternative forced choice identification procedure and an ABX discrimination task using this 11 image stimuli set presented in random order. Second, we utilize kinematic consistency measures (cyclic Spatio-Temporal Index (cSTI); Generalized Orthogonal Procrustes Analysis (GPA)) as a means to assess an S-LPs consistency in the delivery of tactile-kinesthetic-proprioceptive (TKP) inputs. Participants comprised of three S-LPs with different experience levels and one female adult participant who served as a model. All three S-LPs delivered TKP inputs to the model participants speech articulators while kinematic data of mouth and finger movements were collected using an Articulograph system (AG501). Variability of the S-LPs finger-hand movement trajectories in 3-D space (GPA) during delivery of these TKP inputs and the consistency (cSTI) of the resulting lip displacements were extracted from these data.

Results & Discussion For the visual categorical perception experiment, we found a category boundary for both the less experienced and the more experienced S-LPs, in support of a categorical perception mechanism. The location of the boundary differed between the two groups suggesting a greater sensitivity towards detecting smaller lateral jaw deviations in the more experienced clinicians. Kinematic consistency measures indicated that more experienced clinicians had lower intrinsic variability for their finger-hand movement trajectories and delivered TKP inputs with greater consistency relative to the clinician with less experience. The findings will be discussed in terms of treatment fidelity, specifically with respect to tracking a S-LPs skill level as a function of training and clinical experience.

Further Reading

Hayden, D., Namasivayam, A. K., & Ward, R. (2015). The Assessment of Fidelity in a Motor Speech Treatment Approach. *Speech, Language and Hearing, 18*(1), 30-38.

VOICE QUALITY ANALYSIS OF CHILDREN WITH CEREBRAL PALSY DURING SUSTAINED PHONATION AND STORY RETELL

Ignatius S. B. Nip[1], Marc Garellek[2]
[1]*San Diego State University, San Diego, United States of America*
[2]*University of California - San Diego, San Diego, United States of America*

Introduction Many children with cerebral palsy (CP) have dysarthria (Nip, 2017) that is frequently accompanied by voice quality changes. Typically, voice quality differences are captured via rating scales (e.g., CAPE-V); however, they can also be assessed using psychoacoustic measures of the voice source (Kreiman et al., 2014), which are associated with changes in both voice quality perception and their articulatory origins. This information provides insights into the laryngeal impairments causing the voice quality, and can be used to better inform how CP affects the speech mechanism in these children and to better document therapeutic changes. The current investigation examines how acoustic measures in voice quality differ between children with CP and their typically-developing age- and sex-matched peers.

Method Eight children with CP (M=9.43 years; range=4.67-12.33 years, 5M) and eight typically-developing (TD) age- and sex-matched peers (M=9.36 years; range=4.58-13.17 years; 5M) were recorded producing sustained /i, a, u/ vowels and a story retell. Vowels [i, a, u] were extracted from both tasks using PRAAT.

Two measures of voice quality were obtained using VoiceSauce (Shue et al., 2011): H1-A2 (a measure of harmonic spectral tilt, lower values = more constricted or irregular phonation) and CPP (a noise measure, lower values = breathier or more irregular vocal fold vibration).

Two 2x2x3 multilevel models were run for each dependent variable (H1-A2, CPP) with Group (CP, TD), Task (Vowel, Story), and Vowel ([i], [a], [u]) as the independent variables with participants as the repeated factor. Age was used as a covariate for both models.

Results & Discussion For both the vowels in isolation and story re-tell task, the CP group had significantly lower H1-A2 values (i.e., more constricted voice quality) for [a, i] but not [u]. CPP values were overall significantly greater (i.e., more regular or less noisy) for the TD group. The story task had higher CPP values than the isolated vowels for the CP group only. Therefore, findings indicate that both harmonic (H1-A2) and inharmonic (CPP) measures of the voice source can differentiate children with CP from their TD peers. In general, and especially in sustained vowel conditions, children with CP produce vowels with lower spectral tilt and greater noise, suggesting that they use a creakier or rougher voicing characterized by increased vocal fold constriction and noise.

References

Kreiman, J., Gerratt, B. R., Garellek, M., Samlan, R., and Zhang, Z. (2014). Toward a unified theory of voice production and perception. *Loquens, 1*, e009. http://dx.doi.org/10.3989/loquens.2014.009

Nip, I. S. B. (2017). Interarticulator coordination in children with and without cerebral palsy. *Developmental Neurorehabilitation, 20*, 1-13. http://dx.doi.org/10.3109/17518423.2015.1022809

Shue,Y.-L., Keating, P.A., Vicenik, C., & Yu, K. (2011). VoiceSauce: A program for voice analysis. In *Proceedings of the International Congress of Phonetic Sciences* (pp. 1846-1849), Hong Kong.

SUBCORTICAL CONTRIBUTIONS TO AUDITORY-MOTOR INTEGRATION IN SPEECH

Mona Späth[1], Ingrid Aichert[1], Dagmar Timmann[2], Andrés O. Ceballos-Baumann[2], Edith Wagner-Sonntag[3], Wolfram Ziegler[1]
[1]*Ludwig-Maximilians-Universität München, Munich, Germany*
[2]*University of Duisburg-Essen, Essen, Germany*
[3]*Schön Klinik München Schwabing, Munich, Germany*

Background The integration of auditory information plays an important role in speech motor control. Although subcortical structures are assumed to contribute to these processes (Guenther, 2016), they have hardly been investigated in this context. The current study used a covert adaptation experiment to examine auditory-motor integration in individuals with dysarthria due to a dysfunction of the cerebellum (spinocerebellar ataxia type 6, SCA6, a relatively pure form of cerebellar degeneration), and the basal ganglia (Parkinsons disease). A focus will be placed on rhythmic auditory input, as both the cerebellum and the basal ganglia are discussed to be implicated in time and rhythm processing (Ivry & Spencer, 2004; Kotz & Schmidt-Kassow, 2015).

Methods Twelve speakers with ataxic dysarthria due to SCA6, 15 speakers with hypokinetic dysarthria due to Parkinson's and 27 age- and gender-matched healthy controls participated in this study. Speech materials comprised 24 metrically regular sentences (consisting of five trochaic words) and 24 metrically irregular sentences (consisting of five alternately trochaic/iambic words). Model sentences produced by a professional speaker were manipulated using Praat to generate a speech rate continuum (2.9 - 4.0 syllables/sec) and an intensity continuum (64.0 - 75.0 dB). The 48 target sentences were elicited in three different tasks: (a) reading, (b) repetition, (c) turn-taking. Additional tasks comprised (a) a perception task to check whether minor changes in speech rate and intensity in the production experiments could be detected, (b) dysarthria assessment, and (c) cognitive assessment.

The 7776 sentences (54 [participants] x 48 [sentences] x 3 [conditions]) were analysed auditorily and acoustically. Acoustic parameters included speech rate, intensity, rhythm and pauses, as well as speech initiation latency.

Results Data analysis is still incomplete. Data sets from 15 patients with Parkinsons, seven patients with SCA6 and 15 healthy controls so far suggest that individuals with Parkinsons and controls systematically adapted the speech rate of the model speaker, while individuals with SCA6 did not. Intensity was not imitated by any of the groups.

Discussion First analyses suggest that auditory-motor integration as measured by the present adaptation paradigm is impaired in individuals with cerebellar dysfunction, but spared in striatal dysfunction.

References

Guenther, F. H. (2016). *Neural control of speech.* Cambridge: The MIT Press.

Ivry, R. B., & Spencer, R. M. C. (2004). The neural representation of time. *Current Opinion in Neurobiology, 14*, 225-232.

Kotz, S. A., & Schmidt-Kassow, M. (2015). Basal ganglia contribution to rule expectancy and temporal predictability in speech. *Cortex, 68*, 48-60.

EVALUATING ASPECTS OF SPEECH MOTOR STABILITY IN DYSARTHRIA

Frits van Brenk, Anja Lowit
University of Strathclyde, Glasgow, United Kingdom

Dysarthria may manifest in respiratory, phonatory, resonatory, articulatory and prosodic levels of speech production. One way to assess motor control in these areas is to estimate the stability of movement patterns. Previous research has demonstrated changes in motor variability in disordered speakers based on kinematic measures. Our lab has furthermore shown that acoustically based variability measures also have the capability to distinguish between speaker groups, and correlate with clinical assessments (van Brenk & Lowit, 2012).

The aim of this study was to further investigate the potential of acoustic measures to characterize aspects of speech motor stability in dysarthria. A range of speaking conditions and acoustic parameters were evaluated in order to provide a better understanding of the neuromotor factors affecting performance in dysarthria, and to evaluate to what degree the variability estimators hold their promise in being valuable for clinical research.

Participants were 23 speakers with hypokinetic dysarthria, 8 speakers with ataxic dysarthria, and 27 age-matched control speakers. The phrase "Tony knew you were lying in bed" was repeated around twenty times in six speaking conditions varying in rate, complexity and motor load. Contours of sound pressure level, fundamental frequency, and first and second formants were extracted from the acoustic signal, and subjected to functional data analysis (Ramsay et al., 1996) to obtain measures of spatial and temporal variability. The spatiotemporal index was calculated as an additional measure. Furthermore, standardized clinically based measures of speech motor performance and quantifiable details of treatment history were obtained.

We report on the potential of using acoustic variability measures in the assessment of dysarthria, guided by two data treatment procedures. The interplay between all obtained measures was explored by means of a factor analysis used as a data reduction method to detect possible relations. In addition, a logistic regression analysis was performed in which the various measurements were used as predictors and the presence of dysarthria as outcome measure. Based on the results of these analyses we aim to provide guidelines on the optimal selection of speech tasks and acoustic parameters.

References

van Brenk, F., & Lowit, A. (2012). The relationship between acoustic indices of speech motor control variability and other measures of speech performance in dysarthria. *Journal of Medical Speech Language Pathology, 20*(4), 24-29.

Ramsay, J. O., Munhall, K. G., Gracco, V. L., and Ostry, D.J. (1996). Functional data analysis of lip motion. *Journal of the Acoustical Society of America, 99*(6), 3718-3727.

SPEECH TIMING IN CHILDREN WITH AND WITHOUT SPEECH SOUND DISORDERS: ACOUSTIC MEASURES OF ABSOLUTE AND RELATIVE TIMING ACCURACY

Anne van Zelst, Edwin Maas
Temple University, Philadelphia, United States of America

A substantial proportion of children have difficulty acquiring speech production skills (McLeod & Harrison, 2009). For some children with speech sound disorders (SSD) the underlying deficit is in speech motor skill (e.g., childhood apraxia of speech, CAS). We investigated *temporal* control of speech movement sequences in children with and without SSD, including CAS and phonological disorder (PD). Specifically, we examined absolute and relative timing measures derived from the motor control literature (e.g., Shea, Wulf, Park, & Gaunt, 2001). The long-term goal is to better characterize SSD and develop specific and sensitive diagnostic measures.

Thirty-five speakers repeated six phrases with CVC nonwords after a prerecorded model (e.g., “Its the shoob again”) across three blocks. Acoustically-defined vowel and consonant intervals were used to derive two dependent measures, E and PROP (Shea et al., 2001). E is a measure of absolute timing error that combines bias (constant error, CE) and variability (variable error, VE) according to the formula E=$\sqrt{[(CE^2)+(VE^2)]}$. PROP is a measure of relative timing (temporal structure) error, and is computed as PROP=Σ —actualPROP$_i$-targetPROP$_i$—, where PROP$_i$ is the proportion of interval i relative to total phrase duration.

Data analysis has been completed for one adult, two typically-developing 6-year-old (TD6) children, and two children with SSD (1 CAS, 1 PD). TD6 children exhibit numerically greater error than adults for both E and PROP. SSD children do not exhibit overall greater error than TD6 children. However, unlike TD6 children, both SSD children show a change in error across blocks: The child with PD shows a decrease, whereas the child with CAS shows an *increase* for both E and PROP.

Preliminary findings suggest that TD children are less accurate than adults for both absolute and relative timing. Both SSD children showed a change in error across blocks, with rapid improvement for the child with PD and the opposite pattern for the child with CAS. This suggests the potential importance of dynamic measures for differential diagnosis (e.g., Terband, Maassen, & Maas, 2016). [Work supported by NIH-NIDCD (K01-DC010216)].

References

McLeod, S., & Harrison, L. J. (2009). Epidemiology of speech and language impairment in a nationally representative sample of 4-to 5-year-old children. *Journal of Speech, Language, and Hearing Research, 52*, 1213-1229.

Shea, C. H., Wulf, G., Park, J. H., & Gaunt, B. (2001). Effects of an auditory model on the learning of relative and absolute timing. *Journal of Motor Behavior, 33*, 127-138.

Terband, H., Maassen, B., & Maas, E. (2016). Towards a model of pediatric speech sound disorders (SSD) for differential diagnosis and therapy planning. In P. van Lieshout, B. Maassen, & H. Terband, (Eds.), *Speech Motor Control in Normal and Disordered Speech: Future Developments in Theory and Methodology* (pp. 81-110). Rockville, MD: American Speech-Language-Hearing Association.

MOTOR FUNCTION IN THE SERVICE OF ACOUSTIC COMMUNICATION: CLINICAL ASSESSMENT OF DYSARTHRIA

Wolfram Ziegler, Theresa Schölderle, Anja Staiger
Ludwig-Maximilians-Universität, Munich, Germany

Background & objectives A preeminent property of speech motor control is that speaking is in the service of communication. Therefore, a litmus test for the construct validity of clinical assessments of motor speech disorders is whether they measure functions that characterize speech as a motor activity serving communicative goals. The present study examined conventional speech and nonspeech assessment methods (Duffy, 2013) regarding the degree to which they permit appropriate inferences on speech impairment in dysarthria.

Methods
Participants 130 patients with neurogenic movement disorders. 130 healthy controls.
Procedure Participants underwent speech and nonspeech tests of respiratory, laryngeal, and vocal tract motor control: Supported text reading, word- and sentence repetition, rapid syllable repetition, vowel prolongation, single articulator movements, and rapid alternating articulator movements. Speech analyses were performed according to the BoDyS protocol (Ziegler et al., in press). Intelligibility and naturalness were assessed in all participants.

Results Vowel prolongation and rapid articulator movements disqualified because of substantial overlaps between healthy controls and patients with even moderate-to severe speech impairment. Accuracy of single articulator movements was uncorrelated with its closest correspondents, i.e., perceived accuracy of articulation and intelligibility. In split-half analyses, mixed linear regression models revealed that only speech-related measures made significant contributions to the modeling and prediction of intelligibility or naturalness.

Discussion The results corroborate earlier findings of a dissociation between speech and nonspeech measures of vocal tract motor control in dysarthria (Staiger et al., 2016). Articulation and prosody scales were strongly related with naturalness ratings and intelligibility. The validity of nonspeech tasks has still to be proven.

References
Duffy, J. R. (2013). *Motor Speech Disorders: Substrates, Differential Diagnosis, and Management* (3rd ed.). St. Louis: Elsevier Mosby.

Ziegler, W., Staiger, A., Schölderle, T., & Vogel, M. (in press). Gauging the auditory dimensions of dysarthric impairment: Reliability and construct validity of the Bogenhausen Dysarthria Scales (BoDyS). *Journal of Speech, Language, and Hearing Research.*

Staiger, A., Schölderle, T., Brendel, B., Bözel, K., & Ziegler, W. (2016). Task-dependent control of oral-motor timing: A factor analytic approach. *Journal of Motor Behavior.* doi:10.1080/00222895.2016.1241747

Poster session III

ULTRAPHONIX: USING ULTRASOUND VISUAL BIOFEEDBACK TO TEACH CHILDREN WITH SPEECH SOUND DISORDERS NEW ARTICULATIONS

Joanne Cleland[1], James M. Scobbie[2], Zoe Roxburgh[2], Cornelia Heyde[2], Alan Wrench[3]
[1] *University of Strathclyde, Glasgow, United Kingdom*
[2] *Queen Margaret University, Edinburgh, United Kingdom*
[3] *Articulate Instruments Ltd, Musselburgh, United Kingdom*

Background Ultrasound Visual Biofeedback (U-VBF) provides learners with "knowledge of performance" (Maas et al., 2008) that enables them to establish new, more accurate motor programmes for previously in-error speech sounds. The evidence for U-VBF is promising, with around 30 small studies. Although more recent studies differentiate acquisition, retention and generalisation of target articulations (see Sjolie et al., 2016) no studies report in detail when children who begin intervention not stimulable for a particular sound are first able to produce that new articulation. This is an important issue for establishing the dosage of U-VBF. The current study is an investigation of the effectiveness of U-VBF, with particular attention paid to when and how children acquire new articulations.

Participants Fifteen children aged 6 to 15 with SSDs affecting vowels and/or lingual consonants in the absence of structural abnormalities.

Method Single-subject multiple (3) baseline design, with untreated wordlists. Ultrasound was recorded synchronously with audio and used to review dynamic information about the childrens speech errors for diagnostic purposes. Each child received 10-12 sessions of U-VBF with each child required to perform at 80% accuracy at each level of performance before moving on to a motorically more demanding level (for example, from single syllable words to disyllabic words).

Analysis Narrow transcription of wordlists was undertaken and percentage targeted segments correct calculated. Prior to intervention the ultrasound data was analysed both qualitatively and quantitatively to identify errors.

Results Six children were treated for velar fronting; three for post-alveolar fronting; three for the unusual pattern of backing to pharyngeal or glottal; one for production of all syllable onsets as [h]; one for vowel merger and one for lateralised sibilants. Most children (10/15) achieved the new articulation in the first or second session. Four children took until the 6th to 9th session to achieve the new articulation and one never did. Those children who acquired the new articulation earlier in the therapeutic process were able to integrate that new articulation into words and sentences more quickly and generalise to untreated words more successfully. 13/15 children made improvements of more than 20 percentage points increase in the accuracy of targeted segments in untreated wordlists. One child made no improvement and one moved towards a phonetically closer approximation of the target.

References

Maas, E., Robin, D. A., Hula, S. N. A., Freedman, S. E., Wulf, G., Ballard, K. J., & Schmidt, R. A. (2008). Principles of motor learning in treatment of motor speech disorders. *American journal of speech-language pathology, 17*(3), 277-298.

Sjolie, G. M., Leece, M. C., & Preston, J. L. (2016). Acquisition, retention, and generalization of rhotics with and without ultrasound visual feedback. *Journal of Communication Disorders, 64*, 62-77.

INTRODUCING INTER-SYLLABIC INTERVAL AS AN INDICATOR OF (DIS)FLUENCY IN FRENCH

Véronique Delvaux[1,2], Myriam Piccaluga[1], Kathy Huet[1], Bernard Harmegnies[1]
[1]*Université de Mons, Mons, Belgium*
[2]*Fonds National de la Recherche Scientifique, Belgium*

Fluent speech is characterized by appropriate tempo and reduced number of disfluencies. Those disfluencies are mainly silent and full pauses, false starts, and repetitions of sentences and polysyllabic words, signalling issues in linguistic planning and/or monitoring, in relation with cognitive and psycho-affective factors. Other disfluencies typify stuttered speech (although they may also happen in healthy speakers) in association with speech motor control deficits, including single syllable and part word repetitions, and disrhythmic phonations (blocks, prolongations). Other speech motor disorders, typically dysarthrias, are charaterized by a slower articulation rate resulting from impaired motor coordination and reduced speed of articulation.

Thus, the concepts of rhythm and tempo (articulation rate, speech rate) on one hand, and fluency/disfluencies on the other hand, interrelate in complex ways. One of the consequences is the methodological challenge of setting measures and indicators: Which (types of) disfluencies to include/exclude before computing articulation rate? How to select tresholds to define silent pauses and prolongations?

Our team has recently introduced the Inter-Syllabic Interval (ISI) as a potential indicator of fluency in French-speaking healthy and pathological speakers (Piccaluga et al., 2007). ISI is defined as the time interval between intensity peaks characterizing consecutive syllabic nuclei in any spoken utterance, free from the necessity to a priori exclude or categorize portions of the speech signal as potential disfluencies. Mean ISI is close to the notion of speech rate: low ISI signals quick tempo and/or short, few silent pauses while high ISI is expected in dysarthric speech. The overall and time-evolving variance of ISI should allow to discriminate between fluent and disfluent speech, and between different types of disfluencies according to the degree to which they affect syllabic rythm.

The aim of our research programme is to assess the potential of ISI as an indicator of (dis)fluency in French with respect to: other acoustic measures based on conventional syllabic segmentation; other physiological and articulatory measures; subjective assessments from expert and naive listeners. Targeted populations are healthy speakers, PWS and parkinsonians performing speech production tasks with and without time/cognitive pressure.

We present here a first acoustic study performed on 400 healthy francophone speakers, half male half female, of all ages and four dialectal origins. Data were collected using the "MonPaGe" protocol (Fougeron et al., 2016; in this case, the picture description and text reading tasks). We describe results as a function of age, gender and origin, comparing in terms of robustness and cost/benefit ratio statistics on EIS with more conventional measures requiring the a priori detection and classification of disfluencies.

References

Fougeron C., Delvaux V., Pernon M., Lévque N., Borel S., Pellet P., Bagou O., Trouville R., Ménard L., Catalano S., Lopez U., Kocjancic-Antolik T., & Laganaro M. (2016). MonPaGe: un protocole informatisé dévaluation de la parole pathologique en langue franaise. *Actes des 26e Rencontres Internationales dOrthophonie*, Paris, France.

Piccaluga M., Nespoulous J-L., & Harmegnies B. (2007). Disfluency surface markers and cognitive processing; the case of simultaneous interpreting. *Proceedings of the 16th International Congress of the Phonetic Sciences*, Saarbrücken, Germany, 1317-1320.

MAXIMUM REPETITION RATE: NORMATIVE DATA AND CLINICAL USE IN CHILDREN WITH SPEECH SOUND DISORDERS

Sanne Diepeveen[1] Leenke van Haaften[2], Hayo Terband[3], Bert de Swart[1,2], Ben Maassen[4]
[1]*HAN University of Applied Sciences, Nijmegen, The Netherlands*
[2]*RadboudUMC Nijmegen, Nijmegen, The Netherlands*
[3]*Utrecht University, Utrecht, The Netherlands*
[4]*University of Groningen, Groningen, The Netherlands*

Introduction There are only few standardized, normative tools for assessing motor speech functions in children with speech sound disorders (SSD) as compared to typically developing children (Wong, et. al., 2011). We have developed a protocol to assess Maximum Repetition Rate (MRR), which is implemented in the Computer Articulation Instrument (CAI). The aims of the current study were to collect normative, developmental data for the MRR task, and to determine its clinical use in diagnosing children with SSD.

Method A sample of 1524 typically developing children aged between 2;0 and 7;0 years was collected, representative for gender, geographic region and degree of urbanization in the Netherlands. We defined the MRR scores as the number of syllables uttered per second in mono-, bi- and trisyllabic sequences (Thoonen, et. al., 1996). Regression procedures were applied to transform raw scores into continuous age-normalized percentile scores.

Twenty-two children with severe speech- and language difficulties were compared with 96 matched children from the normative sample.

Results Normative data: Under the age of three years, children are unable to perform the MRR-task. In the age range 3 to 7 years, 80% of the children can produce the sequences, and show a steady increase of rate.

Comparison of normative and SSD data: Significant differences were seen for *pa* ($p = .001$), *ka* ($p = .008$); not significantly different were *ta* ($p = .234$, *pata* ($p = .169$), *taka* ($p = .081$), *pataka* ($p = .202$).

Discussion The normative data is comparable with the data of Modolo, et. al. (2011). Further studies of the clinical use of MRR for the differential diagnosis of children with SSD are currently conducted.

References

Modolo, D. J., Berretin-Felix, G., Genaro, K. F., & Brasolotto, A. G. (2010). Oral and vocal fold diadochokinesis in children. *Folia Phoniatrica Et Logopaedica, 63*(1), 1-8.

Thoonen, G., Maassen, B., Gabreels, F., & Schreuder, R. (1999). Validity of maximum performance tasks to diagnose motor speech disorders in children. *Clin Linguist Phon, 13*(1), 1-23.

Wong, A. W., Allegro, J., Tirado, Y., Chadha, N., & Campisi, P. (2011). Objective measurement of motor speech characteristics in the healthy pediatric population. *Int J Pediatr Otorhinolaryngol, 75*(12), 1604-1611. doi:10.1016/j.ijporl.2011.09.023

SPONTANEOUS SPEECH OF PARKINSON'S PATIENTS: AN INVESTIGATION OF PAUSING BEHAVIOUR AND LINGUISTIC CRITERIA

Susanne Ebert[1], Anja Lowit[2], Ulrike Frank[1]
[1] *Universität Potsdam, Potsdam, Germany*
[2] *University of Strathclyde, Glasgow, United Kingdom*

Background Lately, research in Parkinsons Disease (PD) focused on possible language impairments besides dysarthric symptoms (Illes, Metter, Hanson, & Iritani, 1988). Nevertheless, investigations on spontaneous speech production are still the exception (Huber, Darling, Francis & Zhang, 2012; Murray & Lenz, 2001).

Aim This bachelor thesis investigated spontaneous speech in PD patients with regard to pausing behaviour and linguistic criteria in comparison to healthy controls (HC) speech samples.

Method Twenty-nine audio samples of PD patients were transcribed and analysed. In 19 of these recordings a reliable orthographic transcription was possible (group PD: M_{age}=64.9, SD=6.8; severity Hoehn & Yahr Scale M=2.14, SD=0.78). Intelligibility of the remaining 10 patients was limited (group PDsev., M_{age}=67, SD=5; severity Hoehn & Yahr Scale M=3.38, SD=1.16). Consequently, a reduced analysis was conducted. An age and gender matched control group was analysed ($\mathrm{M_{age}}$=67.9, SD=5.9). The following criteria were examined: (1) Pausing behaviour: number and length of silent and filled pauses, type of pause; (2) Linguistic criteria: MLU, number of complete clauses/utterance, proportion of subordinate clauses, ungrammatical sentences, disfluencies and mazes, abortion of sentences.

Results Groups did not differ in the amount and type of pauses produced. Interesting differences showed the post-hoc analysis of silent and filled pauses: Both PD groups showed longer silent pauses than controls. The length of filled pauses showed no difference between controls and less severe PD patients, nevertheless defined the severe PD group. Linguistic criteria did not differ between the PD group and controls, except for the number of disfluencies.

Conclusion In contrast to the literature, this investigation of linguistic criteria showed only few differences between PD patients and controls. Analysis of pausing behaviour could hint towards silent pauses length being an indicator for PD itself, whereas filled pauses length being a possible severity indicator.

References

Huber, J. E., Darling, M., Francis, E. J., & Zhang, D. (2012). Impact of Typical Aging and Parkinson's Disease on the Relationship Among Breath Pausing, Syntax, and Punctuation. *American Journal of Speech-Language Pathology, 21*(4), 368-379.

Illes, J., Metter, E. J., Hanson, W. R., & Iritani, S. (1988). Language production in Parkinson's disease: Acoustic and linguistic considerations. *Brain and Language, 33*(1), 146-160.

Murray, L. L., & Lenz, L. P. (2001). Productive syntax abilities in Huntington's and Parkinson's diseases. *Brain and Cognition, 46*(1-2), 213-219.

TREATMENT OF ACQUIRED APRAXIA OF SPEECH USING REPETITION OF COMPLEX PSEUDOWORDS - A SINGLE CASE STUDY

Franziska Gerzer, Christoph Birke, Judith Heide, Maria Blickensdorff
University of Potsdam, Potsdam, Germany

Introduction Specificity or skill acquisition is one of the principles of experience-dependent neural plasticity that is likely to be relevant to rehabilitation after brain damage (Kleim & Jones, 2008). We report a single-case intervention that was designed to enhance articulatory accuracy of words in a patient demonstrating chronic moderate acquired Apraxia of Speech (AOS) and non-fluent aphasia after stroke. According to the patients individual profile of speech performance, a task of maximal articulatory difficulty was administered in order to ensure specificity of treatment. A range of Principles of Motor Learning (Maas, Robin, Austermann Hula, Freedman, Wulf, Ballard, & Schmidt, 2008) was integrated into the therapy design.

Methods & Materials A multiple baseline control-task design was used to investigate pre-treatment stability of performance, treatment effects and generalization to untreated items (Stadie & Schröder, 2009). Eight therapy sessions of 90 minutes were conducted.

The item set for treatment consisted of 14 three- and four-syllabic pseudowords that were derived from the generalization items and seven three- and four-syllabic foreign language words all with initial CC.- or CCC.-clusters. The item set for generalization consisted of 21 three- and four-syllabic words with word- initial CC.- or CCC.-cluster.

Each therapy session started with a pre-practice phase based on word repetition. Rhythmic cueing techniques, integral stimulation techniques and the metric treatment method were offered to facilitate correct responses.

In the practice phase, a board game was used for word repetition in random order. During the practicephase low frequency feedback on knowledge of results was provided.

Results Improved word production measured as a reduction of phonologic errors and phonetic distorsions in word-initial consonant clusters was observed for treated items ($p<.001$, McNemar Test) and generalization items ($p<.01$, McNemar Test). Improvement transferred to unrelated one-syllabic and two-syllabic words with word-initial CC.-clusters ($p<.05$, McNemar Test). Control task performance was unaltered ($p>.05$ McNemar Test).

Conclusion A specific therapy task using complex pseudoword stimuli in combination with the Principles of Motor Learning were applied to improve articulatory accuracy in a patient with AOS.

References

Kleim, J., A., & Jones, T.A. (2008). Principles of experience-dependent neural plasticity: Implications for rehabilitation after brain damage. *Journal of Speech, Language, and Hearing Research, 51*, S225-S239. doi:10.1044/1092-4388(2008/018)

Maas, E., Robin, D. A., Austermann Hula, S. N., Freedman, S. E., Wulf, G., Ballard, K. J., & Schmidt, R. A. (2008). Tutorial: Principles of motor learning in treatment of motor speech disorders. *American Journal of Speech -Language Pathology, 17*, 277 -298. doi:10.1044/1058- 0360(2008/025)

Stadie, N., & Schröder, A. (2009). *Kognitiv orientierte Sprachtherapie: Methoden, Material und Evaluation für Aphasie, Dyslexie und Dysgraphie.* München: Elsevier, Urban&Fischer Verlag

DURATION OF SHORT SILENT INTERVALS IN THE CONNECTED SPEECH OF INDIVIDUALS WITH PARKINSON DISEASE

Anna C. Gravelin, Jason A. Whitfield
Bowling Green State University, Bowling Green, United States of America

The duration of silent intervals in connected speech is associated with articulatory, prosodic, and/or linguistic processes. A wide range of definitions have been used to examine pause, with some authors defining a pause as a silent interval lasting at least 10 ms and other authors defining pause as a silent interval lasting at least 250 ms. Because speakers with PD exhibit pause abnormalities, several authors have examined the extent to which respiratory, articulatory, prosodic, and linguistic deficits contribute to pause abnormalities in this population. The purpose of this study was to examine the duration of short silent intervals in the connected speech of individuals with and without PD that reflect articulatory processes. Reading samples from 10 individuals with idiopathic PD (5 males, 5 females) and 10 older control speakers (5 males, 5 females) were analyzed. The speakers with PD presented with hypokinetic dysarthria, which ranged from mild to severe with seven speakers falling in the mild to moderate range. Silent intervals longer than 15 ms were labeled in PRAAT using spectrographic and waveform displays as a guide. Each silent interval was categorized according to its syntactic and phonetic context, and whether it coincided with a visually identifiable inspiratory breath. For the current study, within-word intervals and between-word intervals that were not associated with a phrase or clause boundary, occurred in the middle of a phrase, and were not associated with a disfluency or inspiration were examined. No between-group differences in speech rate, syllable number, or total speaking duration were observed, $p>0.05$ for all comparisons. The duration of the silent intervals included in the analysis ranged from 15 to 200 ms. On average, 61.40, $SD=18.83$, silent intervals were identified for speakers in the PD group and 69.70, $SD=11.15$ for the control speakers. Mixed-model analysis revealed that the average duration of these silent intervals was slightly, but significantly longer for the PD group, $Estimate=6.455$ ms, $SE=3.021$, $t=2.137$, $p<0.05$. Examination of the phonetic context indicated that the majority of these short silent intervals were stop gaps. Results of this study suggest that short silent intervals that are unrelated to linguistic or prosodic emphasis or an inspiratory breath are slightly longer for speakers with PD, than for healthy controls. This slightly longer duration between speech segments may result from timing deficits associated with basal ganglia dysfunction secondary to PD. These deficits may reflect a subtle disruption in speech fluency or articulation.

References

Huber, J. E., Darling, M., Francis, E. J., & Zhang, D. (2012). Impact of typical aging and Parkinsons disease on the relationship among breath pausing, syntax, and punctuation. *American Journal of Speech-Language Pathology, 21*(4), 368-379.

Skodda, S., & Schlegel, U. (2008). Speech rate and rhythm in Parkinson's disease. *Movement Disorders, 23*(7), 985-992.

ELICITING SPEECH FROM CHILDREN WITH DYSARTHRIA: A NEW APPROACH FOR CLINICAL ASSESSMENT

Elisabet Haas, Wolfram Ziegler, Theresa Schölderle
Ludwig-Maximilians-Universität, Munich, Germany

Background & Aim Childhood dysarthria often occurs in the context of multiple disabilities. In clinical assessment, speech therapists therefore face the challenge to elicit standard speech samples from children with complex motor, sensory, or cognitive impairment (Morgan & Liégeois 2010). Even though the number of affected children is considered high, there have been hardly any attempts to develop specific, age-appropriate elicitation methods or assessment materials (Patel & Connaghan 2014). The common approach to apply test materials established for adults (e.g., including reading tasks), however, seems hardly feasible. The aim of the current study was therefore to develop a child-appropriate elicitation approach and to evaluate its feasibility.

Methods We introduce a tool that uses sentence repetition to systematically assess motor speech function in children. For motivational reasons, the tool was implemented as a computer game with pictures that are appealing to children and with an exciting framework story connecting the pictures. To cope with the special needs of most children with dysarthria, maximum support is offered to facilitate sentence repetition (e.g. sentences are linked semantically to the pictures; stimuli can be repeated multiple times).

The auditory prompts of the 12 repetition sentences are integrated into the game, which allows for a standardized presentation. Since the elicited speech samples provide the basis of comprehensive speech analyses (see submission Schölderle et al.), the sentences were constructed by systematically controlling relevant phonetic parameters (e.g., length, intonation pattern).

In order to evaluate feasibility, 6- to 7-year-old children with dysarthria and multiple disabilities as well as an age-matched control group (n=7 each) were assessed. Feasibility was measured (1) by clinical-practical aspects (e.g., test duration) and (2) regarding the major aim to elicit informative standard speech samples (e.g., number of incorrect repetitions).

Results We successfully elicited the complete set of sentences in all children. Testing time was 15-20 minutes. No pauses were needed. Only 17% of the sentences were repeated incorrectly (including e.g., word substitutions). A reliability analysis confirmed that all sentences were produced with a consistent speech rate (Cronbach's Alpha α=0,98), suggesting that the sentences were homogeneous with respect to their cognitive demands.

Discussion This study demonstrates that standard speech samples can be elicited from dysarthric children with multiple disabilities if test materials are adapted to the childrens age and needs. This is an important requirement for further investigation and in the clinical assessment of childhood dysarthria.

References

Morgan, A. T. & Liégeois, F. (2010). Re-thinking diagnostic classification of the dysarthrias: a developmental perspective. *Folia Phoniatrica et Logopaedica, 62*(3), 120-126.

Patel, R. & Connaghan, K. (2014). Play Park: a picture description task for assessing childhood motor speech disorders. *International Journal of Speech-Language Pathology, 16*(4), 337-343.

GAME-BASED SPEECH THERAPY USING VISUAL FEEDBACK IN PARKINSON'S DISEASE

Elaine Kearney[1,2], Brandon Haworth[2,3], Jordan Scholl[1], Petros Faloutsos[2,3], Melanie Baljko[2,3], Yana Yunusova[1,2]

[1]*University of Toronto, Toronto, Canada*

[2]*University Health Network - Toronto Rehabilitation Institute, Toronto, Canada*

[3]*York University, Toronto, Canada*

Background Parkinsons disease (PD) is a progressive neurodegenerative disease affecting voluntary movement, including speech movements (Schulz & Grant, 2000). 45% of individuals with PD experience difficulties with articulation, leading to reduced intelligibility (Logemann, Fisher, Boshes, & Blonsky, 1978). Treatment options that address the underlying movement disorder (i.e. hypokinesia) are limited. A recently developed augmented visual feedback system showed promise in training larger tongue movements following a single training session (Yunusova et al., in press).

Purpose An ABA single-subject design study was conducted to establish the preliminary efficacy of a treatment program using game-based visual feedback on speech movements and intelligibility in PD.

Method Five males (mean age = 75, SD = 8.79) diagnosed with PD attended a 10-session program over 5 weeks, aimed at increasing the size of tongue movements during sentence production. Each session involved training 5 sentences, in an acquisition and test phase. During acquisition, participants were instructed to use large tongue movements and were provided with visual feedback relating to the size of their tongue movement for all trials. Feedback frequency was reduced to 50% during the test phase. Tongue movements were tracked using the Wave Speech Research System, a 3D electromagnetic articulography. Movement size was indexed via Articulatory Working Space (AWS). In addition, measures of tongue movement and speech intelligibility for both trained and untrained sentences were taken pre- and post- treatment. Response to treatment was examined by visual analysis, while treatment effect and generalization were assessed using Cohens d effect sizes.

Results 4/5 participants increased tongue movements during therapy sessions. Posttreatment, 4/5 participants demonstrated large effect sizes for trained sentences ($d > 1.0$), and two participants generalized the effect of treatment to untrained sentences ($d > 1.0$). Analysis of corresponding intelligibility data is ongoing.

Conclusion This study demonstrated that movement-based therapy using augmented visual feedback may facilitate the treatment of speech movement impairment in PD. Further work that manipulates motor learning variables, such as treatment intensity and feedback frequency, may be useful in enhancing generalization of treatment effects.

References

Logemann, J. A., Fisher, H. B., Boshes, B., & Blonsky, E. R. (1978). *Journal of Speech and Hearing Disorders, 43*(1), 47-57.

Schulz, G. M., & Grant, M. K. (2000). *Journal of Communication Disorders, 33*(1), 59-88.

Yunusova, Y., Kearney, E., Kulkarni, M., Haworth, B., Baljko, M., & Faloutsos, P. (in press). *Journal of Speech, Language, and Hearing Research.*

ARTICULATORY KINEMATICS DURING STOP CLOSURE IN SPEAKERS WITH PARKINSON'S DISEASE

Yunjung Kim[1], Jeffrey Berry[2], Christina Kuo[3]
[1]*Louisiana State University, Baton Rouge, United States of America*
[2]*Marquette University, Milwaukee, United States of America*
[3]*James Madison University, Harrisonburg, United States of America*

Background Little is known about lingual movements during stop closure although several studies have reported on inter-articulator interactions (Löfqvist & Gracco, 1997) as well as the extent and nature of tongue movement during the brief interval. While these studies typically have a small number of speakers of varied native languages, one consistent finding is lingual movement during this closure interval is sensitive to the surrounding vowel context. In the dysarthria associated with Parkinsons disease (PD), a general reduction in articulatory motions is well known, but little has been reported on segment-specific movement. The overall objective of this presentation is to gain segment-specific information on speech movements in PD during stop closure by examining the characteristics of the labial/lingual movement during stop closure of PD compared to neurologically-healthy speakers.

Methods *Participants* Articulatory kinematic data from individuals with PD and healthy controls will be reported. Participants were asked to repeat three sets of phrases three times in a conversational voice: (1) *bilabial*: lee pool, lease pool, lee spool, lease spool, (2) *alveolar*: lee tool, lease tool, lee stool, lease stool, and (3) *velar*: lee cool, lease cool, lee school, lease school. These were adapted from Bell-Berti and Harris (1979) to vary place of articulation, the durations of the consonant strings, the duration of consonants in each string, and the location of syllable boundaries within the strings. The objective was to vary the duration of stop closure by contexts rather than volitional rate adjustments.
Analysis Labial/lingual function during 1) pre-stop, 2) post-stop vocalic nuclei and 3) stop closure was recorded by a 3D EMA system. 120 stop closures (10 participants * 3 stops * 4 contexts) are being analyzed regarding the time-varying position, speed and acceleration of four measurement points: tongue front, tongue back, lower lip, and upper lip.

Expectation Based on the prior findings that in VCV sequences lingual motions for the second vowel start well before the stop closure is completed (Öhman, 1966), we expect similar observations for healthy speakers but different patterns for talkers with PD. Discussions will be devoted to the two following question: (1) Are segmentally based movements of the tongue during stop closures reduced in PD, and (2) if so, how does the reduction affect coarticulatory behavior?

References

Bell-Berti, F., and Harris, K. S. (1979). Anticipatory coarticulation: some implications from a study of lip rounding. *Journal of Acoustical Society of America, 65*, 1268-1270.

Löfqvist, A. & Gracco, V. L. (1997). Lip and jaw kinematics in bilabial stop consonant production. *Journal of Speech, Language, and Hearing Research, 40*, 877-893.

Öhman, S., E., G. (1966). Coarticulation in VCV utterances: spectrographic measurements. *Journal of Acoustical Society of America, 39*, 151-168.

TREATMENT OF DYSPROSODIC SYMPTOMS IN ACQUIRED APRAXIA OF SPEECH - A SINGLE CASE STUDY

Ragna Krug, Hanna Stübner, Sophie Hoffmann, Judith Heide, Maria Blickensdorff
University of Potsdam, Potsdam, Germany

Introduction Dysprosody is a key symptom in acquired Apraxia of Speech (AOS; McNeil, Doyle, & Wambaugh, 2000). We report an intervention which was based on the Principles of Motor Learning (Maas, Robin, Austermann Hula, Freedman, Wulf, Ballard & Schmidt, 2008) and that was designed to reduce dysprosodic symptoms, namely intersyllabic pauses in consonant clusters and additions of schwa, in a patient demonstrating chronic mild AOS.

A range of different tasks was applied during the reported intervention period, including stress and focus realization in declarative and interrogative sentences, as well as word production in variable contexts. We report on implementation and outcome with respect to the word production task.

Methods & Materials A multiple baseline control-task design was used to investigate pre-treatment stability of performance, treatment effects and generalization to untreated items (Stadie & Schröder, 2009). Six therapy sessions of 45 minutes were conducted.

Item sets for treatment and generalization each consisted of 28 two-syllabic, both words and pseudowords with a word- medial CC.CCC- or CCC.CC-cluster.

Each therapy session started with a pre-practice phase in which two tasks were applied: a receptive phoneme detection tasks and a word-repetition task. Integral stimulation techniques were offered to facilitate correct responses. In the practice phase, an affective intonation task was used in which the items had to be uttered with different emotions. During the practice phase, delayed feedback on knowledge of results was provided.

Results Improved word production, measured as a significant reduction in intersyllabic pauses was observed for treated items ($p=<.01$, McNemar Test). Although no generalization to untreated items was observed, the analysis of spontaneous speech samples revealed a reduction of intersyllabic and wordfinal additions of schwa and pauses, leading to an improvement in fluency of speech. Control task performance was unaltered ($p>.05$ McNemar Test).

Conclusion Principles of Motor Learning were successfully applied to treat dysprosodic symptoms on word-level in a patient with AOS.

References

McNeil, M.R., Doyle P.J., & Wambaugh J. (2000). Apraxia of Speech: A Treatable Disorder of Motor Planning and Programming. In S. E. Nadeau, L. J. Gonzales Rothi & B. Crosson (Eds.), *Aphasia and language: theory to practice* (pp. 221-266). New York: Guilford Press.

Maas, E., Robin, D. A., Austermann Hula, S. N., Freedman, S. E., Wulf, G., Ballard, K. J., & Schmidt, R. A. (2008). Tutorial: Principles of motor learning in treatment of motor speech disorders. *American Journal of Speech - Language Pathology, 17*, 277 -298. doi:10.1044/1058- 0360(2008/025)

Stadie, N., & Schröder, A. (2009). *Kognitiv orientierte Sprachtherapie: Methoden, Material und Evaluation für Aphasie, Dyslexie und Dysgraphie.* München: Elsevier, Urban & Fischer.

AUTOMATIC CONTRASTIVE STRESS MEASUREMENT IN DYSARTHRIA

Lukas Latacz[1,2], Werner Verhelst[1,2], Hichem Sahli[1,2]
[1] *Vrije Universiteit Brussel, Brussels, Belgium*
[2] *Imec, Leuven, Belgium*

Objective A new algorithm for automated determination of contrastive stress (i.e. focus) in dysarthric speech was evaluated. We investigated the reliability of the algorithm by comparing the automated measurements to the judgements of professional listeners.

Background Healthy speakers employ changes in fundamental frequency (F0), intensity and/or duration to indicate whether a word in a sentence is stressed, e.g. "The TEACHER arrived late" versus "The teacher arrived LATE". Acoustic differences in the realizations of contrastive stress between healthy speakers and speakers with dysarthria have previously been reported, see e.g. (Patel and Campellone, 2009). Listeners might therefore experience greater trouble in recognizing the intended meanings produced by dysarthric speakers compared to healthy speakers.

Method We used annotated recordings of 18 Dutch-speaking speakers with dysarthria secondary to Parkinsons disease (Martens et al., 2011). Each speaker was presented with three pairs of sentences in which one keyword needed to be emphasized. These sentence pairs were used in a reading task and an imitation task. Three professional listeners judged these samples. A single judgement for each sample was obtained by majority voting.

Our algorithm uses linear regression models to map several acoustic features based on F0, energy and duration to prominence ratings - ratings on a continuous scale to determine how prominent a word or syllable is. Next, a logistic regression model determines which words are emphasized based on the predicted prominence ratings. The latter model was trained and evaluated on recordings of 36 healthy Dutch-speaking speakers who participated in the same tasks as the dysarthric speakers and which were annotated in the same manner. None of the models were trained using pathological speech.

Results Our algorithm obtained an overall accuracy of 85% on dysarthric speech. Cohens kappa was 0.70. The performance on recordings of healthy speakers was evaluated using 10-fold cross-validation, leading to 90% overall accuracy and Cohens kappa of 0.80.

Conclusions A substantial agreement between the listeners judgements and the automated predictions was obtained. The performance on recordings of healthy speakers was moderately better, which might be explained by differences in the realizations of contrastive stress such as acoustic cue exchange or the presence of pitch irregularities such as jitter and shimmer. Further research is needed to evaluate the performance of our algorithm on a broader range of sentences.

The algorithm is currently being integrated in a clinical software tool for assessing and managing prosody, articulation and speech intelligibility in dysarthric speech. Further clinical evaluations using this tool are planned.

References

Martens et al. (2011). Assessment of Prosodic Communicative Efficiency in Parkinson's Disease as Judged by Professional Listeners. *Parkinsons Disease, 2011.*

Patel, R., and Campellone, P. (2009). Acoustic and Perceptual Cues to Contrastive Stress in Dysarthria. *J Speech Lang Hear Res, 52*(1), 206-222.

DO PATIENTS WITH SPEECH IMPAIRMENTS BENEFIT FROM AUDITORY PRIMING WITH A REGULAR METRICAL PATTERN?

Katharina Lehner[1], Ingrid Aichert[1], Simone Falk[1,2], Mona Späth[1], Wolfram Ziegler[1]
[1] *Ludwig-Maximilians-Universität, Munich, Germany*
[2] *Université Sorbonne Nouvelle Paris-3, Paris, France*

Background In earlier investigations based on repetition tasks we found that the regular metrical pattern in German (trochee), unlike the iambic pattern, facilitates articulation in patients with apraxia of speech (AOS), confirming that segmental and prosodic aspects of speech production interact (Aichert et al., 2016). This result provided further evidence for a nonlinear architecture of speech motor representations, in which the articulatory primitives of speech motor planning, i.e., phonetic gestures for the articulation of vowels and consonants, are dominated by the rhythmical patterning of words (cf. Ziegler & Aichert, 2015). In this study we investigated if articulation in apraxic speakers also benefits from auditory priming by regular metrical patterns. Furthermore, we asked if the advantage of regular speech rhythm, if present, is confined to impairments at the motor planning stage of speech production (i.e., AOS), or if it also applies to phonological encoding impairments (PI).

Method 10 patients with Apraxia of speech, 10 aphasic patients with postlexical phonological impairment, and 36 neurologically healthy speakers were examined. A sequential synchronization paradigm based on a sentence completion task was conducted in four different conditions: (1) Regular prime sentence, trochaic target word, (2) Regular prime sentence, iambic target word, (3) Irregular prime sentence, trochaic target word, (4) Irregular prime sentence, iambic target word. A total of 96 sentences were presented in a randomized order. Speech errors were classified by error types (e.g. segmental errors, prosodic errors, groping). Furthermore, responses were analyzed for onset latency and duration of the target word.

Results The study revealed an influence of the metrical regularity of speech input in patients with AOS as well as in patients with PI. They therefore seemed to exploit intrinsic rhythmic cues in speech for the initiation and the segmental realisation of the target words. Our study also comfirmed one more time the faciliating effect of regular (trochaic) word stress on speech accuracy in patients with AOS. The same effect could be also observed in individuals with PI. There seems to be a robust metrical influence on speech at both the phonological and the phonetic planning stages of speech production.

Discussion Results will be interpreted on the background of models of speech production and apraxia of speech. In particular, they will be discussed with reference to rhythmic synchronization accounts of betweenspeaker interaction (e.g., Cummins, 2009).

References

Aichert, I., Späth, M., & Ziegler, W. (2016). The Role of Metrical Information in Apraxia of Speech. Perceptual and Acoustic Analyses of Word Stress. *Neuropsychologia, 82*, 171-178.

Cummins, F. (2009). Rhythm as entrainment: The case of synchronous speech. *Journal of Phonetics, 37*, 16.28.

Ziegler, W., & Aichert, I. (2015). How much is a word? Predicting ease of articulation planning from apraxic speech error patterns. *Cortex, 69*, 24-39.

THE COMMON ENDOPHENOTYPES OF SPEECH AND READING DISORDERS: WHAT DO THEY TELL US ABOUT THE PHONOLOGY - PHONETICS DIVIDE?

Ben Maassen
University of Groningen, Groningen, The Netherlands

Clinically, considerable overlap exists between speech sound disorder (SSD), among those Childhood Apraxia of Speech (CAS) and reading disorder (RD). Studies have revealed common early precursors in oral motor skills, auditory perception, phonological awareness, phonological memory, vocabulary, and speech automation processes (Lewis et al. 2011; Maassen et al. 2012). Also, there is quite some overlap between candidate genes associated with SSD and RD (Pennington & Bishop, 2009), suggesting common underlying deficits.

This paper gives an overview of the literature on endophenotypes at the neurological and cognitive level underlying each of these disorders, thereby analyzing associations and dissociations. Some remarkable findings are the following. Longitudinal studies have shown that poor auditory-phonetic processing, such as tested in speech-in-noise or categorical perception tasks, is a strong, general predictor of any speech or language disorder. Likewise, nonword repetition and rapid serial naming are sensitive tasks to diagnose specific language impairment (SLI), childhood apraxia of speech (CAS) and other SSDs, as well as RD. One of the explanations for not finding (double) dissociations in developmental disorders, as compared to acquired disorders in adulthood, is the strong dependency of functions during speech and language acquisition; in particular, phonology develops on the basis of auditory and articulatory phonetics on the one hand, and the top-down pressure of a growing vocabulary on the other.

In the Dutch Dyslexia Program (DDP) 200 children with familial risk (FR) of dyslexia and over 100 non-FR children were followed from age 2 months to 12 years. Significant early predictors of RD were deficits in auditory perception, phonological awareness, vocabulary, nonword repetition and rapid serial naming. Very few cases of SSD occurred, and no cases of CAS. Comparison of these results with studies of the endophenotypes that play a role in the co-morbidity RD in children diagnosed with SSD, gives insight in the possible associations and dissociations across the phonology - phonetics divide.

References

Lewis, B. A., et al. (2011). Literacy outcomes of children with early childhood speech sound disorders: Impact of endophenotypes. *JSLHR, 54*(6), 1628-1643.

Maassen, B., et al. (2012). Neurolinguistic and neurophysiological precursors of dyslexia: Selective studies from the dutch dyslexia programme. In A. A. Benasich, & R. H. Fitch (Eds.), *Developmental dyslexia: Early precursors, neurobehavioral markers and biological substrates.* (pp. 119-132). Baltimore, Maryland: Paul Brookes Publishing Comp.

Pennington, B. F., & Bishop, D. V. M. (2009). Relations among speech, language, and reading disorders. *Annual Review of Psychology, 60*, 283-306.

WHAT HAPPENS WHEN THEY GROW UP: A SURVEY OF ADULTS WHO HAD CHILDHOOD APRAXIA OF SPEECH

Patricia McCabe[1,3], Jonathan Preston[2], Elizabeth Murray[1], Geraldine Bricker[1], Angela Morgan[3]
[1]*University of Sydney, Sydney, Australia*
[2]*Syracuse University, Syracuse, United States of America*
[3]*Murdoch Childrens Research Institute, Melbourne, Australia*

Childhood Apraxia of Speech (CAS) is an uncommon speech disorder attributed to poor motor speech planning and programming. It has been demonstrated that children with CAS may have difficulties with spoken and written language (Lewis, et al 2004) which may persist into adulthood (Carrigg, et al 2015).

CAS has been now been identified as both heritable and significant in childhood outcomes (Fedorenko et al 2016). Although limited case examples exist, what is largely unknown is what happens to people with CAS as they reach and traverse adulthood. This means we lack prognostic information including educational and socioeconomic outcomes. Additionally, the well described combination of limited service availability and poor therapy outcomes means that those who are now adults and who were diagnosed with CAS in childhood can provide baseline measures which will allow us to evaluate the long term current and future improvements in services on people with CAS.

This research aimed to provide an initial description of the impact of lifelong motor speech disorder through a survey of adults who were diagnosed before the age of 16 as having CAS. An international online survey was conducted between June and December 2016. Fifty three participants commenced the survey and 28 valid responses were recorded. 71% of responses were from the United States; 14% were from Australia with 4 other countries making the remainder. Thirtyeight percent of respondents had completed only high school and 28% had an undergraduate degree with 37% still studying.

Regarding familial trends, 74% had no other family member with CAS; although 32% had a family member with a communication disability more broadly while 14% of respondents had participated in genomic testing. 56% of respondents say they still have speech problems and 48% had difficulty with reading while 20% believe they have recovered. The majority also reported avoiding words and speaking tasks. The voices of the respondents will be used to flesh out these numerical results. This paper will provide clinicians with prognostic information not currently available and which parents need in order to understand their childs potential and to obtain services.

References

Carrigg, B., Baker, E., Parry, L., & Ballard, K. J. (2015). Persistent speech sound disorder in a 22yearold male: Communication, educational, socioemotional, and vocational outcomes. *SIG 16 Perspectives on SchoolBased Issues, 16*(2), 3749.

Fedorenko, E., Morgan, A., Murray, E., Cardinaux, A., Mei, C., TagerFlusberg, H., Fisher, S.E. & Kanwisher, N. (2016). A highly penetrant form of childhood apraxia of speech due to deletion of 16p11.2. *European Journal of Human Genetics, 24*(2), 302306.

Lewis, B. A., Freebairn, L. A., Hansen, A. J., Iyengar, S. K., & Taylor, H. G. (2004). Schoolage followup of children with childhood apraxia of speech. *Language, Speech, and Hearing Services in Schools, 35*(2), 122140

ARTICULATORY MOVEMENT VARIABILITY IN TALKERS WITH ALS: COMPARISONS OF CLEAR AND SLOW SPEECH EFFECTS

Antje Mefferd[1], Mili Kuruvilla-Dugdale[2]
[1] *Vanderbilt University Medical Center, Nashville, United States of America*
[2] *University of Missouri, Columbia, United States of America*

Background and Aim During the early stages of speech decline, individuals with ALS are often instructed to use a clear speech strategy or to reduce their speaking rate to improve speech intelligibility. Although it is well-known that both tasks can impact a talkers ability to produce consistent speech movement patterns, direct comparisons of slow and clear speech effects on movement variability in ALS are currently lacking. Therefore, the current study sought to identify articulatory kinematic changes in response to these two global speech interventions.

Method The complete study will include 12 talkers with ALS and 12 controls. Data from seven individuals with ALS (mild to moderate dysarthria) and seven healthy controls have been collected so far. Participants were asked to produce the sentence *I owe you a yoyo* ten times each using habitual, clear, and slow speech. Articulatory movements were recorded with a 3D electromagnetic articulograph (WAVE, NDI Inc.). 5DOF (degrees of freedom) sensors were placed on the sagittal midline of the tongue (tip and dorsum), lower lip, and jaw (right and left). To correct the speech kinematic recordings for head movement, a 6DOF sensor was placed on the forehead as a reference sensor. For each participant the spatiotemporal variability index (STI) was calculated for the tongue tip, tongue dorsum, lower lip, and jaw based on ten vertical displacement signals of each sensor in each speech condition (Smith et al., 1995). Further, utterance movement durations, vocal intensity, and articulatory working space of the tongue, lip, and jaw were used to evaluate rate and clarity manipulations.

Results & Discussion Data analysis is currently under way. Based on previous findings, it was hypothesized that in both groups of talkers speech movements will be significantly more variable during slow speech compared to typical speech across all articulators. Predictions for clear speech effects on speech movement variability in ALS were difficult to make because so far clear speech effects on articulatory kinematics have only been studied in talkers with Parkinsons disease (Dromey, 2000). With regards to between-group differences, we expected to replicate previous findings of significantly lower jaw and lip movement variability in talkers with ALS compared to controls during habitual speech (Mefferd et al., 2014). Because the significantly reduced lip and jaw movement variability in talkers with ALS may be suggestive of an impaired speech motor system, similar between-group differences were expected for tongue movement variability during habitual speech.

References

Dromey, C. (2000). Articulatory kinematics in patients with Parkinson disease using different speech treatment approaches. *Journal of Medical Speech-Language Pathology, 8*(3), 155-161.

Mefferd, A. S., Pattee, G. L., & Green, J. R. (2014). Speaking rate effects on articulatory pattern consistency in talkers with mild ALS. *Clinical Linguistics & Phonetics, 28*(11), 799-811.

Smith, A., Goffman, L., Zelaznik, H. N., Ying, G., & McGillem, C. (1995). Spatiotemporal stability and patterning of speech movement sequences. *Experimental Brain Research, 104*(3), 493-501.

ARTICULATING VOWELS IN PARKINSON'S SPEECH: A CROSS-LINGUISTIC STUDY IN ITALIAN AND ENGLISH

Caroline Menezes[1], Barbara Gili Fivela[2], Massimiliano Iraci[2], Marisa Lucarelli[1], Alexis Garon[2]
[1]*University of Toledo, Ohio, United States of America*
[2]*Universitá del Salento, Lecce, Italy*

This crosslinguistic study compares articulatory variations in vowel production between Parkinsons disease (PD) and healthy speakers (HC) of Italian and American English. Dysfunction of the basal ganglia leads to slowing down and an increase in variability of articulatory movements (bradykinesia) (Ackermann, et al., 1997; Darley, et al., 1975). This study aims at understanding 1) how bradykinesia affects vowel production in Parkinsons speech and if differences in vowel density across the two languages would influence this effect (Italian showing fewer phonemes than English) and 2) how sentence length affects vowel production across health conditions and if this is universal to Italian and English. We hypothesize greater changes may be observed in low-density systems and longer sentences, because of the lower impact of changes on comprehension and the result of fatigue, respectively. Therefore, we investigate the articulation of closed (/i/, /u/) and open (/a/or//ɑ/)/ vowels in stressed/unstressed CV syllables, where C=/m/. Syllables are produced in controlled disyllabic target words (i.e., [ma.mi], [mi.ma], [mu.ma], [ma.mu]) at the initial and final position of short and long sentences (7 and 16 syllable each, with constant inter-stress and inter-accent distance between target syllables). Both English and Italian utterances are consistent for the above criteria.

For each language, one mild PD-speaker and one peer HC participated in the study. All subjects read the corpus 7 times at normal speech rate. Both in Italy and in the USA, articulatory and acoustic data were collected using the AG500 (Carstens-Med. GmbH). Reference coils were placed on the left and right mastoids and on the upper incisor. Three coils were also used to determine the bite plane. Data were recorded from 5 coils distributed between the tongue dorsum, tongue tip, upper and lower lips, lower incisors.

Acoustic measurements were made on Praat (Boersma & Weenink) and included vowel durations and first and second formants calculated as an average over the entire vowel. Kinematic measurements were performed after segmentation and labelling of the signal (gesture attainment, labelled at the zero velocity, and maximum velocity, labelled at the velocity peak of the relevant coil). The duration, amplitude and velocity of tongue dorsum gestures were calculated.

Preliminary results on Italian reveal interesting differences between the two subjects. For the PD-speaker only, utterance length significantly influenced the acoustic duration of the unstressed vowel ($\chi^2(1) = 25.94$, p <0.001) as well as the second formant of both stressed ($\chi^2(1) = 5.91$, p < 0.015) and unstressed vowels ($\chi^2(1) = 23.46$, p <0.001). Cross-linguist analysis of vowel quality is still to be conducted.

References

Ackermann, H., Hertrich, I., Daum, I., Scharf, G., & Spieker, S. (1997). Kinematic analysis of articulatory movements in central motor disorders. *Movement Disorders, 6*, 1019 1027.

Darley, F.L., Aronson, A .E., Brown, J. R. (1975). *Motor speech disorders.* Philadelphia: W.B. Saunders

THE EFFECT OF LEXICAL STRESS ON EYE MOVEMENT AND PROSODY DURING READING IN PARKINSON'S DISEASE

Eunsun Park, Frank R. Boutsen, Justin D. Dvorak
University of Oklahoma Health Sciences Center, Oklahoma City, United States of America

Background A majority of individuals with Parkinsons disease (PD) manifest hypokinetic dysarthria, which is characterized by dysprosody. Some also experience reading difficulties and impaired eye movements (Anderson & MacAskill, 2013). However, currently there exists no objective approach to evaluate both prosody and eye-movement behavior during silent and oral reading in individuals with PD. The goal of this study is to examine the effects of lexical stress on eye movements and prosody planning during reading in individuals with mild to moderate PD, as compared to age-matched healthy individuals.

Method 13 participants with mild to moderate PD (mean age 54.8±7.8 yrs; mean duration of PD 4.04±3.98 yrs) and 16 healthy participants (mean age 55.3±6.5 yrs) read sentences under silent and oral conditions. Each sentence was either short (10 syllables) or long (16 syllables) and contained one of two possible target words with one or two stressed syllables. Eye-movement variables (e.g., dwell duration, firstpass fixation duration, average fixation duration, first fixation landing position (letters counted from word onset), and eye-voice span (EVS)) and speech variables (e.g., normalized intensity range, normalized pitch range, and speech rate) were measured. All statistical tests were conducted in SAS 9.4.

Results During silent reading, participants with PD demonstrated significantly longer first-fixation duration (p=0.02) and average fixation duration (p=0.01). Participants with PD descriptively demonstrated longer dwell duration, first fixation duration, second fixation duration, and average fixation duration in both silent and oral reading condition as compared to the controls. The number of lexical stresses was not associated with eye-movement variables such as dwell duration (p=0.20), first fixation duration (p=0.32), and average fixation duration (p=0.97). The average first-fixation landing positions in the 10-letter target words during oral reading were statistically different between groups (p=0.01; PD group 2.8±1.6 letters; Control group 3.3±2.1 letters). Participants with PD had a significantly slower speech rate than controls (p=0.04).

Discussion This study challenges results from earlier studies showing that prosodic planning reveals itself in durational eye-movement variables (Ashby & Clifton, 2005). That said, the results suggest that participants with PD require additional cognitive processing time to encode and decode target words during silent and oral reading.

References

Anderson, T.J. & MacAskill M.R. (2013). Eye movements in patients with neurodegenerative disorders. *Nat. Rev. Neurol.* doi:10.1038/nrneurol.2012.273.

Ashby, J. and Clifton, C., Jr. (2005). The prosodic property of lexical stress affects eye movements during silent reading. *Cognition, 96*(3), B89-B100.

SENSORIMOTOR INTEGRATION IN ADULTS WHO STUTTER: EVIDENCE FROM BEHAVIOR AND NEUROIMAGING

Anastasia G. Sares[1], Mickael Deroche[1], Hiroki Ohashi[2], Douglas M. Shiller[3], Vincent Gracco[1,2]
[1] *McGill University, Montréal, Canada*
[2] *Haskins Laboratories, New Haven, United States of America*
[3] *Université de Montréal, Montréal, Canada*

The integration of sensory perception and motor action is a hallmark of both music and speech production. For individuals who stutter, deficient sensorimotor integration has been suggested as a primary trait of the speech disorder. What is unclear is whether the apparent deficit in sensorimotor integration for speech extends to other vocal behavior. Here we compare the performance and brain activity of adults who stutter (AS) and adult controls (AC) under a sensorimotor manipulation: altered auditory feedback during vocalization.

Participants sustained the vowel /a/ while hearing on-line feedback of their own voice through headphones. In some trials, feedback was briefly shifted up or down by 100 cents to simulate a vocal production error. Participants vocal responses were recorded and analyzed to see how they reacted to these shifts in pitch. As previously shown, participants tended to compensate for the auditory pitch change by moving in the opposite direction of the shift. The average compensatory response was smaller for AS than for AC: the current analysis reveals that this may be due to both actual magnitude differences and more variability in timing in the group with a stutter.

Multiple differences in brain function have been previously identified during the fluent speech of individuals who stutter. In a second session, AS and AC participants carried out the same experiment in an MRI scanner to assess differences in functional activation between groups. Preliminary analyses indicate that, in response to shifted feedback versus no shift, participants with a stutter had greater activity in the cerebellum, and reduced activation in other areas. Cerebellar over-activation has been observed in previous tasks for adults who stutter, and may reveal more conscious or effortful vocal control in this group.

Further reading

Loucks, T., Chon, H., & Han, W. (2012). Audiovocal integration in adults who stutter. *Int J Lang Commun Disord, 47*(4), 451-456. http://doi.org/10.1111/j.1460-6984.2011.00111.x.

Parkinson, A. L., Flagmeier, S. G., Manes, J. L., Larson, C. R., Rogers, B., & Robin, D. A. (2012). Understanding the neural mechanisms involved in sensory control of voice production. *NeuroImage, 61*(1), 314-22. http://doi.org/10.1016/j.neuroimage.2012.02.068.

De Nil, L. F., Kroll, R. M., & Houle, S. (2001). Functional neuroimaging of cerebellar activation during single word reading and verb generation in stuttering and nonstuttering adults. *Neuroscience Letters, 302*(2-3), 77-80. http://doi.org/10.1016/S0304-3940(01)01671-8.

DEBATING THE USE OF NONSPEECH TASKS IN DYSARTHRIA ASSESSMENT: THE ASPECT OF FEASIBILITY

Theresa Schölderle, Wolfram Ziegler, Anja Staiger
Ludwig-Maximilians-University, Munich

Background & Aim There is a continuing debate on the clinical value of nonspeech oral motor tasks for the assessment of dysarthria (Maas, 2016). While previous investigations scrutinized if such tasks allow for valid conclusions on speech functions (Staiger, Schölderle, Brendel, Bötzel, & Ziegler, 2016), the aspect of feasibility has gathered little attention. Traditionally, nonspeech tasks are assumed to be more feasible than standard speech tasks, especially in cognitively impaired patients, as they are believed to impose lower cognitive-linguistic demands on the examined person (compare, for instance, single articulator movements like "purse your lips" vs. reading a text) (Folkins et al., 1995). However, this assumption has never been investigated empirically.

Methods Twenty-one adults with dysarthria and cognitive impairment due to cerebral palsy participated. They were assessed with six standard speech tasks (e.g., text reading, sentence repetition) and three standard nonspeech tasks (e.g., rapid syllable repetition, repetitive articulator movements). The tasks were evaluated for their feasibility on the basis of common clinical criteria.

Results Comparisons between all pairs of speech and nonspeech tasks showed no significant difference in most instances. Interestingly, however, in a substantial number of cases speech tasks were significantly more feasible than nonspeech tasks. Among the speech tasks, only the reading of a standard text turned out to be challenging, whereas a modified supported reading task was less demanding. Several nonspeech tasks induced dysexecutive symptoms (e.g., perseverations, disinhibition) and increased motor effort.

Discussion The present study enhances reservations against the diagnostic use of nonspeech tasks, since they lack the alleged advantage of being more feasible than speech tasks. Among standard speech tasks, only text reading can be challenging in individuals with cognitive impairments, but reading demands can be diminished through specific adaptations.

References

Folkins, J. W., Moon, J. B., Luschei, E. S., Robin, D. A., Tye-Murray, N., & Moll, K. L. (1995). What can nonspeech tasks tell us about speech motor disabilities? *Journal of Phonetics, 23*(1), 139-147.

Maas, E. (2016). Speech and nonspeech: What are we talking about? *International Journal of Speech-Language Pathology*, 1-15.

Staiger, A., Schölderle, T., Brendel, B., Bötzel, K., & Ziegler, W. (2016). Oral Motor Abilities Are Task-Dependent: A Factor Analytic Approach to Performance Rate. *Journal of Motor Behavior, 9*, 1-12.

DEVIANT COARTICULATION IN CHILDREN WITH CHILDHOOD APRAXIA OF SPEECH (CAS) DOES NOT INCLUDE HYPERARTICULATION

Hayo Terband
Utrecht University, Utrecht, The Netherlands

Childhood Apraxia of Speech (CAS) is a developmental motor speech disorder with a core deficit in planning and/or programming the spatiotemporal parameters of movement sequences. One of the hallmark characteristics of the speech of children with CAS are "lengthened and disrupted coarticulatory transitions between sounds and syllables" (ASHA, 2007, p. 4). Whereas coarticulation effects usually change the characteristics of a speech sound in the direction of the neighbouring speech sound, coarticulation has been found to be both stronger and more extended as well as the opposite, more segmental (or hyperarticulation) in children with CAS as compared to their typically developing peers.

The present study investigates the hypothesis that this paradox is due to reduced phonological distinctiveness. The smaller the distinction between speech sounds, the weaker the possible coarticulatory influence on surrounding speech sounds. Children with CAS demonstrated less coarticulation in studies where they also showed a decreased differentiation of speech sounds as compared to their typically developing peers (e.g., Sussman, Marquardt, & Doyle, 2000). In studies that feature similar phonological distinctiveness in the speech of children with CAS in comparison with normally developing children, coarticulation was found to be stronger and more extended (e.g., Nijland et al., 2003).

The dataset has been collected previously as part of the studies of Nijland and colleagues. The speech samples comprise three repetitions of simple bisyllabic meaningless utterances of the type [dəCV] (C = /b,d/ and V = /i,u/) in the context of the sentence /he wɪːr/ of 16 children with CAS (ages 5,5 - 7,5 years) and 8 typically developing children (aged 5,0 - 7,3 years). Coarticulation is analysed in F2 of [ə] and calculated as the vowel context dependent ratio relative to the size of the produced vowel contrast, for each consonant context separately.

Preliminary results show coarticulation to be stronger in children with CAS as compared to the controls in the /d/ context, but not in the /b/ context. While the control group shows a differentiation in coarticulation between consonant contexts, the children with CAS do not. These results indicate that deviant coarticulation in CAS does not include hyperarticulation. The increased coarticulation in CAS compared to controls is limited to certain articulatory contexts. Clinical implications and directions for future research will be discussed.

References

ASHA. (2007). Childhood Apraxia of Speech [technical report]. *American Speech-Language-Hearing Association.* Available from www.asha.org/policy.

Nijland, L., Maassen, B., Vander Meulen, S., Gabreels, F., Kraaimaat, F. W., & Schreuder, R. (2003). Planning of syllables in children with developmental apraxia of speech. *Clinical Linguistics & Phonetics, 17* (1), 1-24.

Sussman, H., Marquardt, T., & Doyle, J. (2000). An acoustic analysis of phonemic integrity and contrastiveness in developmental apraxia of speech. *Journal of Medical Speech-Laguage Pathology, 8*, 301-313.

LISTENING TO DISORDERED SPEECH RESULTS IN EARLY MODULATIONS OF AUDITORY EVENT-RELATED POTENTIALS

Catherine Theys, Megan J. McAuliffe
University of Canterbury, Christchurch, New Zealand

In the last decade, research on motor speech disorders has increasingly taken into account the bidirectionality between speaker and listener. Listening to disordered speech (e.g., dysarthria) may result in substantial phonemic uncertainty. In turn, a larger set of potential word target candidates may be activatedcontributing to intelligibility deficits. To resolve this uncertainty, a combination of both bottom-up and top-down processes are thought to play a role (Liss, 2007). The goal of the present study was to investigate the contribution of these processes by analysing listeners neurophysiological processing when listening to dysarthric speech.

Thirty healthy native English speakers (12 males, 18-44 years) participated in a speech perception experiment while undergoing 32-channel EEG recording. Similarly to Obleser and Kotz (2011), we focused on the auditory N100 as a marker for earlier sensory processing and the N400- like peak representing information on later cognitive-linguistic processes. Participants listened to 55 moderate hypokinetic dysarthric sentences and 55 control sentences. The experiment was repeated one week later to investigate the effects of repeated exposure to disordered speech. The amplitudes and latencies of the event-related potentials over Cz were analysed.

Repeated measures GLM statistics of the N100 with speech type (dysarthria vs. control) and test session as independent variables showed a main effect of speech type, with increased amplitude ($F_{ampl}(28)=12.18$, $p<.01$) and decreased latency ($F_{lat}(28)=6.77$, $p=.02$) when listening to dysarthric versus control speech. There was no significant main effect of test session or interaction effect. In contrast, no significant effects of speech type and test session were observed on the amplitude of the N400-like peak. For latency, only a significant interaction effect was present ($F_{lat}(28)=4.16$, $p=.05$), evidenced by decreased latency for dysarthric sentences during the first test session, and the reverse during the second session.

The N100 results show that the quality of the auditory signal in naturally degraded dysarthric speech influences early sensory auditory processing, indicating an increase in the initial allocation of neurophysiological resources (Obleser & Kotz, 2011). The N400 latency results show that later, more cognitive-linguistic processes are not only influenced by the degradation of the signal itself but also by the amount of exposure to that signal, a finding consistent with previous behavioural research on dysarthric speech (Borrie et al., 2012).

References

Liss, J.M. (2007). The role of speech perception in motor speech disorders. In G. Weismer (Ed.), *Motor speech disorders: Essays for Ray Kent* (pp. 195-231). San Diego, CA: Plural.

Obleser, J. & Kotz, S.A. (2011). Multiple brain signatures of integration in the comprehension of degraded speech. *NeuroImage, 55*(2), 713-723.

Borrie, S.A., McAuliffe, M.J., Liss, J.M., Kirk, C., O'Beirne, G.A., and Anderson, T. (2012). Familiarization conditions and the mechanisms that underlie improved recognition of dysarthric speech. *Language and Cognitive Processes, 27*(7-8), 1039-1055.

REACHING GOALS WITH LIMITED MEANS: AN ULTRASOUND STUDY OF SPEECH PRODUCED BY CHILDREN WITH MYOTONIC DYSTROPHY

Pamela Trudeau-Fisette, Marie Bellavance-Courtemanche, Christine Turgeon, Lucie Ménard
Université du Québec à Montréal, Montréal, Canada

Introduction Speech production entails appropriately timed contractions of facial and other muscles. Type 1 myotonic dystrophy (DM1) or Steinert's disease, a neurodegenerative disease that causes muscle weakness and difficulties in muscle relaxation after contraction, frequently affects orofacial articulatory dynamics (Harper, 2001). The effects are particularly severe in children diagnosed with DM1 (Guirnaraes, Suazo, and Nagahashi, 2010; Sjögreen, Lohmander, and Kiliaridis, 2011). Even though they have a limited range of lip, jaw, and tongue movements compared to typically-developing children, those children can produce some vowels and consonants with good intelligibility scores, especially when they are required to speak clearly. How do they compensate for their limited motor control abilities? To investigate the different strategies used to reach phonemic goals, we conducted a study of vowels and consonants produced by children with DM1.

Method We recruited fourteen 6- to 14-year-old French-speaking children diagnosed with DM1 and 14 aged-matched typically developing children. They were asked to produce repetitions of CV syllables where C was a consonant (/b d m n/) and V was a vowel (/i a u/) in two conditions: clear speech and conversational speech. A synchronized ultrasound, motion tracking system (Optotrak), and audio recording system was used to track lip and jaw displacements as well as tongue shape and position. Duration and formant values were also extracted. The Euclidean distances between vowels, in the formant space, were measured. Differences in lip and tongue contrasts between vowels were also calculated and the effects of speaker group, speaking condition, and phoneme, was measured through mixed models.

Results & Discussion The vowel contrasts in the formant space were reduced in children with Steinert disease compared to the control children. Different patterns of articulatory contrasts were observed among the children, with some children using more tongue contrasts than lip contrasts. Children with DM1 increased articulatory lip contrasts to a greater extent than tongue contrasts in the clear speech condition compared to the conversational speech condition. These results suggest that children with DM1 produce highly efficient compensatory strategies to make up for their muscular deficits.

References

Harper, S. P. (2001). *Myotonic Dystrophy*, Third Edition (Vol. 37). London: W. B. Saunders.

Guirnaraes, A S., Suazo, G. L, & Nagahashi M, S. K. (2010). Fenomeno miotonico orofacial en pacientes con distrofia miotonica de Steiner. *Avances en odontoestomatologia, 26*(3), 139-142.

Sjögreen, L., Lohmander, A , & Kiliaridis, S. (2011). Exploring quantitative methods for evaluation of lip function. *J Oral Rehabil, 38*(6), 410-422. doi:10.1111/j.1365-2842.2010.02168.x 110

GENERALIZATION EFFECTS OF TREATMENT FOR ACQUIRED APRAXIA OF SPEECH

Julie Wambaugh[1,2], Christina Nessler[2], Sandra Wright[2], Shannon Mauszycki[1,2]
[1] *University of Utah, Salt Lake City, United States of America*
[2] *VA Salt Lake City Health Care System, Salt Lake City, United States of America*

Introduction Acquired apraxia of speech (AOS) is a neurogenic motor speech disorder that is characterized by disruptions in articulation, speech rate, and prosody. Numerous treatments have been shown to result in improved articulation in AOS (Ballard et al., 2015). An aim of treatment is to promote improvement that extends beyond treated items. Response generalization to untreated exemplars of treated items (i.e., trained = sun, soup, seat; generalization = sip, sail, sing) has been found to be relatively robust across many AOS treatment studies. Stimulus generalization, which is improved production of trained items in untrained contexts (which can take many forms), has been studied less frequently with AOS treatments. When reported, stimulus generalization effects of AOS treatment have been variable within and across participants. The purpose of this investigation was to examine the generalization effects of Sound Production Treatment (SPT; Wambaugh et al., 1998; see Bailey, Eatchel, & Wambaugh for meta-analysis) which is an AOS treatment that has been repeatedly shown to have positive effects on treated items and untrained exemplars.

Method Twenty English speakers with chronic AOS and aphasia participated. SPT was administered to all participants in the contexts of single-subject experimental designs. A larger purpose of the study was to examine the effects of blocked versus random schedules of practice. Each participant received two phases of treatment: SPT-Blocked and SPT-Random. Generalization measures were conducted pre- and post-treatment. Generalization was measured in three conditions: 1) *Generalization to Longer Utterances.* Treatment items were elicited in the context of a semantically plausible utterance of longer length, e.g., target = flower, generalization = red flower. 2) *Generalization to Sentence Completion.* Treated words were elicited via sentence completion, e.g., "Flower. She picked a pretty" 3) *Speech Intelligibility.* The Assessment of Intelligibility of Dysarthric Speech (Yorkston & Beukelman, 1981) was administered.

Results At 6 weeks post treatment, increases of 16% (blocked items) and 18% (random items) above highest baseline accuracy levels were found for production of trained items in untrained longer contexts. Increases of 28% (blocked) and 35% (random) were found for production of trained items in untrained sentence completion contexts. Speech intelligibility scores increased by an average of 3% post-treatment. Wide ranges in performance were observed across participants for all measures.

Summary/Conclusions Stimulus generalization effects were positive for this group of speakers and reflected the range of increased accuracy that has been reported previously for SPT. Generalization to sentence completion was stronger than generalization to longer utterances.

References

Bailey, D., Eatchel, K., & Wambaugh, J.L. (2015). Sound Production Treatment: Synthesis and quantification of outcomes. *American Journal of Speech-Language Pathology, 24*, S798-S814.

Ballard, K., Wambaugh, J., Duffy, J., Layfield, C., Maas, E., Mauszycki, S., & McNeil, M. (2015). Updated treatment guidelines for acquired apraxia of speech: A systematic review of intervention research between 2004 and 2012. *American Journal of Speech-Language Pathology, 24*, 316-337.

Wambaugh, J.L., Kalinyak-Fliszar, M.M., West, J.E., & Doyle, P.J. (1998). Effects of treatment for sound errors in apraxia of speech. *Journal of Speech, Language, and Hearing Research, 41*, 725-743.

PLANUM TEMPORALE ASYMMETRY IN PEOPLE WHO STUTTER

Kate E. Watkins[1], Emily L. Connally[1], Peter Howell[2], David Ward[3], Jennifer Chesters[1], Patricia M. Gough[4]

[1] *University of Oxford, Oxford, UK*

[2] *University College London, London, UK*

[3] *University of Reading, Reading, UK*

[4] *Maynooth University, Maynooth, Ireland*

Previous studies have reported that the planum temporale - a language-related structure that normally shows a leftward asymmetry - had reduced asymmetry in people who stutter (PWS; Foundas et al., 2001) and reversed asymmetry in those with severe stuttering (Foundas et al., 2004). These findings are consistent with the theory that altered language lateralization may be a cause or consequence of stuttering.

Here, we re-examined these findings in a large sample of PWS ranging in age from 12 to 54 years. We evaluated planum temporale asymmetry in structural MRI scans obtained from 67 PWS and 63 agematched controls using: 1) manual measurements of the surface area; 2) voxel-based morphometry to automatically calculate grey matter density. We examined the influences of gender, age and stuttering severity on planum temporale asymmetry.

Results showed that the size of the planum temporale and its asymmetry were not different in PWS compared with Controls using either the manual or the automated method. Both groups showed a significant leftwards asymmetry on average (about one-third of PWS and Controls showed rightward asymmetry). Importantly, and contrary to previous reports, the degree of asymmetry was not related to stuttering severity. In the manual measurements, women who stutter had a tendency towards rightwards asymmetry but men who stutter showed the same degree of leftwards asymmetry as male Controls. In the automated measurements, Controls showed a significant increase in leftwards asymmetry with age but this relationship was not observed in PWS.

We conclude that reduced planum temporale asymmetry is not a prominent feature of the brain in PWS and that the asymmetry is unrelated to stuttering severity.

References

Foundas, A.L., Bollich, A.M., Corey, D.M., Hurley, M., & Heilman, K.M. (2001). Anomalous anatomy of speech-language areas in adults with persistent developmental stuttering. *Neurology, 57*(2), 207-15.

Foundas, A.L., Bollich, A.M., Feldman, J., Corey, D.M., Hurley, M., Lemen, L.C., & Heilman, K.M. (2004). Aberrant auditory processing and atypical planum temporale in developmental stuttering. *Neurology, 63*(9), 1640-6.

NEUROANATOMICAL CORRELATES OF BULBAR AMYOTROPHIC LATERAL SCLEROSIS (ALS)

Yana Yunusova[1,2], Sanjana Shellikeri[1,2], Matthew Myers[1,2], Sandra E. Black[1,2,3], Lorne Zinman[1,2,3]
[1]*University of Toronto, Toronto, Canada*
[2]*Hurvitz Brain Sciences Research Program, Sunnybrook Research Institute, Toronto, Canada*
[3]*Sunnybrook Health Sciences Centre, Toronto, Canada*

Background ALS is a multisystem neurodegenerative disorder with motor and extramotor - cognitive-linguistic - involvement; 30% of patients present with bulbar-onset ALS, however, up to 85% of all patients exhibit bulbar disease as ALS progresses. Some behavioural studies propose that the degree of bulbar dysfunction may be associated with an increased burden of extra-motor impairments, suggesting a unique neurodegenerative profile(Massman et al., 1996). Yet, measures that are able to detect relevant neuroanatomical changes are unavailable. Furthermore, MRI studies examining associations between bulbar ALS and extramotor involvement are limited in their methodologies (Schuster et al., 2014). Identification of bulbar signs as early as possible is important in order to access timely clinical care and experimental clinical trials. The objectives of this study are to: (1) clinically validate a novel method of primary motor cortex (PMC) partitioning; and (2) examine associations between bulbar motor dysfunction and cognitive/language neuroanatomy.

Methods T1-weighted (130 slices, res = 1 mm^3) and DTI images (36 directions, res = 1.4 mm^3) were used to measure gray (GM) and white matter (WM) integrity in 16 patients with varying degrees of bulbar ALS and 13 healthy controls (HC). Measures included cortical thickness, gray matter volume, surface area, and DTI indices of bulbar motor, language, and cognitive regions in the cortex and underlying white matter tracts. Bulbar-specific motor areas were identified using a novel semi-automatic method (Freesurfer) and reliability and validity were established. Associations were examined between extramotor regions and bulbar motor, limb motor, and disease severity clinical measures.

Results The novel method for PMC partitioning had excellent interrater reliability and was validated with clinical measures of motor dysfunction. Group effects (ALS vs. HC) indicated increased WM pathology in left Broca (fractional anisotropy, p=.050) and Wernicke (axial diffusivity, p=.036) areas. Articulatory rate was significantly correlated with GM changes in bilateral Broca areas and Heschls gyri. Neither disease severity nor limb dysfunction correlated with extramotor neuroanatomy.

Discussion Patients with bulbar ALS show gray and white matter involvement in speech and language motor and extramotor cortical areas. Further, the degree of clinical bulbar impairment is related to the neuroanatomic changes in these regions, validating the potential use of MRI metrics as markers of bulbar disease.

References

Massman, P. J., Sims, J., Cooke, N., Haverkamp, L. J., Appel, V., & Appel, S. H. (1996). Prevalence and correlates of neuropsychological deficits in amyotrophic lateral sclerosis. Journal of Neurology, *Neurosurgey and Psychiatry, 61*(5), 450-455.

Schuster, C., Kasper, E., Dyrba, M., Machts, J., Bittner, D., Kaufmann, J., . . . Vielhaber, S. (2014). Cortical thinning and its relation to cognition in amyotrophic lateral sclerosis. *Neurobiology of Aging, 35*(1), 240-246.

CORTICAL OSCILLATIONS DURING SIMPLE AND COMPLEX VERBAL AND FINGER SEQUENCING TASKS

Luc De Nil[1], Soonji Kwon[1], Cecilia Jobst[2], Anna Mersov[1], Douglas Cheyne[1,2]
[1] *University of Toronto, Toronto, Canada*
[2] *The Hospital for Sick Children, Toronto, Canada*

The human somatosensory cortex demonstrates robust event-related desynchronization (ERD) of cortical beta band (15-30 Hz) oscillations prior to and during movement, followed by a reversed beta event-related synchronization (ERS) (Cheyne, 2013; Mersov, et al. 2016). To date, research directly comparing the modulation of these sensorimotor rhythms in speech and between speech and nonspeech motor tasks is sparse.

The aim of the current study was multifold: (1) to investigate whether motor preparation for speech can be reliably visualized using magnetoencephalography (MEG); (2) to compare cortical oscillatory modulations during speech and non-speech tasks; and (3) to investigate the effects of task complexity on these oscillatory modulations. Participants were sixteen healthy English-speaking adults (8 males and 8 females). For the nonverbal tasks, participants performed repeated button-presses of randomized complexity. The simple task consisted of a sequence of 4 identical button presses (1 finger), while the complex task consisted of a sequence of 4 different response buttons (3 different fingers). For the verbal tasks, participants produced a sequence of 4 identical (simple task) or 4 different syllables (complex task). Continuous brain activity was recorded throughout the tasks using MEG. Synthetic Aperture Magnetometry (SAM) was used to measure the location and time course of beta oscillations.

In the nonspeech task, motor onset generated a single, contralateral peak activation in the hand region of the sensorimotor cortex. Motor response time courses in this region showed significantly stronger pre-movement beta ERD in the simple task compared to the complex task. In contrast, the complex task generated a stronger beta ERS peak following the movement. Similar findings were found for the speech task with simple tasks showing stronger contralateral beta ERD as well as stronger ERS following movement in comparison to simple task. Speech, however, revealed more pronounced bilateral activation that was more localized towards the lateral sensorimotor cortex. While further detailed analysis is currently ongoing, our results confirm that MEG is an effective tool to visualize modulations of somatotopically organized cortical oscillatory activity during overt speech tasks; that there are similarities between speech and nonspeech motor tasks; and that cortical responses are influenced by task complexity. The relevance of these findings for our understanding of neural processes involved in motor tasks will be discussed.

References

Mersov, A.-M., Jobst, C., Cheyne, D.O., De Nil, L. (2016). Sensorimotor oscillations prior to speech onset reflect altered motor networks in adults who stutter. *Frontiers in Human Neuroscience, 10*, art. no. 443.

Cheyne, D.O. (2013). MEG studies of sensorimotor rhythms: A review. *Experimental Neurology, 245*, 27-39.

CORTICAL-CEREBELLAR CIRCUITS DRIVE SPEECH MOTOR ADAPTATION

Daniel R. Lametti, Phoebe Freidin, Harriet Smith, Kate E. Watkins
University of Oxford, Oxford, United Kingdom

Motor learning is studied in the lab by having participants interact with robotic arms or joysticks to control a cursor on a computer screen. Rapid movements of the cursor are made to targets arranged around a circle. A rotation of the cursors position is then introduced and participants adapt their movements to counter the error. Recent studies have investigated the impact of single-session transcranial direct current stimulation (tDCS) on this type of visuomotor learning (Galea et al. 2011). This problem has not been investigated in the context of speech motor adaptation. Speech motor adaptation is similar to visuomotor adaptation but it has important advantages - mainly, sensorimotor learning can be studied in a more complex movement without requiring participants to interact with a robotic arm or joystick. To that end, sixty young adults participated in a double-blind sham-controlled study involving a speech-motor-adaptation task and tDCS to either the cerebellum or primary motor cortex (M1).

Participants were asked to produce the words bed, dead, and head into a microphone. The first formant frequency (F1) of their vowel productions was then increased by 20played back to them over headphones with an unnoticeable delay. Compensatory changes in production in response to the auditory alteration were used to measure speech motor adaptation (Lametti et al. 2014ab). 16 minutes of 2-mA tDCS was applied to either the cerebellum or M1 throughout speech adaptation. Modulation of task performance related to stimulation would indicate a role (or roles) for these brain areas in speech motor adaptation. In the case of cerebellar stimulation, the anode was placed on the right cerebellum and the cathode was placed on the right cheek. For M1 stimulation the anode was placed over left articulatory motor cortex while the cathode was placed on the right forehead. Sham tDCS was applied to either the cerebellum or the motor cortex. The sixty participants were divided equally between the stimulation conditions. The participants and the experimenter did not know if real or sham stimulation was being applied.

Changes in the first two formants (F1 and F2) as well as changes in the distance in Hz between these measures were examined. When compared to sham stimulation, tDCS applied to either the cerebellum or M1 increased the rate and amount of speech motor adaptation as measured by changes in F1. Intriguingly, M1 stimulation drove a change in the distance between F1 and F2 that was not observed to the same extent in the cerebellar and sham stimulation groups. The effect of this change was to move participants productions closer to prototypical productions of bid, did, and hid. Taken together, the results show that modulation of activity in circuits shared by the cerebellum and motor cortex can facilitate speech motor adaptation.

References

Galea J.M., Vazquez A., Pasricha N., de Xivry J.J., Celnik P. (2011). Dissociating the roles of the cerebellum and motor cortex during adaptive learning: the motor cortex retains what the cerebellum learns. *Cereb Cortex, 21*(8), 1761-70.

Lametti D.R., Rochet-Capellan A., Neufeld E., Shiller D.M., Ostry D.J. (2014). Plasticity in the human speech motor system drives changes in speech perception. *J Neurosci, 34*(31), 10339-46.

Lametti D.R., Krol S.A., Shiller D.M., Ostry D.J. (2014). Brief periods of auditory perceptual training can determine the sensory targets of speech motor learning. *Psychol Sci, 25*(7), 1325-36.

MOTOR SKILL LEARNING AND LONG-TERM POTENTIATION IN ADULTS WHO STUTTER

Aravind Namasivayam[1,2], Kaviraja Udupa[1,3], Paul Saha[1,3], Vina Law[1], Anna Huynh[1], Robert Kroll[4], Pascal van Lieshout[1,2], Milos R. Popovic[1,2], Robert Chen[1,3]
[1] *University of Toronto, Toronto, Canada*
[2] *Toronto Rehabilitation Institute, Toronto, Canada*
[3] *Toronto Western Research Institute, Toronto, Canada*
[4] *The Speech and Stuttering Institute, Toronto, Canada*

Background Behavioural studies have suggested that adults who stutter (AWS) have limitations in learning of new hand and speech gestures. Despite these behavioural data, at a neurophysiological level, the evidence for limitations in motor learning in AWS is still lacking. This information is critical since limitations of motor learning have implications for treatment: some interventions with AWS rely heavily on learning of new speech patterns. We examined long-term potentiation (LTP)-like plasticity in the motor cortex of AWS using a neurophysiological paradigm known as paired associative stimulation (PAS). PAS involves repetitive peripheral median nerve stimulation followed by transcranial magnetic stimulation (TMS) of the contralateral motor cortex (M1 representation of the abductor pollicis brevis (APB) muscle) at 25-ms interstimulus intervals (PAS-25). This repeated pairing results in LTP-like changes as shown by an increase in amplitude and lower thresholds of motor evoked potentials (MEP) in the contralateral hand muscles following single pulse TMS.

Aim We tested the following two hypotheses: (a) AWS have motor learning deficits and they will demonstrate decreased PAS-25 induced LTP-like plasticity compared to adults who do not stutter (ANS), and (b) the extent of decreased PAS-25 induced LTP-like plasticity in AWS will correlate with decreased performance in a rotary pursuit motor learning task.

Participants 16 healthy adults (10 AWS and 6 ANS) matched for age, gender, educational level, and handedness participated in the study.

Procedure TMS was delivered via a 7-cm figure-of-eight coil connected to a Magstim 200 stimulator. MEPs were recorded from the contralateral right APB and abductor digiti minimi (ADM) muscles. Cortical excitability was measured by TMS at 5 time points at baseline, Time 0, 20, 60 min, and 1 week after PAS-25. Motor learning was measured using the rotary pursuit apparatus at baseline, 40 min and 1 week after PAS-25) to assess motor skill acquisition and learning at the behavioural level. Additionally, we tested several intracortical circuits including short interval intracortical inhibition, intracortical facilitation and long interval intracortical inhibition to examine impairments in the cortical neural circuits.

Results The results revealed a significant increase in MEP and behavioural variables after PAS-25 (peak increase at 60 min), as expected. No significant Group or Group x Time effects were found. However, there were trends of decreased slope of recruitment curve, altered inhibitory circuits and reduced plasticity effects in AWS compared to controls. The study is ongoing with more participants being recruited at this time.

Further reading

Chen, R., & Udupa, K. (2009). Measurement and modulation of brain plasticity of the motor system in humans using transcranial magnetic stimulation. *Motor Control, 13*, 422-453.

NEUROANATOMICAL DISTRIBUTION OF ADVERSE SPEECH EFFECTS DUE TO DEEP BRAIN STIMULATION

Linda Sandström[1], Patric Blomstedt[1], Lena Hartelius[2], Fredrik Karlsson[1]
[1]*Umeå University, Umeå, Sweden*
[2]*University of Gothenburg, Gothenburg, Sweden*

Background Impaired speech production is one of the most common adverse effects with thalamic deep brain stimulation (DBS) affecting approximately 24% of patients.(Alomar et al., 2017) It has been suggested that a crucial factor for both beneficial and adverse DBS effects is the spread of electrical stimulation, from the targeted structure to neighbouring brain regions(Fytagoridis, Åström, Samuelsson, & Blomstedt, 2016), but the neuroanatomical origin of adverse speech effects has been found to be elusive(Dembek et al., 2017). The absence of a clear spatial distribution of DBS-induced speech effects may in part be due to methodological issues, most importantly the lack of thorough and detailed investigations into how DBS affect speech motor control in the individual patient.

Aim We present a systematic stimulation evaluation designed to link reduced speech motor control in patients treated with DBS to neuroanatomical regions.

Method 36 patients diagnosed with Essential tremor treated with DBS targeting the posterior subthalamic area have been recorded 1) with DBS turned off, 2) with DBS turned on using the patient's ordinary clinical settings, and 3) in an experimental procedure with increasing stimulation amplitude, from 0.5V up to a maximum of 4.5V. A speech protocol, designed to provide a comprehensive, objective assessment of patients articulatory and phonatory proficiency, was performed in-between stimulation adjustments. Identified DBS-induced speech and voice effects for the individual patient will be evaluated in relation to the verified electrode position and simulations of the electrical field spread for each stimulation setting, thus linking specific speech and voice effects to neuroanatomical regions. Based on all patient data, a neuroanatomical map will be created indicating the probability of stimulation-induced adverse effects on speech.

Significance The presented study design is proposed to be well positioned to be able to increase our understanding of the contribution of neuroanatomical structures in the subthalamic region to speech motor control. The outcome may further be used to reduce DBS-induced adverse effects on speech, which would benefit all patients treated with DBS.

References

Alomar, S., King, N. K. K., Tam, J., Bari, A. A., Hamani, C., & Lozano, A. M. (2017). Speech and language adverse effects after thalamotomy and deep brain stimulation in patients with movement disorders: A meta-analysis. *Movement Disorders, 32*(1), 53-63.

Dembek, T. A., Barbe, M. T., Åström, M., Hoevels, M., Visser-Vandewalle, V., Fink, G. R., & Timmermann, L. (2017). Probabilistic mapping of deep brain stimulation effects in essential tremor. *NeuroImage Clinical, 13*, 164-173.

Fytagoridis, A., Åström, M., Samuelsson, J., & Blomstedt, P. (2016). Deep Brain Stimulation of the Caudal Zona Incerta: Tremor Control in Relation to the Location of Stimulation Fields. *Stereotactic and Functional Neurosurgery, 94*(6), 363-370.

DEPRESSED BRAIN, DEPRESSED SPEECH: LIMBIC SYSTEM INTERACTIONS WITH SPEECH MOTOR CONTROL

Kevin R. Sitek[1,2], Gregory A. Ciccarelli[1,2], Mathias Goncalves[1], Thomas F. Quatieri[1], Satrajit S. Ghosh[1]
[1]*Massachusetts Institute of Technology, Cambridge, United States of America*
[2]*Harvard University, Cambridge, United States of America*

Speech communication involves the complex coordination of linguistic, articulatory, and sensory processing. Since speech is often used to communicate emotionally salient information about the internal state of the speaker, speech production also requires the coordination of brain networks involved in emotional processing and internal state representation. These brain networks are largely impaired in emotional processing disorders such as major depressive disorder. Indeed, psychiatrists often observe speech production patterns to help diagnose specific symptoms, and machine learning algorithms are able to detect the severity of a speakers depression with only the acoustic features of speech (Williamson, Quatieri, Helfer, Ciccarelli, & Mehta, 2014).

We sought to understand the relationship between depression, affect, and speech. Using functional MRI, we investigated how individuals with and without depression produce implicitly emotional speech. We found that, when producing speech with emotional content, brain regions involved in speech production as well as medial prefrontal areas associated with emotional processing were more active. However, depressed participants activated regions of the canonical speech network more strongly than non-depressed participants.

Within the depressed participant group, depression severity was strongly related to activation in the right anterior insula when producing implicitly sad speech. This region has previously been implicated in both depression and speech processing: activation in the right insula is highly correlated with depression symptom severity in general, and activation and connectivity patterns during speech processing suggest it may serve as a hub between higher level cognitive processing and linguistic and motor speech planning, particularly with respect to prosody (Oh et al. 2014).

We also examined the effective connectivity between networks while performing the speech production task. In non-depressed participants, a seed in the amygdala (involved in both limbic processing and internal state representations) was less correlated with bilateral inferior frontal gyri (IFG, including "Brocas area" in the left hemisphere) and anterior insulae during speech production. Depressed participants showed even stronger anticorrelations between the amygdala and the right IFG and insula, extending into sensorimotor cortex.

These results suggest that internal state representations, which are overactive in major depressive disorder, must be inhibited in order to produce speech. Indeed, the right insula, which plays a crucial role in both depression and speech processing, is uniquely situated to modulate speech production in depression. Overall, our findings suggest a strong coupling between limbic and speech production networks in major depressive disorder.

References

Oh, A., Duerden, E. G., & Pang, E. W. (2014). The role of the insula in speech and language processing. *Brain and Language, 135*, 96-103.

Williamson, J. R., Quatieri, T. F., Helfer, B. S., Ciccarelli, G., & Mehta, D. D. (2014). Vocal and Facial Biomarkers of Depression Based on Motor Incoordination and Timing. In *Proceedings of the 4th International Workshop on Audio/Visual Emotion Challenge* (pp. 65-72). New York, NY, USA: ACM.

LANGUAGE LEARNING IN THE ADULT BRAIN: DISRUPTING THE DORSOLATERAL PREFONTAL CORTEX FACILITATES WORD-FORM LEARNING

Eleonore H.M. Smalle[1], Muriel Panouilleres[2], Arnaud Szmalec[1,3], Riikka Möttönen[2]
[1] *Université Catholique de Louvain, Louvain-la-Neuve, Belgium*
[2] *Oxford University, Oxford, United Kingdom*
[3] *Ghent University, Ghent, Belgium*

Adults do not learn languages as easily as children do. It has been hypothesized that the late-developing prefrontal cortex that supports executive functions competes with striatum-dependent procedural learning mechanisms that are important for language learning. To address this hypothesis, we tested whether a temporary neural disruption of the left Dorsolateral Prefrontal Cortex (DLPFC) can improve implicit, procedural learning of word-forms in adults. Young adults were presented with repeating audio-visual sequences of syllables for immediate serial recall in a Hebb repetition learning task that simulates word-form learning. Inhibitory theta-burst Transcranial Magnetic Stimulation was applied to the left DLPFC or to the control site before the Hebb task. The DLPFC-disrupted group showed enhanced learning of the novel phonological sequences relative to the control group. Moreover, learning was negatively correlated with executive functions that rely on the DLPFC in the control group, but not in the DLPFC-disrupted group. The results support the hypothesis that a mature prefrontal cortex competes with implicit learning of word-forms. The findings provide new insight into the competition between brain mechanisms that contribute to language learning in the adult brain.

References

Galea, J. M., Albert, N. B., Ditye, T., & Miall, R. C. (2010). Disruption of the dorsolateral prefrontal cortex facilitates the consolidation of procedural skills. *J Cogn Neurosci, 22*(6), 1158-1164.

Nemeth, D., Janacsek, K., Polner, B., & Kovacs, Z. A. (2013). Boosting Human Learning by Hypnosis. *Cereb Cortex, 23*(4), 801-805. doi:10.1093/cercor/bhs068

Poldrack, R. A., Clark, J., Pare-Blagoev, E. J., Shohamy, D., Creso Moyano, J., Myers, C., & Gluck, M. A. (2001). Interactive memory systems in the human brain. *Nature, 414*(6863), 546-550.

MAPPING NON-INVASIVE BRAIN ACTIVITY (MEG) SIGNALS TO SPOKEN PHRASES

Jun Wang[1], Myungjong Kim[1], Angel W. Hernandez-Mulero[2], Daragh Heitzman[3], Paul Ferrari[4,5]

[1] *University of Texas at Dallas, Dallas, United States of America*

[2] *Helen DeVos Childrens Hospital, Grand Rapids, United States of America*

[3] *Texas Neurology, Dallas, United States of America*

[4] *Dell Childrens Medical Center, Austin, United States of America*

[5] *University of Texas at Austin, Austin, United States of America*

Introduction Patients with locked-in-syndrome (fully paralyzed but aware) struggle in their daily communication. Enabling them a level of communication significantly improves their life quality. Current EEG-based BCI communication requires users to build words from single letters selected on a screen, which is extremely inefficient (Brumberg et al., 2010). A direct mapping from brain activity signals to text may provide a faster communication rate than current BCIs. While recent studies showed potential of implanted/invasive signals (Bouchard et al., 2013; Brumberg et al., 2010; Herff et al., 2016), decoding spoken speech from non-invasive signals has been rarely studied (Wang et al., 2017). This project explored phrase decoding from single-trial Magnetoencephalography (MEG) signals, which provides real-time temporal resolution and millimeter-level spatial resolution.

Method

Data collection. Neuromagnetic brain activity signals were recorded from three right-handed adult English speakers using a 306 channel Elekta Triux MEG machine. Stimuli were presented on a screen for 1 second (perception segment), followed by a 1 second fixation cross (preparation segment). Termination of the fixation heralded a blank screen that signaled the subject to overtly produce the phrase that was just previously shown (production segment). Subjects had 2 seconds to produce speech before the next stimulus was presented. Five short, daily used phrases (e.g., how are you doing?) were displayed in a pseudo-random order and multiple times. A total of 1,219 phrase samples (trials) were collected.

Data analysis. First, dynamic time warping and Wiener filtering were used to reduce noise. Then, Gamma band energy (30-100Hz) was calculated from 50 ms sliding windows with 25 ms steps and was fed into a machine learning classifier neural network. Cross validation, the standard testing strategy in machine learning, was used, where testing training samples (trials) were unique.

Results & Discussion The classification accuracy for the production segment with whole-head (200) gradiometers was 93.66% (SD = 1.61) and 61.99% (SD = 9.06) with speech related (20) gradiometers (including these around Brocas area, Wernickes area, and motor cortex). The accuracy for the pre-stimulus segment was just about chance level (20%): 43.76% (SD = 12.89%) for the whole 200 gradiometers and 33.02% (SD = 6.58) for the speech related 20 gradiometers. The production segment yielded a significantly higher level ($p < 0.05$), which demonstrated the possibility for decoding spoken speech directly from non-invasive (MEG) signals.

References

Brumberg, J. S., Nieto-Castanon, A., Kennedy, P. R., & Guenther, F. H. (2010). Brain-Computer Interfaces for Speech Communication. *Speech communication, 52*(4), 367-379.

Wang, J., Kim, M., Hernandez-Mulero, A. H., Heitzman, D., & Ferrari, P. (2017). Towards decoding speech production from single-trial MEG signals. In *Proc. IEEE ICASSP* (pp. 3036-40).

Conference organization

The **7^{th} International Conference on Speech Motor Control** is held in Groningen, the Netherlands, July 5 - 8, 2017.

The conference is organized by:
Center for Language and Cognition Groningen (CLCG), Faculty of Arts, University of Groningen; School of Behavioral and Cognitive Neurosciences (BCN), University Medical Centre Groningen; Department of Speech Language Pathology, University of Toronto; Utrecht Institute of Linguistics - OTS, Utrecht University

In collaboration with:
Motor Speech Conference, Madonna Rehabilitation Hospital, Lincoln, NE, USA.

Local organizing committee

- Ben Maassen, chair (CLCG & BCN, University of Groningen)
- Hayo Terband, co-chair (Utrecht Institute of Linguistics - OTS, Utrecht University)
- Mariska Pater, Christien Koopstra (Groningen Congres Bureau)

Venue
The conference will be held in the Academy Building of the University of Groningen, in the heart of the city-centre of Groningen. Visiting address: Broerstraat 5, Groningen.

Secretariat of the Conference
Groningen Congres Bureau
Att. Ms. Christien Koopstra
Address: Griffeweg 5
Postcode: 9724 GE Groningen
The Netherlands
Telephone: +31 (0) 50 3168877
E-mail: info@groningencongresbureau.nl
Conference website: www.slp-nijmegen.nl/smc2017

Acknowledgements

The conference organization gratefully acknowledges sponsoring of the conference by the following institutions:

- Royal Netherlands Academy of Arts and Sciences (KNAW)
- Centre for Language and Cognition Groningen (CLCG), University of Groningen
- School of Behavioural and Cognitive Neurosciences (BCN), University Medical Centre Groningen

The conference is also sponsored by Carstens Medizinelektronik GmbH, Germany