Jaargang 27, Supplement, augustus 2022

Stem-, Spraak- en Taalpathologie

8th International Conference on Speech Motor Control Groningen: Abstracts

rijksuniversiteit groningen / bibliotheek

Conference organization
The 8^{th} **International Conference on Speech Motor Control** is held in Groningen, the Netherlands, August 24 - 27, 2022.

The conference is hosted 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. There is close collaboration with the Motor Speech Conference, Madonna Rehabilitation Hospital, Lincoln, NE, USA, and we thank the organizers of the Conference on Motor Speech for their support.

The Program Committee consists of the topic chairs:
- Marina Laganaro (University of Geneva, Switzerland)
- Susanne Fuchs (Leibniz-Zentrum Allgemeine Sprachwissenschaft (ZAS), Berlin, Germany)
- Phil Hoole (Ludwig-Maximilians-University, Munich, Germany)
- Anja Lowit (University of Strathclyde, Glasgow, UK)
- Aude Noiray (CNRS, Laboratoire Dynamique du Langage, Univ. Lyon 2, Lyon, France)
- Aravind Namasivayam (University of Toronto, Canada)
- Angela Morgan (Murdoch Children's Research Institute & University of Melbourne, Australia)
- Adam Vogel (University of Melbourne, Australia)

The program committee is supported by the local (Dutch-speaking) organizing committee:
- Ben Maassen, chair (CLCG & BCN, University of Groningen, the Netherlands)
- Maria Valbuena, chair organization (Groningen Congres Bureau)
- Hayo Terband, co-chair (University of Iowa, Iowa City, United States of America)
- Edwin Maas (Temple University, Philadelphia, United States of America)
- Pascal van Lieshout (University of Toronto, Canada)
- Frits van Brenk (Utrecht University, the Netherlands)
- Teja Rebernik (University of Groningen, the Netherlands)

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. Maria Valbuena
Address: Griffeweg 5
Postcode: 9724 GE Groningen
The Netherlands
Telephone: +31 (0) 50 3168877
E-mail: info@gcb.nl
Conference website: `www.slp-nijmegen.nl/smc2022/`

Acknowledgments
The conference organization gratefully acknowledges sponsoring of the conference by the following institutions:
- Centre for Language and Cognition Groningen (CLCG), University of Groningen
- School of Behavioural and Cognitive Neurosciences (BCN), University Medical Centre Groningen

CLCG
Center for Language and Cognition Groningen

bcn

Stem-, Spraak- en Taalpathologie
Vol. 27, 2022, Supplement, pp. i-iv
https://doi.org/10.21827/32.8310/2022-115

Preface

This Supplement of 'Tijdschrift voor Stem-, Spraak- en Taalpathologie' (Journal of Voice, Speech and Language Pathology) contains the abstracts of the eighth edition of the International Conference on Speech Motor Control, held in Groningen, The Netherlands, August 24 - 27, 2022. With this eighth conference, a well-established Nijmegen (5 editions) - Groningen (6th & 7th edition) tradition continues. 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 five years since the previous conference in 2017 have yielded not only further insights in genetic, neural, physiological and developmental aspects of speech production, stuttering and other speech motor conditions, but have also advanced theoretical modelling. Combined with ongoing studies of populations that are increasingly better characterized genetically and neurobiologically, this quantitative boost of interdisciplinary results is now leading to a qualitative turning point in which large data sets are analyzed with powerful artificial intelligence and machine learning algorithms. The implementation of theories into computational models allows for the explicit testing of multifactorial interactions, thereby going beyond the more traditional single-factor experiments. Machine learning and the data sharing required to make this feasible are a special topic of the 2022 conference.

Brief history of previous conferences
Speech motor control is a dynamic research field. The tremendous, multidisciplinary progress made during the past decades is reflected in the Nijmegen - Groningen series of conferences. In the first edition in **1985**, the 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 motor speech conditions other than stuttering. At the third conference (**1996**), more emphasis was placed on the emerging field of brain imaging. In addition, development of speech motor control became a prominent topic. Since the fourth edition in **2001**, we witnessed the introduction of important neurophysiological and neurobehavioral concepts, and a growing interest in the 'interface' between higher order cognitive/psycholinguistic processes and speech production. Thus, the conferences of **2006**, **2011**, and **2017** have witnessed tremendous progress in integrating genetic, neurobiological, biomechanical, cognitive and behavioral levels of research in interdisciplinary collaborations. At the **2017** conference special attention was payed to the evolution of speech: phylogenetic evolution in homo sapiens; ontogenetic evolution in infants; and evolution of speech conditions and disorders in diverse contexts. With the exponentially growing computational power and data collection facilities, in recent years applications for diagnosis and monitoring of speech conditions by acoustic signal analysis algorithms have been developed, and trials have been run to test the validity of such approaches. This field is introduced as special topic at the **2022** conference under the label 'Machine Learning & Data Sharing'.

The 2022 conference program is organized around five topics; the invitation of speakers and the review process of submitted papers was coordinated by the chair and co-chair assigned to the specific topic. The topics are:

1. Neural Anatomy & Physiology of Speech Production (Chair: Marina Laganaro)
2. Speech Production Modeling & Action-Perception (Chairs: Susanne Fuchs & Phil Hoole)
3. Conditions Affecting Speech Motor Control (Chair: Anja Lowit)
4. (A)typical Speech Motor & Speech Sound Development in Children (Chairs: Aude Noiray & Aravind Namasivayam)
5. Machine Learning & Data Sharing (Chair: Angela Morgan, Adam Vogel)

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 organizers of this conference. In order to underscore this policy, ample time is scheduled for the poster sessions and special prizes are awarded for the most informative and well-designed posters.

In recognition of the importance of including a wide diversity of perspectives and experiences, and following recommendations during and after the previous conference, we have made considerable effort in forming a program committee and invited speaker line-up that is balanced with respect to gender, geographic location, and career stage, as well as theoretical orientation. We hope and trust that this effort will enhance both the scientific and social experience of the conference.

The organizing committee and colleagues in the field are keen on having an in-person conference. By having postponed the conference from 2021 to **August 2022**, we hope attendees feel free to interact at the conference and in social gatherings. The organization supports measures to make you feel safe during the conference, such as providing masks, hand sanitizers, self-tests, and badge-stickers to indicate your preferred way of social interaction. We recognize that despite these efforts, not everyone is able or willing to attend, and of course we respect the decision of potential attendees not to travel. Nevertheless, we believe we have assembled an extraordinary program and are confident that the conference will be stimulating and invigorating for those in attendance.

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, The Netherlands
Hayo Terband	Iowa City, USA
Edwin Maas	Philadelphia, USA

August 2022

https://doi.org/10.21827/32.8310/2022-115

Books/special issues of previous editions of the conference

Peters, H.F.M. & Hulstijn, W. (Eds.) (1987). *Speech motor dynamics in stuttering.* New York, USA: Springer Verlag.

Peters, H.F.M., Hulstijn, W. & Starkweather, C.W. (Eds.) (1991). *Speech motor control and stuttering.* Amsterdam, Netherlands: Excerpta Medica.

Hulstijn, W., Peters, H.F.M., & van Lieshout, P.H.H.M. (Eds.) (1996). *Speech production: motor control, brain research and fluency disorders.* Amsterdam, Netherlands: Elsevier.

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.

Maassen, B., Terband, H., Maas, E., & Namasivayam, A. (2019). Select Papers from the 7th International Conference on Speech Motor Control [Special issue]. *Journal of Speech, Language, and Hearing Sciences, 62*(8S).

Oral presentations

TOWARDS PROCESS-ORIENTED, DIMENSIONAL APPROACHES FOR DIAGNOSING SSD IN CHILDREN. STATE-OF-THE-ART AND FUTURE PERSPECTIVES.

Ben Maassen
Center for Language and Cognition, University of Groningen (CLCG)
Department of Neurosciences/BCN, University Medical Center Groningen (UMCG)

Children with speech sound disorders (SSD) form a heterogeneous group, with respect to severity, etiology, proximal causes, speech error characteristics and response to treatment. Infants develop speech and language in interaction with neurological maturation, and general perceptual, motoric and cognitive skills in a social-emotional context. Part 1 of this paper gives an overview and state-of-the-art of studies that aim to 'predict' speech phenomena and speech symptoms from underlying models. This includes studies on babbling (MacNeilage & Davis, 2001); on early stages in the hierarchy of motor control over articulatory mechanisms (Namasivayam et al., 2020; 2021); on the acquisition of systemic and phoneme-specific mappings according to the DIVA-model (Guenther, Ghosh, & Tourville, 2006); on computational modelling of feedback and feedforward control in childhood apraxia of speech (CAS) (Terband, Maassen, Guenther, & Brumberg, 2009); on co-articulation patterns within and between syllables in CAS (Nijland et al., 2003); and on kinematic studies of functional synergies and relative phase in syllabic contrasts, such as 'ta:p-ta:p' vs. 'pta:-pta:' or 'spa:-spa:' vs. 'pa:s-pa:s' (Nijland, Maassen, Hulstijn, & Peters, 2004; Terband, Maassen, van Lieshout, & Nijland, 2011). These studies challenge the phonological – phonetic dualism in mature speech production and speech development (Ziegler, 2016).

Part 2 presents the construction and first implementation studies of the Computer Articulation Instrument (CAI), a battery of speech production tasks to diagnose speech delay and subtypes of SSD in Dutch children. The CAI contains an analysis tool that automatically compares the phonetic transcriptions of speech targets with realizations. The CAI was designed and standardized to be able to assess and compare performance on four tasks: (a) picture naming, (b) nonword imitation, (c) word and nonword repetition, and (d) maximum repetition rate, thereby covering phonological and speech motor skills. Rather than focusing on single diagnostic markers, over 20 measures are assessed objectively and quantitatively, such as percentages consonants, vowels and syllable structures correct; phonological processes, and maximum repetition rates. All measures are normalized, and since the tasks are designed such that they tap into partly overlapping and partly different psycholinguistic and speech motor processes, the clinical result is a speech profile that can be interpreted in terms of strengths and weaknesses of the child's speech production system (Diepeveen, van Haaften, Terband, de Swart, & Maassen, 2019; van Haaften et al., 2019). Clinical studies will be presented in which the extracted speech parameters were analyzed by means of factor and cluster analyses. Results support a dimensional rather than categorical diagnostic model, in which clusters of children showed different performance profiles.

In order to strengthen the interpretation of the speech profile into underlying deficits, techniques are going to be implemented that directly interfere with particular subprocesses (Terband, Maassen, & Maas, 2019). Examples are: auditory or kinesthetic masking during speech output in a nonword imitation task; auditory feedback perturbation; speeding up speech rate; applying a dual-task condition.

In an ideal world, one would like to have longitudinal assessments of children with

SSD from their babbling stage onwards. In practice, we must work with a snapshot at the age of 3 or 4 years, possibly combined with further evaluation by means of response to intervention. With a process-oriented, dimensional analysis of speech skills, we hope to deepen our insight in the developmental mechanisms underlying the speech profile, and thereby obtain optimal guidelines for clinical intervention.

References

Diepeveen, S., van Haaften, L., Terband, H., de Swart, B., & Maassen, B. (2019). A standardized protocol for maximum repetition rate assessment in children. *Folia Phoniatrica Et Logopaedica.* `https://doi.org/10.1159/000500305`

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*(3), 280-301. `https://doi.org/S0093-934X(05)00115-X`

MacNeilage, P. F., & Davis, B. L. (2001). Motor mechanisms in speech ontogeny: Phylogenetic, neurobiological and linguistic implications. Current *Opinion in Neurobiology, 11*(6), 696-700.

Namasivayam, A. K., Huynh, A., Bali, R., Granata, F., Law, V., Rampersaud, D., & Hayden, D. (2021). Development and validation of a probe word list to assess speech motor skills in children. *American Journal of Speech-Language Pathology, 30*(2), 622-648. `https://doi.org/10.1044/2020_AJSLP-20-00139`

Namasivayam, A. K., Coleman, D., O'Dwyer, A., & van Lieshout, P. (2020). Speech sound disorders in children: An articulatory phonology perspective. *Frontiers in Psychology, 10*, 2998.

Nijland, L., Maassen, B., Hulstijn, W., & Peters, H. F. M. (2004). Speech motor coordination in dutch-speaking children with DAS studied with EMMA. *Journal of Multilingual Communication Disorders, 2*(1), 50-60.

Nijland, L., Maassen, B., Van der Meulen, S., Gabreëls, F., Kraaimaat, F. W., & Schreuder, R. (2003). Planning of syllables by children with developmental apraxia of speech. *Clinical Linguistics and Phonetics, 17*(1), 1-24.

Terband, H., Maassen, B., Guenther, F. H., & Brumberg, J. (2009). Computational neural modeling of speech motor control in childhood apraxia of speech (CAS). *Journal of Speech, Language, and Hearing Research, 52*(6), 1595-1609.

Terband, H., Maassen, B., & Maas, E. (2019). A psycholinguistic framework for diagnosis and treatment planning of developmental speech disorders. *Folia Phoniatrica Et Logopaedica, 71*, 216-227. `https://doi.org/10.1159/000499426`

Terband, H., Maassen, B., van Lieshout, P., & Nijland, L. (2011). Stability and composition of functional synergies for speech movements in children with developmental speech disorders. *Journal of Communication Disorders, 44*(1), 59-74. `https://doi.org/10.1016/j.jcomdis.2010.07.003`

van Haaften, L., Diepeveen, S., van den Engel, L., Jonker, M., de Swart, B., & Maassen, B. (2019). The psychometric evaluation of a speech production test battery for children: The reliability and validity of the computer articulation instrument. *Journal of Speech, Language, and Hearing Research, 62*(7), 2141-2170. `https://doi.org/10.1044/2018_JSLHR-S-18-0274`

Ziegler, W. (2016). Phonology versus phonetics in speech sound disorders. In P. Van Lieshout, B. Maassen & H. Terband (Eds.), *Speech motor control in normal and disordered speech: Future developments in theory and methodology* (pp. 250-288). Rockville, MD: ASHA Press.

GENETIC ARCHITECTURE OF CHILD SPEECH DISORDER

Angela Morgan
Speech and Language, Murdoch Children's Research Institute
Department of Audiology and Speech Pathology, University of Melbourne
Speech Pathology Department, Royal Children's Hospital

Childhood apraxia of speech (CAS) is the prototypic severe childhood speech disorder. CAS is characterized by motor programming and planning deficits. Knowledge of the aetiology of CAS was limited for decades, and largely since the condition was first recognised over 70 years ago. Fortunately, due to technological advances in genetic sequencing and analysis methods, there has been an exponential leap in our understanding of the genetic bases of CAS in the last four years. We now understand that CAS has a genetic basis, with a genetic cause identified in a third of cases (Eising et al., 2019; Hildebrand et al., 2020). Over 30 single candidate genes and many more chromosomal conditions have been implicated in CAS to date. Here we provide an overview of research on the genetic basis of CAS and also present new data on gene findings in a large cohort of 70 individuals with CAS (Kaspi et al., in press). Genetic conditions were identified in 18/70 (26%) probands, almost doubling the current number of candidate genes for CAS. Three of the 18 variants affected genes previously identified in the earlier cohorts; *SETBP1*, *SETD1A* and *DDX3X* (Eising et al., 2019; Hildebrand et al., 2020), thus confirming the role of these genes in explaining CAS, while the remaining 15 had not been previously associated with this disorder. Further insights into the biology of CAS are discussed, such as highlighting the roles of chromatin organization and gene regulation in child speech disorder, and confirming that genes involved in CAS are co-expressed during brain development (Eising et al., 2019; Hildebrand et al., 2020; Kaspi et al., in press). Further, this third and largest cohort confirmed a diagnostic yield comparable to, or even higher, than other neurodevelopmental disorders with substantial *de novo* variant burden such as epilepsy or autism (Kaspi et al., in press). Data also support the increasingly recognised overlaps between genes conferring risk for a range of neurodevelopmental disorders. Understanding the aetiological basis of CAS is critical to end the diagnostic odyssey, identify comorbidities, open doors to further societal support and ensure patients are poised for precision medicine trials.

References

Eising, E., Carrion-Castillo, A., Vino, A., Strand, E.A., Jakielski, K.J., Scerri, T.S. et al. (2019). A set of regulatory genes co-expressed in embryonic human brain is implicated in disrupted speech development. *Mol Psychiatry, 24*(7), pp. 1065-1078.

Hildebrand MS, Jackson VE, Scerri TS, Van Reyk O, Coleman M, Braden RO et al. (2020). Severe childhood speech disorder: Gene discovery highlights transcriptional dysregulation. *Neurology, 94*(20): e2148-e2167.

Kaspi, A., Hildebrand, M.S., Jackson, V.E., et al. (in press). Genetic aetiologies for childhood speech disorder: novel pathways co-expressed during brain development. *Molecular Psychiatry.*

DATA-DRIVEN CARE PATHWAY FOR CHILDREN OVER 36 MONTHS OF AGE WITH MOTOR SPEECH DISORDERS.

Aravind K Namasivayam[1], Margit Pukonen[2], Pascal van Lieshout[1]
[1] *University of Toronto, Toronto, Canada*
[2] *The Speech and Stuttering Institute, Toronto, Canada*

A significant body of research has been carried out in the area of optimal intervention parameters for subtypes of speech sound disorders (SSD) in children (e.g., Allen, 2013; Baker 2012; Williams, 2012). However, there is only limited information available regarding optimal intervention protocols for children with Motor Speech Disorders (MSD). We report outcomes from a large-scale motor speech research project (N = 98) funded by the Ontario Ministry of Children, Community and Social Services. The project assessed treatment outcomes and factors affecting these treatment outcomes in children with Motor Speech Disorders (MSD) between the ages of 3 and 10 years. The aim of the project was three-fold: (a) To establish the magnitude of treatment effects for outcome measures related to the speech sound system (articulation/phonology), speech intelligibility and functional communication in children with MSD including those with and without childhood apraxia of speech (CAS), (b) Explore the relationship between treatment dose frequency (1x vs. 2x week treatment sessions) and outcome measures, to aid the development of an evidence-based service model or "care pathway" for children with MSD living in Ontario, and (c) To identify factors that contribute to positive treatment outcomes in children with MSD.

Overall, the results indicated that the motor-speech treatment provided (motor speech treatment protocol; Namasivayam et al., 2015a) was effective in improving articulation, speech intelligibility and functional outcomes in children with MSD. In general, the magnitude of change for the higher dose frequency groups (2x/ week) was larger compared to the lower dose frequency groups (1x/week). However, for children with CAS only higher dose frequency (2×/week) led to significantly better outcomes for articulation and functional communication compared with lower 1×/week intervention (Namasivayam et al., 2015b). Both low- and high-dose frequency treatments yielded similar results for children with Speech Motor Delay (Namasivayam et al., 2019). Recently (Namasivayam et al., in prep), we estimated the association between minimal clinically important difference (MCID) in functional outcomes and multiple predictors (diagnostic features, amount of home training in minutes, intervention dose, age of child, gender, severity (percent consonants correct; Goldman-Fristoe Test of Articulation–Second Edition- Standard score)) using multivariable logistic regression models. The presence or absence of CAS features (on the diagnostic rating scale of the Kaufman Speech to Praxis Test) was the strongest and only significant predictor of whether a child demonstrated clinically relevant change in functional outcomes at the end of treatment. The presence of CAS features typically resulted in limited treatment gains. The research project has led to the development of a evidence-based pilot care pathway to support identification and intervention planning for children over 36 months of age with MSD. The evidence-based care pathway is currently in the clinical implementation phase in the province of Ontario, Canada.

References

Allen, M. M. (2013). Intervention efficacy and intensity for children with speech sound disorders. *Journal of Speech, Language and Hearing Research, 56*, 865–877.

Baker, E. (2012). Optimal intervention intensity in speech–language pathology: discoveries, challenges, and unchartered territories. *International Journal of Speech-Language Pathology, 14*(5), 478–485.

Namasivayam, A.K., Pukonen, M., Hard, J., Jahnke, R., Kearney, E., Kroll, R., & Van Lieshout, P.H.H.M. (2015a). Motor Speech Treatment Protocol for Developmental Motor Speech Disorders. *Developmental Neurorehabilitation, 18*(5), 296–303.

Namasivayam, A. K., Pukonen, M., Goshulak, D., Hard, J., Rudzicz, F., Rietveld, T., Maassen, B., Kroll, R., & Van Lieshout, P.H.H.M. (2015b). Treatment Intensity and Childhood Apraxia of Speech. *International Journal of Language & Communication Disorders, 50*(4), 529-46.

Namasivayam, A.K., Pukonen, M., Goshulak, D., Francesca, G., Le, D. L., Kroll, R., & Van Lieshout, P. H. H. M. (2019). Investigating intervention dose frequency for children with speech sound disorders and motor speech involvement. *International Journal of Language & Communication Disorders, 54*(4), 673–686.

Williams, A. L. (2012). Intensity in phonological intervention: is there a prescribed amount? *International Journal of Speech-Language Pathology, 14*(5), 456–461.

SPEECH AND OROFACIAL BIOMARKERS AND TOOLS IN THE ASSESSMENT OF NEUROLOGICAL DISEASES

Yana Yunusova
University of Toronto, Toronto, Canada

Speech is a window into the neurological health of an individual. It is established that changes in speech can be among the earliest signs and symptoms of a neurological disease; these changes follow a predictable pattern that can describe the course of disease progression and point to different profiles in disease presentation. Traditionally, experts in motor speech disorders played a key role in differential diagnosis of neurological diseases through providing perceptual (visual & auditory) assessment of speech and orofacial functions. With realization that reliable perceptual assessment requires a highly advanced skill set and with development of technology, instrumental speech and orofacial analytics emerged as a potential front runner in clinical assessment as well as a sought-after component in clinical trials.

Researchers have developed methods to extract, automatically, quickly and accurately, granular and reliable acoustic and kinematic measurements from a large number of speakers and tasks. Novel mathematical/ artificial intelligence approaches have enabled the extraction of thousands of features, or properties, per sample. Based on our research in amyotrophic lateral sclerosis, Parkinson's disease, and stroke, we demonstrated that automatic speech and orofacial assessment show a great promise in detecting early disease state, sensitivity to disease progression, and a potential to show disease subtypes. Several needs/ goals emerged in the process, however.

These needs/goals are to (1) validate consumer-grade/ easy to use technologies to achieve clinical goals; (2) adopt, through additional development and fine-tuning, existing algorithms to clinical populations; (3) establish a framework for feature selection; (4) develop accurate clinical characterizations of various diseases and dysarthria types; and (5) consider the fundamentals of tool/ biomarker development. Digital measures can be used to construct digital biomarkers when they demonstrate a relationship with a physiological construct of interest, linked to a biological or pathological process, and are clinically interpretable. To achieve these goals, these measures must undergo a rigorous assessment of their psychometric properties, following established measurement development frameworks (e.g., The COnsensus-based Standards for the selection of health Measurement INstruments or COSMIN). In this talk, we will discuss the steps addressing the above goals in a stepwise fashion, based on the history of our own research and research in the field.

SPEECH MOTOR PROFILES IN PRIMARY PROGRESSIVE APHASIA

Anja Staiger[1], Matthias L. Schroeter[2], Wolfram Ziegler[1], Danièle Pino[2], Frank Regenbrecht[2], Theresa Schölderle[1], Theresa Rieger[1], Felix Müller-Sarnowski[3,4], Janine Diehl-Schmid[3,5,6]
[1]*Ludwig-Maximilians-University, Munich, Germany*
[2]*Max-Planck-Institute for Human Cognitive and Brain Sciences Leipzig & University Hospital Leipzig, Leipzig, Germany*
[3]*Technische Universität, Munich, Germany*
[4]*Augsburg University, Augsburg, Germany*
[5]*Munich Cluster for Systems Neurology (SyNergy), Munich, Germany*
[5]*kbo-Inn-Salzach-Klinikum Wasserburg, Wasserburg, Germany*

Background & objectives In the last three decades, increasing attention has been paid to motor speech disorders (MSD) in primary progressive aphasia (PPA). While the main focus has been on progressive apraxia of speech, much less is known about dysarthria and other forms of motor speech impairment. Moreover, there have been virtually no studies that detailed the overall speech motor profiles of individuals with PPA, and if so, they have been exclusively related to persons with the non-fluent variant of PPA (Ogar et al., 2007). The aim of the present study was to investigate the speech motor profiles across the auditory dimensions of respiration, voice, articulation, prosody, as well as speech behavior in a prospectively collected sample of individuals with PPA independent of subtype.

Methods So far, we included 38 German speaking individuals with a neurological root diagnosis of PPA (18f, mean disease duration 3.3 y). Neuropsychological and language data were collected to assess all consensus-proposed diagnostic domains (Gorno-Tempini et al., 2011). We purposely refrained from an a-priori classification into subtypes, since motor speech dysfunction is among the criteria on which, according to current "gold standard", subtyping of PPA is based.

Speech tasks comprised various speech modalities and levels of speech motor complexity (repetition of structurally simple and complex words, repetition of sentences of increasing length, spontaneous speech, picture description). Auditory speech analyses were performed by three expert raters (AS, WZ, TS) according to a novel protocol that aimed at capturing speech motor symptoms as inclusively as possible. In addition, syndrome and severity of MSDs were rated.

Results Preliminary data suggests that more than 40% of the participants had some form of an MSD. In addition to apraxia of speech (prosodic and phonetic-prosodic types), we observed different dysarthria syndromes, neurogenic stuttering, adynamic speech, as well as mixed forms. Degrees of severity ranged from mild to severe, with moderate severity being most common. Performance in the different speech dimensions showed a great heterogeneity in the patient sample. We found MSDs also in patients whose language profiles were not compatible with the non-fluent subtype of PPA.

Discussion The results confirm previous findings that MSDs are common in PPA and can manifest in different syndromes (Duffy et al., 2014; Staiger et al., 2021). The findings emphasize that future studies of MSDs in PPA should be extended to all clinical variants and should take into account the qualitative characteristics of motor speech dysfunction across speech dimensions.

References

Duffy, J. R., Strand, E. A., & Josephs, K. A. (2014). Motor speech disorders associated with primary progressive aphasia. *Aphasiology, 28*(8-9), 1004-1017.

Gorno-Tempini, M., Hillis, A., Weintraub, S., Kertesz, A., Mendez, M., Cappa, S. E., ...& Boeve, B. (2011). Classification of primary progressive aphasia and its variants. *Neurology, 76*(11), 1006-1014.

Ogar, J. M., Dronkers, N. F., Brambati, S. M., Miller, B. L., & Gorno-Tempini, M. L. (2007). Progressive nonfluent aphasia and its characteristic motor speech deficits. *Alzheimer Disease & Associated Disorders, 21*(4), S23-S30.

Staiger, A., Schroeter, M. L., Ziegler, W., Schölderle, T., Anderl-Straub, S., Danek, A., ...& Jahn, H. (2021). Motor speech disorders in the nonfluent, semantic and logopenic variants of primary progressive aphasia. *Cortex, 140*, 66-79.

DISEASE- AND TREATMENT-RELATED CHANGES OF TONGUE BODY MOVEMENTS IN PARKINSON'S DISEASE: AN ELECTROMAGNETIC ARTICULOGRAPHY STUDY

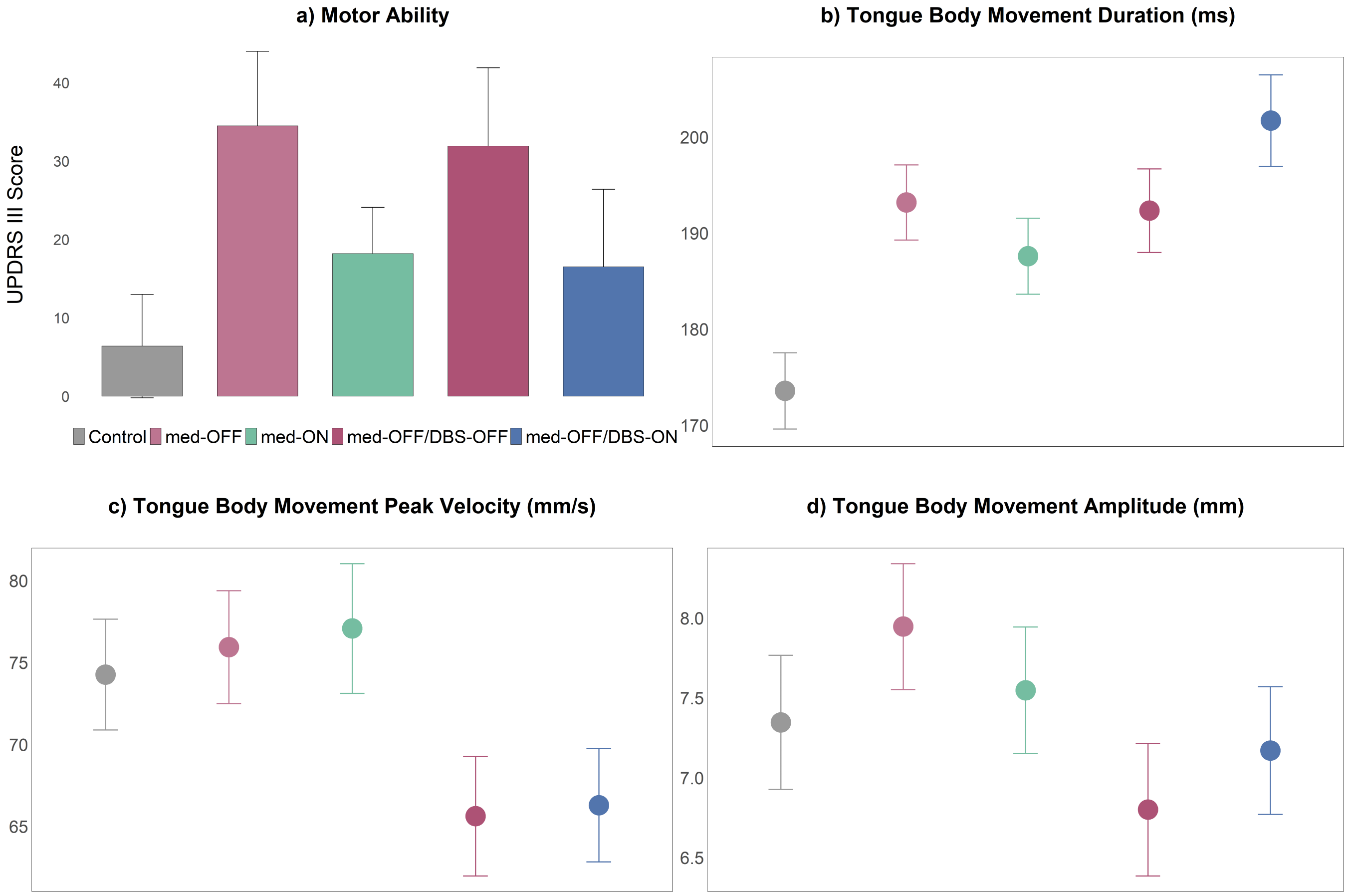
Tabea Thies[1], Michael T. Barbe[1], Veerle Visser-Vandewalle[2], Doris Mücke[1]
[1] *University of Cologne, Cologne, Germany*
[2] *University Hospital Cologne, University of Cologne, Cologne, Germany*

Introduction Parkinson's Disease (PD) affects not only gross motor but also speech motor control. Speakers with PD exhibit slower, smaller and unprecise articulatory movements, resulting in an overall reduced vowel space and less intelligible speech (Duffy, 2019). The reduced vowel space is related to overall disease severity, but can increase with dopaminergic treatment (Thies et al., 2020; Thies et al., 2021). Standard treatment options are levodopa intake or – in later disease stages – a deep brain stimulation (DBS). For a DBS, two electrodes are implanted at specific brain areas, such as the nucleus subthalamicus (STN). The effect of STN-DBS on motor speech is under debate. This study aims to investigate the effect of PD, levodopa, and STN-DBS on vocalic tongue movements in speakers with PD.

Methods 13 patients which were diagnosed with PD 8 (± 5) years prior to study inclusion and 13 age- and sex-matched healthy controls participated in the study (4 females, 9 males, 59 ± 9 years). All participants were recorded with an electromagnetic articulograph (AG 501) producing ten different target words embedded in a carrier sentence in controlled broad focus condition. Target words were disyllabic girl names (C1V1.C2V2-structure). The vowel V1 was one of the following five German vowels: /i, e, a, o, u/. Acoustic vowel durations of V1 and underlying tongue body kinematics (durations, amplitudes, peak velocities) were calculated. Whereas, the healthy speakers were recorded once, the speakers with PD were recorded in four different treatment conditions: Before the implantation of DBS electrodes in medication-OFF (without levodopa) and medication-ON condition (with a standardized dose of soluble levodopa), and 8 (± 2.8) months postoperatively without any medication in DBS-OFF (deactivated DBS) and DBS-ON condition (activated DBS). To determine the motor ability of all participants, part III of the Unified Parkinson's Disease Scale was assessed (UPDRS, Fahn & Elton, 1987); higher scores indicate a more severe motor impairment.

Results As Figure 1a depicts, UPDRS scores are higher in individuals with PD compared to healthy controls, but decrease in therapeutic conditions (med-ON, DBS-ON). Figure 1b-d reveal longer, slower and larger tongue body movements in speakers with PD compared to healthy speakers (control vs. med-OFF). Comparing med-OFF and med-ON condition, tongue movements are shorter, faster and smaller. Speech movements after the DBS-implantation are slower and smaller (med-OFF vs. DBS-OFF). The activated stimulation leads to longer, faster and larger tongue movements (DBS-OFF vs. DBS-ON). Whereas, acoustic vowel durations are longer in speakers with PD compared to controls, they do not change with therapy; but acoustic durations increase comparing med-OFF and DBS-OFF condition.

Conclusion Both, levodopa and activated DBS, improve speech motor control in this sample of patients with PD. Under levodopa, tongue body movements are more agile (shorter and faster). With activated DBS, movement duration and amplitude increase without changing agility (speed). Our data suggests that the mere presence of the DBS electrodes (med-OFF vs. DBS-OFF) worsens speech ability, as postoperative movement patterns are slower and smaller compared to the preoperative state.



a) UPDRS III motor scores across the groups and conditions. Higher scores indicate a more severe motor impairment. b-d) Means and standard errors of articulatory speech parameters per group and condition.

References

Duffy, J. R. (2019). *Motor Speech Disorders: Substrates, Differential Diagnosis, and Management*, 4th ed. Edinburgh: Elsevier.

Fahn, S., & Elton, R. L. (1987). UPDRS program members. Unified Parkinsons disease rating scale. *Recent developments in Parkinson's disease, 2*, 153-163.

Thies, T., Mücke, D., Lowit, A., Kalbe, E., Steffen, J., & Barbe, M. T. (2020). Prominence Marking in Parkinsonian Speech and Its Correlation with Motor Performance and Cognitive Abilities. *Neuropsychologia 137*: 107306.

Thies, T., Mücke, D., Dano, R., & Barbe, M. T. (2021). Levodopa-Based Changes on Vocalic Speech Movements during Prosodic Prominence Marking. *Brain Sciences 11*: 594.

SPATIAL AND TEMPORAL VARIABILITY OF SPEECH GESTURES DURING FAST SYLLABLE REPETITION IN PARKINSON'S DISEASE: AN ARTICULATORY STUDY

Jidde Jacobi[1,2], Teja Rebernik[1,3], Roel Jonkers[1], Ben Maassen[1], Michael Proctor[2], Martijn Wieling[1,4]

[1] *University of Groningen, Groningen, The Netherlands*

[2]*Department of Cognitive Science, Macquarie University, Sydney, Australia*

[3]*Netherlands Cancer Institute, Amsterdam, The Netherlands*

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

Introduction Fast syllable repetition tasks are commonly used in the assessment of speech difficulties in Parkinson's disease (Duffy, 2013). In such tasks, subjects are asked to repeat monosyllables such as /pa/ as quickly as possible. Previous studies have demonstrated that some individuals with Parkinson's disease (PD) show atypical temporal patterns, such as lower or higher syllable repetition rates, in comparison to typical speakers (Canter, 1963; Gurd et al., 1998; Hirose et al., 1981). Additionally, self-paced timing of syllable production has been found to be impaired in fast syllable repetition tasks performed by individuals with PD (Ackermann et al., 1997). The goal of the present study was to replicate these findings and to additionally examine articulatory kinematics during a fast syllable repetition task. More specifically, we examined temporal variability, spatial variability, and the relationship between the two. The kinematic data that was acquired in this study allowed for a fine-grained examination of speech at the articulatory level. This study, therefore, provides new insight into the ways in which the speech motor system is disordered in PD.

Method We acquired kinematic speech data using electromagnetic articulography (EMA) during a fast syllable repetition task that was performed by both individuals with PD (n = 24) and typical speakers (n = 25). Participants were asked to repeat the syllables /pɑ/, /tɑ/, and /kɑ/ as fast as possible. EMA sensors were placed on the tongue, lips, and jaw to track articulatory trajectories. A measure of spatial accuracy was obtained by calculating the bivariate endpoint deviation for the relevant articulator, and a measure of temporal accuracy was obtained by calculating the standard deviation of the time intervals between successive repetitions. Generalized additive modeling (GAM) was used to examine group differences.

Results We found lower spatial variability and higher temporal variability in the speech data of individuals with PD. In addition, we found an atypical relationship between temporal variability and spatial variability in their speech productions. We believe that some individuals with PD in this study may have been unable to successfully regulate articulatory timing during the task, and therefore allocated most of their speech motor resources to spatial accuracy rather than to syllable timing. We suggest that this may be due to deficiencies in the speech-motor regulation mechanism that underlies speech production in PD.

References

Ackermann, H., Konczak, J., & Hertrich, I. (1997). The temporal control of repetitive articulatory movements in Parkinson's disease. *Brain and Language, 56*(2), 312–319. https://doi.org/10.1006/brln.1997.1851

Canter, G. J. (1963). Speech characteristics of patients with Parkinson's disease: I. Intensity, pitch, and duration. *Journal of Speech and Hearing Disorders, 28*(3), 221–229. https://doi.org/10.1044/jshd.2803.221

Duffy, J. R. (2013). *Motor speech disorders: Substrates, differential diagnosis, and management.* Elsevier.

Gurd, J. M., Bessell, N., Watson, I., & Coleman, J. (1998). Motor speech versus digit control in Parkinson's disease: A cognitive neuropsychology investigation. *Clinical Linguistics & Phonetics, 12*(5), 357–378. https://doi.org/10.1080/02699209808985231

Hirose, H., Kiritani, S., Ushijima, T., Yoshioka, H., & Sawashima, M. (1981). Patterns of dysarthric movements in patients with Parkinsonism. *Folia Phoniatrica et Logopaedica, 33*(4), 204–215. https://doi.org/10.1159/000265595

WHY PROSODY MATTERS IN APRAXIA OF SPEECH - THEORETICAL AND CLINICAL ISSUES

Ingrid Aichert
Ludwig-Maximilians-University, Munich, Germany

Speaking is a motor activity of unrivalled complexity that we acquire during childhood and, as adults, master effortlessly, with a high automatedness, and largely without errors. An integral part of our ability to articulate fluently and comprehensibly is the rhythmic structuring of speech (Poeppel & Assaneo, 2020).

Apraxia of speech is a clinical model of how this ability breaks down after lesions to brain areas engaged in speech motor planning. It causes a loss of the "ease of articulation" that characterizes speech production in neurologically healthy adults and a variety of symptoms affecting the integrity of speech sounds and the rhythmic and melodic properties of spoken language. It has long been recognized that prosodic disturbances are among the primary clinical characteristics of apraxia of speech (McNeil et al., 1997). Interestingly, it has long been known that speech apraxic symptoms, especially articulation errors at a segmental level, can be positively modulated by interventions that focus on the rhythmical aspects of speaking (Wambaugh, 2021), even though there has been no theoretical model to explain the transfer of rhythmic interventions to the phonemic level. One of the most traditional rhythmic-melodic approaches in language therapy, Melodic Intonation Therapy, has recently been shown to be particularly useful for patients with apraxia of speech (Zumbansen & Tremblay, 2019).

Based on the premise that apraxia of speech provides a window into the internal make-up of learned speech motor behavior, I will present evidence for an interaction of segmental and prosodic information in speech planning by analyzing factors that influence the occurrence of speech errors in apraxic speakers (e.g., Aichert et al., 2016). These data will be discussed as evidence for the assumption that metrical aspects of articulation are part of speech motor plans (Ziegler et al., 2021). Furthermore, I will present evidence for an influence of the rhythm of perceived speech on segment-level accuracy in speech production (Aichert et al., 2019, 2021), which points to action-perception links of speech rhythm in apraxia of speech.

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. https://doi.org/10.1016/j.neuropsychologia.2016.01.009

Aichert, I., Lehner, K., Falk, S., Späth, M. & Ziegler, W. (2019). Do patients with neurogenic speech sound impairments benefit from auditory priming with a regular metrical pattern? *Journal of Speech Language and Hearing Research, 62*, 3104–3118. https://doi.org/10.1044/2019_JSLHR-S-CSMC7-18-0172

Aichert, I., Lehner, K., Falk, S., Franke, M., Ziegler, W. (2021). In Time with the Beat: Entrainment in Patients with Phonological Impairment, Apraxia of Speech, and Parkinson's Disease. *Brain Sciences, 11*, 152-4. https://doi.org/10.3390/brainsci11111524

McNeil, M. R., Robin, D. A., & Schmidt, R. A. (1997). Apraxia of speech: definition, differentiation and treatment. In M. R. McNeil (Ed.), *Clinical Management of Sensorimotor Speech Disorders* (pp. 311-344). Thieme.

Poeppel, D. & Assaneo, M.F. (2020). Speech rhythms and their neural foundations.

Nature Reviews Neuroscience, 21, 322–334 (2020). https://doi.org/10.1038/s41583-020-0304-4

Wambaugh, J. L. (2021). An expanding apraxia of speech (AOS) treatment evidence base: An update of recent developments. *Aphasiology, 35*(4), 442-461. https://doi.org/10.1080/02687038.2020.1732289

Ziegler, W., Lehner, K., Pfab, J., & Aichert, I. (2021). The Nonlinear Gestural Model of Speech Apraxia: Clinical Implications and Applications. *Aphasiology, 35*(4), 462-484. https://doi.org/10.1080/02687038.2020.1727839

Zumbansen, A., & Trembley, P. (2019). Music-based interventions for aphasia could act through a motor-speech mechanism: a systematic review and case–control analysis of published individual participant data. *Aphasiology, 33*(4), 466-497. https://doi.org/10.1080/02687038.2018.1506089

AN FMRI STUDY OF OVERT SPEECH PREPARATION AND PRODUCTION IN CHILDREN WHO STUTTER

Ho Ming Chow[1], Emily Garnett[2], Nan Ratner[3], Soo-Eun Chang[2]
[1]*University of Delaware, Newark, Delaware, United States of America*
[2]*University of Michigan, Ann Arbor, Michigan, United States of America*
[3]*University of Maryland, College Park, Maryland, United States of America*

Recent neuroimaging studies of children who stutter (CWS) have indicated that developmental stuttering is associated with functional and structural anomalies in the cortical and subcortical areas classically associated with speech motor control (Neef et al., 2015). FMRI is arguably the best non-invasive tool for the investigation of cortical and subcortical motor area activity during speech production. However, a major obstacle that hinders the use of fMRI to study speech production is that speech-related movements are a major source of contamination in fMRI. Thus, most of the previous fMRI studies employed relatively simple speaking tasks (e.g., single-word production) or non-speaking contexts (e.g., wakeful rest) to examine speech motor functions. However, these procedures do not optimally inform the neural correlates of childhood stuttering because moments of stuttering typically do not occur during simple speaking tasks (Usler & Walsh, 2018). Spontaneous, continuous speech, on the other hand, is ecologically-valid for studying the disorder because symptoms of stuttering are most apparent in this context. However, continuous speech production is also more likely to induce severe movement-related artifacts that contaminate the fMRI signal, especially in children. To overcome this obstacle, we used a technique, based on spatial independent component analysis (sICA), to remove fMRI artifacts during continuous speech production (Xu et al., 2014). This technique decomposes individual functional images into a number of components. Components with spatial biologically implausible patterns could be identified automatically and removed from the functional images.

Using this denoising technique, we examined two aspects of speech production known to induce dysfluencies in CWS, formulation and preparation of speech. To examine the effect of speech formulation, we compared brain activation during continuous (propositional) and automatic (non-propositional) speech production in CWS with their fluent peers. For the effect of speech preparation, we compared brain activation during the period shortly before (2 seconds) the initiation of speech in CWS and controls. We also examined age-related changes in brain activity patterns during speech preparation and production. In total, our final analysis included 28 CWS (13 boys and 15 girls) and 24 controls (11 boys and 13 girls) in the 5-12.5 age range.

Comparing continuous and automatic speech conditions, the control group exhibited increased left-lateralized activation in the inferior frontal gyrus, premotor cortex and pre-supplementary motor area as well as the bilateral posterior middle/inferior temporal gyri. The direct group comparison in the basal-ganglia-thalamocortical (BGTC) loop showed that CWS exhibited decreased activation in the left premotor cortex (PMC). Activation in the bilateral PMC during speech production furthermore increased with age in controls but decreased with age in CWS. Age-related activation decreases were observed in the putamen and the thalamus for CWS during speech preparation. Taken together, the results support the notion that stuttering is associated with aberrant development and functional deficits in the BGTC loop affecting the preparation and production of spontaneous speech.

References

Neef, N. E., Anwander, A., & Friederici, A. D. (2015). The neurobiological grounding of persistent stuttering: from structure to function. *Current Neurology and Neuroscience Reports, 15*(9), 63. https://doi.org/10.1007/s11910-015-0579-4

Usler, E. R., & Walsh, B. (2018). The effects of syntactic complexity and sentence length on the speech motor control of school-age children who stutter. *Journal of Speech, Language, and Hearing Research, 61*(9), 2157–2167. https://doi.org/10.1044/2018_JSLHR-S-17-0435

Xu, Y., Tong, Y., Liu, S., Chow, H. M., AbdulSabur, N. Y., Mattay, G. S., & Braun, A. R. (2014). Denoising the speaking brain: toward a robust technique for correcting artifact-contaminated fMRI data under severe motion. NeuroImage, 103, 33–47. https://doi.org/10.1016/j.neuroimage.2014.09.013

INNER SPEECH AS AN EXAPTATION OF HIERARCHICAL PREDICTIVE CONTROL OF SPEECH: ARTICULATING CONDENSATION, DIALOGALITY AND INTENTIONALITY

Hélène Lœvenbruck
CNRS, Université Grenoble Alpes, Grenoble, France

And I think to myself
What a wonderful world

As you silently read these lyrics, can you hear Louis Armstrong's voice? Who covertly speaks to whom when we think or say to ourselves? How can we control for our various inner voices?

Inner speech, also known as interior monologue or endophasia, can take various forms depending on the situation, cognitive demand or emotional context. It may also vary across cultures and individuals. People with auditory verbal aphantasia do not hear an inner voice when they produce inner language. Conversely, people with verbal hyperphantasia have an extremely vivid inner voice. On another dimension, although inner speech is usually identified as self-produced, some individuals have voice hearing experiences during which the voice feels externally controlled. These hallucinations may be considered as endophasia with a disrupted feeling of agency.

The diversity of endophasia can be accounted for by considering a gradual variation along three essential dimensions: condensation, dialogality and intentionality. Along thecondensation dimension, some extremely condensed forms are observed, that are abbreviated and deprived of sensory qualities. At the other end, there are expanded forms, with full syntactic, lexical and phonological specification, and with gestural, articulatory and auditory properties. Along dialogality, endophasia is sometimes monologal, when we engage in internal soliloquy. Other times, it is dialogal, when we recall past dialogues or imagine future conversations, and alternate between speaker and listener roles. It can also be dialogic when, as one speaker, we consider different perspectives. Along intentionality, endophasia can be intentional, when we deliberately rehearse material in memory for example. And it can feel unintentional and irruptive, during mind wandering. A related property is the feeling of agency, which may fluctuate, with endophasia not always identified as self-produced.

To account for varieties of endophasia along these three dimensions, we have introduced the *ConDialInt* Model, a neurocognitive sketch cast within a predictive control framework. In *ConDialInt*, the phenomenon of inner voice is seen as an exaptation of the sensory predictions occurring in the predictive control of overt speech. Speech production is considered as hierarchically controlled, from conceptualisation to articulation, via formulation and articulatory planning. Expanded forms of endophasia, associated with inner voice percepts, are viewed as predicted sensory signals deriving from internal simulation of desired speech goals. Condensed, non-sensory, forms of endophasia are construed as hierarchical precursors of expanded forms. Dialogal forms are viewed as verbal simulations accompanied with indexical and perspective properties. Intentional forms are seen as rigorously monitored predictions, whereas unintentional forms are loosely monitored simulations. Agency attribution is assumed to rely on the sequencing of predicted and desired signals. The *ConDialInt* model is compatible with neuroanatomical data obtained for a variety of inner speech situations. It also accounts for atypical endophasia. Verbal hyperphantasia vs aphantasia are explained by variation in condensation, with activation vs inhibition of lower-level sensory simulation. Verbal hallucinations are construed as unintentional forms of dialogal endophasia bestowed with a lack of agency. These propositions have implications for levels of representation in theories of speech control.

TEMPORAL DYNAMICS OF MOTOR SPEECH ENCODING

Marina Laganaro
University of Geneva, Geneva, Switzerland

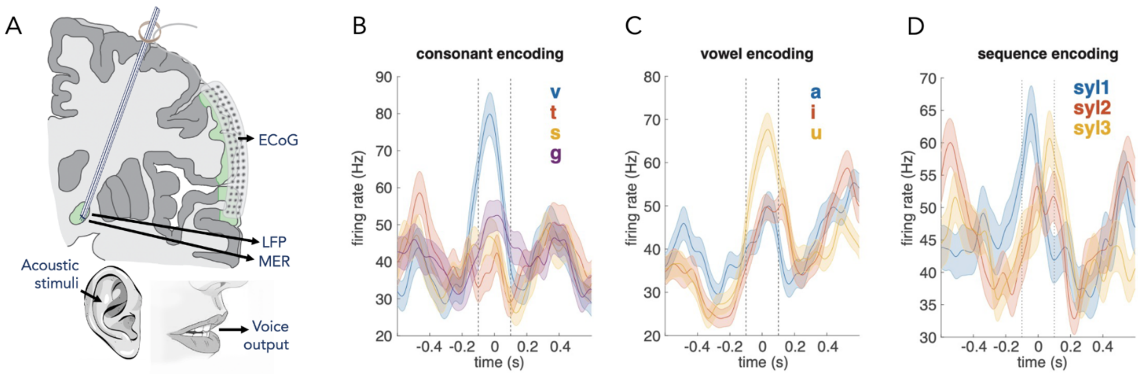
Whereas there is currently some insight into the brain structures involved in motor speech control, the temporal dynamics of motor speech encoding is much less investigated, probably because neuroimaging techniques allowing high temporal resolution (EEG and MEG) are very sensitive to articulatory movement. Here we will review existing evidence on the time-course of motor speech planning and programming and will show that motor speech encoding may require a larger proportion of time and of processes relative to linguistic encoding than suggested in current models on the dynamics of language and speech production.

CORTICAL-SUBTHALAMIC ACTIVITY IN SPEECH PRODUCTION

Mark Richardson[1], Witold Lipski[2], Matteo Vissani[1], Anna Chrabaszcz[2], Christina Dastolfo-Hromack[4], Alan Bush[1], Susan Shaman[2], Julie Fiez[2], Robert Turner[2]
[1] *Harvard Medical School, Boston, Massachusetts, United States of America*
[2] *University of Pittsburgh, Pittsburgh, Pennsylvania, United States of America*
[4] *West Virginia University, Morgantown, West Virginia, United States of America*

The production and modulation of speech, like other behaviors, relies on basal ganglia-thalamocortical networks, not solely on cortical processes. Current knowledge about how the basal ganglia modulate speech is primarily theoretical, due to a lack of empirical data. Notably, no models of speech production include the subthalamic nucleus (STN), a basal ganglia node that has been implicated in multiple cognitive processes relevant to speech production. Using access to the brain that can only be provided by deep brain stimulation (DBS) surgery, we established a novel experimental paradigm to simultaneously record STN activity and electrocorticography (ECoG) from speech-related cortical areas. We previously discovered separate populations of individual STN neurons whose inhibition or excitation was selective for either speech planning or production (Lipski et al., 2018). We also found that STN local field potential activity tracks with specific articulatory motor features (Chrabaszcz et al., 2019) and with vocal gain adjustment (Dastolfo-Hromack et al., 2021). Using STN stimulation and cortical evoked potential mapping, we also recently found evidence that the cortical-basal ganglia hyperdirect pathway extends to the STN from the inferior frontal gyrus and the superior temporal gyrus, in addition to the known connections from motor cortex (Jorge et al., 2022). These unique data suggest that the STN modulates cortical networks for speech planning, production, and perception. Here, we report two additional lines of study from single unit recordings in the STN during an intraoperative syllable triplet repetition task, in patients with Parkinson's disease undergoing DBS surgery.

First, 227 single units recorded in 25 subjects were analyzed using linear regression models that employed 7 categorical phoneme and sequence variables. 23% of neurons specifically encoded phoneme-related information and 14% encoded sequence-specific information. Encoding of phoneme and sequence features was independent of the direction of firing rate change, suggesting that multiple types of information are processed in parallel. Second, to assess STN-cortical interactions, we measured the phase consistency of neuronal spikes with narrowband-filtered ECoG fluctuations by computing the pairwise phase consistency index, which is unbiased by the number of spikes. STN neurons displayed significant pairwise phase consistency values with extended cortical regions, including sensory and premotor areas. Notably, significant suppression of beta-frequency range spike-phase synchrony was observed during both speech preparation and production. In sum, these data suggest that dynamic information flow between speech-related cortical areas and the STN contributes to multiple aspects of speech production.



STN neural firing encodes phoneme and sequence features of speech*. Schematic of experimental paradigm (A). Examples of mean firing rate raster plots indicating consonant (B), vowel (C) and sequence (D) encoding. For each feature category, the spike density function is averaged by feature, across all syllables in a task session. Data are aligned to the consonant-vowel transition.*

References

Chrabaszcz, A. et al. (2019). Subthalamic Nucleus and Sensorimotor Cortex Activity During Speech Production. *J Neurosci 39*, 2698–2708.

Dastolfo-Hromack, C. et al. (2021). Articulatory Gain Predicts Motor Cortex and Subthalamic Nucleus Activity During Speech. *Cereb Cortex*. `https://doi.org/10.1093/cercor/bhab251`

Jorge, A. et al. (2022) Hyperdirect connectivity of opercular speech network to the subthalamic nucleus. *Cell Reports 38*, 110477.

Lipski, W. J. et al. (2018). Subthalamic Nucleus Neurons Differentially Encode Early and Late Aspects of Speech Production. *J Neurosci 38*, 5620–5631.

DECODING OF SPEECH IMAGERY AS A WINDOW TO SPEECH PLANNING AND PRODUCTION

Joan Orpella[1], Francesco Mantegna[1], Florencia Assaneo[2], David Poeppel[1,3]
[1]*New York University, New York, United States of America*
[2]*Universidad Autónoma de México, Ciudad de México, Mexico*
[3]*Ernst Strüngman Institute, Frankfurt am Main, Germany*

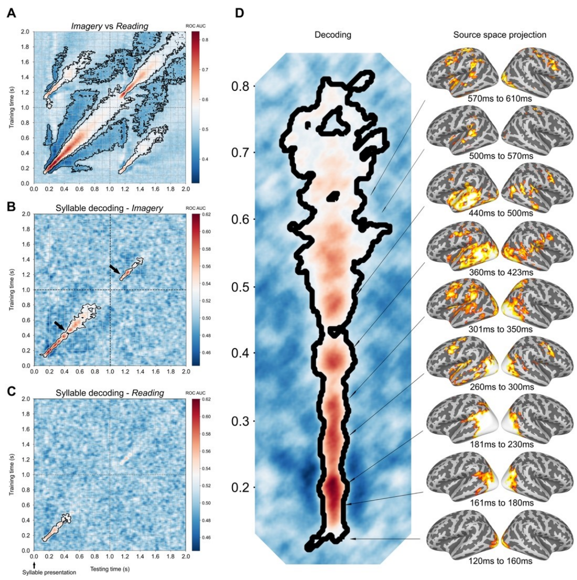
Introduction Speech imagery (the ability to generate internally quasi-perceptual experiences of our own or others' speech) is widely recognized as a fundamental ability tightly linked to important cognitive functions such as inner speech, phonological working memory, and predictive processing. Speech imagery is also considered an ideal medium to test theories of overt speech. Despite its pervasive nature, the study and use of speech imagery for clinical or basic research has been tremendously challenging. Particularly, the lack of direct observable behavior and difficulties in aligning imagery events across trials and individuals have prevented a thorough understanding of the underlying neural dynamics and limited its use as a research tool. In this work, we aimed to map out the generation of speech imagery by pairing magnetoencephalography (MEG) with a novel experimental protocol designed to overcome these difficulties.

Method Thirty participants (22 women; mean age = 26, std = 7) imagined producing isolated syllables (e.g., pa, ta, ka) immediately after these were presented on a screen and a second time 1000ms later while we recorded their neural activity with MEG (157-channel whole-head axial gradiometer). This Imagery condition was contrasted with a Reading condition, in which participants simply read the syllables and were asked not to imagine them. Each condition comprised 4 blocks of 120 trials (40 randomized presentations of each syllable). We recorded electromyographic (EMG) data from participants' upper lip and jaw to monitor their micromovements. We also acquired structural magnetic resonance imaging data (T1) from a subset of 10 participants to source project their speech imagery MEG data. Participants were trained on an overt version of the task prior to the MEG session. Their overt productions were recorded to estimate timings for imagery and syllable/consonant-vowel transition durations.

Data analysis We used a decoding approach (King & Dehaene, 2014) to 1. classify participants' MEG data as Imagery or Reading, 2. classify the imagined syllables, 3. explore different levels of representation (syllable, consonant-vowel transition) during imagery, and 4. ensure that our results could not be explained by participants' micromovements. Participants' MEG data was also projected to source space to investigate the temporal dynamics of speech imagery. Results: Robust classification scores were obtained for the contrast between Imagery and Reading and between the syllables (Fig 1a). Syllable decoding revealed a rapid sequence of relatively encapsulated representations from visual encoding to the imagined speech event (Fig 1b). Participants' micromovements did not discriminate between the syllables.

The neural correlates of the decoded sequence of representations (Fig 1d) neatly maps onto the predictions of current models of speech production (e.g., State Feedback Control; Hickok et al., 2011; Hickok, 2012) providing some evidence for hypothesized internal and external feedback loops for speech planning and production, respectively. Additionally, a novel decoding approach (Windowed Multinomial Classification) revealed the presence of two nested and concurrent levels of representation (syllable and consonant-vowel transition) and exposed the compressed nature of representations during planning which contrasts with the natural rate at which syllables are internally produced.

Conclusions The results show an evolving sequence of representations for speech imagery with neural dynamics and characteristics consistent with SFC. It is assumed that the same sequence underlies the generation of sensory predictions (Keller & Mrsic-Flogel, 2018) through the speech motor system that modulates speech perception and subserves the articulatory loop of phonological working memory (Hickok et al., 2011). The results highlight the potential of speech imagery for research based on these new experimental approaches and analytical methods, and further paves the way for successful non-invasive brain-computer interface applications.



Temporal generalization matrices track the development of neural representations during speech Imagery and Reading. A. Average temporal generalization (TG) matrix for the contrast Imagery vs Reading. *Participants imagined producing one of three syllables immediately after its visual presentation (time = 0) and a second time at time = 1. ROC AUC = Receiver Operative Curve Area Under the Curve (chance = 0.5).* ***B. and C. TG matrices for the pairwise contrasts between syllables for each condition*** *(pa vs. ka, pa vs. ta, and ta vs. ka) (Imagery and Reading, respectively). TG matrices were first averaged within subject and then across subjects. Black arrows indicate the median syllable onset time of participants' overt productions during training (event 1 median: 436ms; event 2 median: 1175ms). Clusters of statistically significant decoding ($p < 0.05$; black contour lines) were in all cases determined at the second level of analysis via a cluster-based permutation test across subjects (1000 permutations; two-tailed). Statistical significance indicates consistence across subjects, while high ROC AUC values reflect robust classifier performance on discriminating the contrasts.* ***D. Clusters of syllable decoding reveal a processing cascade during speech imagery.*** *Evoked activity for a subgroup of participants estimated with sLORETA corresponding to the clusters of significant syllable decoding in the Imagery condition (Fig 1B). Source space activity was thresholded at minima ranging between 1.88 and 3.34 and maxima between 2.37 and 4.51 units for display purposes.*

References

King, J. R. & Dehaene, S. (2014). Characterizing the dynamics of mental representations: The temporal generalization method. *Trends Cogn. Sci. 18*, 203–210.

Hickok, G., Houde, J. & Rong, F. (2011). Sensorimotor Integration in Speech Processing: Computational Basis and Neural Organization. *Neuron 68*, 1–28.

Hickok, G. (2012). Computational neuroanatomy of speech production. *Nat. Rev. Neurosci. 13.*

Keller, G. B. & Mrsic-Flogel, T. D. (2018). Predictive Processing: A Canonical Cortical Computation. *Neuron 100*, 424–435.

THE ENCODING OF SPEECH MOVEMENTS AND PLANNING IN THE HUMAN MOTOR CORTEX

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

We will review recent advances in understanding the neural encoding of speech movement parameters and the planning of speech sequences.

MOTOR-TO-SENSORY INFLUENCES DURING SPEECH MOVEMENT PLANNING IN INDIVIDUALS WHO STUTTER

Ludo Max
University of Washington, Seattle, Washington, United States of America

Multiple lines of research regarding the neurological basis of stuttering suggest that the characteristic breakdowns in speech fluency can be attributed to fundamental difficulties with sensorimotor integration. Although neuroimaging studies have identified various structural and functional differences throughout the brains of individuals who stutter, it remains critical to identify the basic mechanisms that give rise to the actual stuttered speech dysfluencies. Surprisingly, work in this area continues to be driven mostly by isolated hypotheses rather than theoretical perspectives with a solid foundation in the neuroscience of sensorimotor control. This presentation will provide an overview of a series of studies designed based on a theoretically-motivated neurobiological perspective on stuttering. The approach is based on empirical and computational work indicating that, for the control of fast voluntary movements, the central nervous system predicts future sensory states by means of adaptive internal forward models. Such predictions are believed to underlie the movement-related modulation of afferent signals in various sensory modalities, including the auditory system while speaking. In our basic paradigm, we investigate this modulation during speech movement planning by recording auditory evoked potentials in response to probe tones presented immediately prior to speaking or at the equivalent time in no-speaking control conditions.

As a measure of pre-speech auditory modulation (PSAM), we calculate changes in the auditory evoked potential amplitude for the speaking condition relative to the no-speaking conditions. Our lab initially discovered that typical speakers already modulate auditory processing mechanisms prior to the initiation of speech movements whereas this phenomenon is lacking in adults who stutter. Across studies, the between-group difference in PSAM is consistently associated with large effect sizes, and sometimes even differentiates perfectly between stuttering and nonstuttering speakers. In subsequent studies, we also found that PSAM is not correlated with speech auditory–motor adaptation in typical speakers but negatively correlated in stuttering speakers.

As one of the most intriguing findings, the PSAM difference between stuttering and nonstuttering adults disappeared when completing the same speaking task with delayed auditory feedback. In our most recent studies, we have focused on the functional relevance of PSAM for speech production by studying (a) its perceptual consequences for the speaker's own auditory monitoring ability, (b) its relation to the speaker's inter-trial production variability, and (c) its characteristics during the planning of sentence-level rather than word-level utterances. Data from the entire series of studies will be integrated to develop an interim-hypothesis regarding the role of PSAM in speech production in general and the potentially negative impact of its absence in individuals who stutter.

Acknowledgment Experiments conducted by Yusuf Ali, Ayoub Daliri, Kwang S. Kim, Elise LeBovidge, and Hantao Wang. Funded by NIH/NIDCD grants R01DC017444 and R01DC014510.

References

Daliri, A., & Max, L. (2018). Stuttering adults' lack of pre-speech auditory modulation normalizes when speaking with delayed auditory feedback. *Cortex, 99*, 55–68.

Max, L., & Daliri, A. (2019). Limited pre-speech auditory modulation in individuals who stutter: data and hypotheses. *Journal of Speech, Language, and Hearing Research, 62*, 3071–3084.

ALTERED AUDITORY FEEDBACK AS A POWERFUL TOOL TO EXPLORE SPEECH ACTION-PERCEPTION RELATIONSHIPS: A (NON-EXHAUSTIVE) REVIEW OF EXPERIMENTS USING FORMANTS PERTURBATION

Tiphaine Caudrelier
GIPSA-lab, Grenoble, France
CNRS, Grenoble, France
University of Grenoble, Grenoble, France

Action-perception relationship is a key topic in speech production models (as well as in speech perception theories). What could be a better way to experimentally investigate the perception-action relationship in speech, than altering perception and observing the impact on speech production (or vice versa)? This is exactly what altered auditory feedback studies do, by altering the auditory feedback of speakers in real time while they produce speech: isolated words, pseudo-words or entire sentences (Caudrelier & Rochet-Capellan, 2019; Houde & Jordan, 1998; Lametti, Smith, Watkins, & Shiller, 2018).

In this talk, I will focus on a particular type of feedback alteration: formants shift, which enables to transform vowels into other vowels. Given this ability to manipulate phonological representations, this type of alteration has the potential to address questions in both speech motor control and psycholinguistics and to help bridging the gap between these fields. This could lead to a better overall understanding of speech production.

In the first part of the talk, I will give an overview of a few studies which shed light on the perception-action relationship in speech, the contribution of different sensory modalities (auditory, somatosensory, visual) and their integration. These studies give some support to the idea that speech production units are sensorimotor representations associating abstract units with motor representations and multiple sensory goals. Thus arises another important question which I will focus on in the second part of the talk. This question has been traditionally more debated in the field of psycholinguistics but is also key for speech motor control models: what are the main speech production units? More specifically, what is the grain-size of the above introduced sensorimotor units? Again I will present a few formant perturbation studies that have addressed these questions (e.g. Caudrelier et al., 2019; Caudrelier, Schwartz, Perrier, Gerber, & Rochet-Capellan, 2018). Finally I will briefly review other topics that have greatly benefited from formant perturbation studies. The implications for speech production models will be discussed, as well as future perspectives in the field of formants perturbation.

References

Caudrelier, T., Ménard, L., Perrier, P., Schwartz, J., Gerber, S., Vidou, C., & Rochet-Capellan, A. (2019). Transfer of sensorimotor learning reveals phoneme representations in preliterate children. *Cognition, 192*(11), 103973. https://doi.org/10.1016/j.cognition.2019.05.010

Caudrelier, T., & Rochet-Capellan, A. (2019). Changes in speech production in response to formant perturbations: An overview of two decades of research. In S. Fuchs, J. Clelland, & A. Rochet-Capellan (Eds.), *Speech production and perception: Learning and memory* (pp. 15–75). Peter Lang Publisher.

Caudrelier, T., Schwartz, J., Perrier, P., Gerber, S., & Rochet-Capellan, A. (2018). Transfer of learning: What does it tell us about speech production units? *Journal of Speech Language and Hearing Research, 61*(7), 1613–1625.

Houde, J. F., & Jordan, M. I. (1998). Sensorimotor adaptation in speech production. *Science, 279*(5354), 1213–1216. https://doi.org/10.1126/science.279.5354.1213

Lametti, D. R., Smith, H. J., Watkins, K. E., & Shiller, D. M. (2018). Robust sensorimotor learning during variable sentence-level speech. *Current Biology, 28*(19), 3106–3113. https://doi.org/10.1016/j.cub.2018.07.030

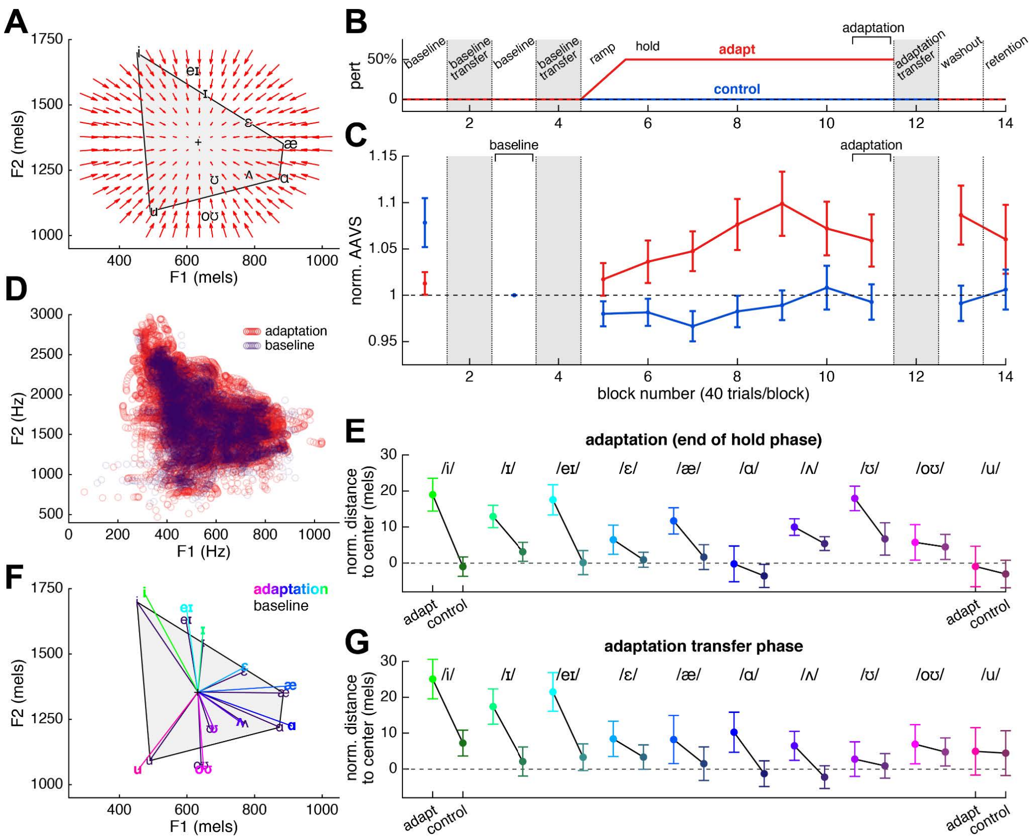
INCREASED VOWEL CONTRAST IN CONNECTED SPEECH INDUCED BY SENSORIMOTOR ADAPTATION

Caroline A. Niziolek*, Benjamin Parrell*, Sophie A. Johnson
University of Wisconsin-Madison, Madison, United States of America
**equal contribution*

Alterations to sensory feedback can drive robust adaptive changes to the production of consonants and vowels, but these changes often have no behavioral relevance or benefit to communication effectiveness (e.g., causing "head" to be more similar to "had"). This work aims to align the outcomes of adaptation with changes known to increase speech intelligibility. Specifically, we target adaptations that would increase speakers' working vowel space area and increase the acoustic contrast between vowels.

To this end, we implemented a vowel centralization feedback perturbation paradigm that pushes all vowels towards the center of the vowel space (Fig. 1A), making them sound less distinct from one another. Recent work has shown that speakers can successfully adapt to this vowel centralization during the production of isolated monosyllabic words, producing corner vowels farther from the center of the vowel space and increasing global measures of vowel contrast (Parrell & Niziolek, 2021). In the current study, we investigated whether speakers can adapt to vowel centralization in connected speech, which has been shown to support learning a constant formant shift applied globally across the entire vowel space (Lametti et al., 2018). Adult speakers of American English (n=32; 39 planned, ages 18-71) read a subset of 40 phonetically-balanced sentences covering the English vowel space (IEEE Recommended Practice for Speech Quality Measurements, 1969). After a baseline phase with normal feedback, the vowel centralization perturbation was introduced in a ramp phase and held at maximum magnitude (50% of the distance to the vowel space center) for 6 blocks (Fig 1B). Feedback was returned to normal in a subsequent washout phase as well as in a separate retention phase that occurred after a 10-minute delay. Additionally, noise-masked transfer blocks before the ramp and after the hold phase assessed changes in the production of isolated monosyllabic words (Lametti et al., 2018). To control for potential vowel changes over the course of the experiment, speakers also completed a control session, identical in structure to the main adapt session and counterbalanced in order, in which no auditory perturbations were applied.

Speakers adapted to the centralization perturbation during sentence production. A global measure of the working vowel space for connected speech, the articulatory-acoustic vowel space (AAVS; Whitfield & Goberman), increased over the course of the adaptation session (Fig. 1C,D; $p = 0.04$). AAVS did not change in the control session ($p = 0.71$), which had a lower AAVS than the adapt session (one-tailed t-test, $p = 0.03$). Individual vowels, particularly high front vowels, were produced farther from the center of the vowel space (Fig. 1E,F) at the end of the hold phase. These changes partially persisted after the feedback shift was removed, including after a 10-minute silent period. Additionally, this learning showed a robust transfer to isolated monosyllabic words that were not present in the sentence stimuli (Fig. 1G). These findings establish the validity of a sensorimotor adaptation paradigm to increase vowel contrast in connected speech, an outcome that has the potential to enhance intelligibility in individuals who present with a reduced vowel space due to motor speech disorders, such as the hypokinetic dysarthria associated with Parkinson's disease.



A: Perturbation field applied to speech. All perturbations point towards the center of each speaker's vowel space. B: Magnitude of the perturbation applied throughout the experiment. In the adapt session (red), the perturbation during the hold phase is 50% of the 2D distance (in F1/F2 space) between the current formant values and the vowel center. In a separate control session (blue), no perturbation is applied. C: Articulatory-acoustic vowel space (AAVS) for all participants over the course of the experiment. D: Formant trajectories from an example participant during sentence reading. AAVS is increased after adaptation (red dots). E,G: Increases in individual vowels' distance to the center of the vowel space during sentence production in the hold phase (E) and single word production in the transfer phase (G). F: Average formants for an example participant at baseline (black) and transfer (bright colors), illustrating increased distance from the center of the vowel space.

References

IEEE Recommended Practice for Speech Quality Measurements (1969). *IEEE Transactions on Audio and Electroacoustics, 17*, 227–246.

Lametti, D. R., Smith, H. J., Watkins, K. E., & Shiller, D. M. (2018). Robust Sensorimotor Learning during Variable Sentence-Level Speech. *Current Biology, 28*(19), 3106-3113.e2. https://doi.org/10.1016/j.cub.2018.07.030

Parrell, B., & Niziolek, C. A. (2021). Increased speech contrast induced by sensorimotor adaptation to a nonuniform auditory perturbation. *Journal of Neurophysiology, 125*(2), 638–647. https://doi.org/10.1152/jn.00466.2020

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. https://doi.org/10.1016/j.jcomdis.2014.06.005

SPEECH MOTOR ADAPTATION DURING SYNCHRONOUS SPEECH

Abigail R. Bradshaw[1,2], Daniel R. Lametti[3], Douglas M. Shiller[4], Kyle Jasmin[5], Carolyn McGettigan[1]

[1]*University College London, London, United Kingdom*

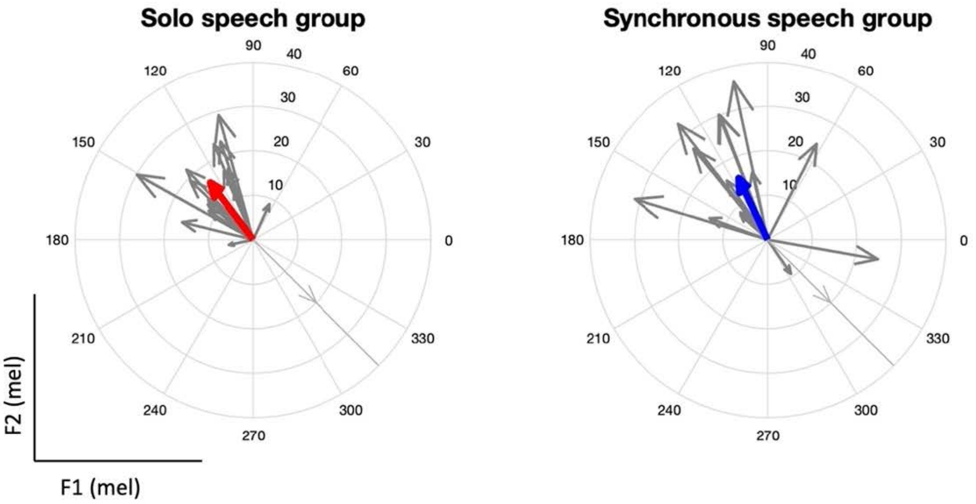
[2]*University of Cambridge, Cambridge, United Kingdom*

[3]*Acadia University, Wolfville, Canada*

[4]*University of Montréal, Montréal, Canada*

[5]*Royal Holloway University of London, London, United Kingdom*

Research with the altered auditory feedback paradigm has provided evidence for the critical role played by processing of self-generated speech auditory feedback in speech motor control. Specifically, speakers exposed to a predictable sustained perturbation of real-time auditory feedback (e.g. a change in the first or second formant) gradually start to adapt to this perturbation; for example, by shifting their produced formant frequencies in an opposite direction to the perturbation. Less is known however about the impact of other voices on such speech motor adaptation, and whether adaptation is robust across different speaking contexts. In particular, despite speech typically being a social act, few previous studies have examined speech adaptation in contexts involving a social element. In this study, we tested the effect of synchronous speech (the act of speaking in synchrony with another voice) on the adaptation response. Using the sentence-level adaptation paradigm developed by Lametti, Smith, Watkins, and Shiller (2018), we measured participant's adaptation to a joint F1-F2 perturbation during production of sentences, either while speaking alone (solo reading group) or while synchronising their speech with another voice ("the accompanist"; synchronous speech group). We found that both groups exhibited a significant adaptation response, with no significant difference between the magnitude of the average group responses (see Figure 1). There was however a significant difference in the level of between-participant variability in adaptation within the two groups, with more variable adaptation responses in the synchronous speech group. It was further found that participants in the synchronous speech group showed evidence of convergent changes in the F1, F2 and F0 of their speech productions towards those of the accompanist voice prior to introduction of the feedback perturbation, replicating previous work (Bradshaw & McGettigan, 2021). Individual variability in the magnitude of this initial convergence however did not correlate with the magnitude of adaptation to the subsequent perturbation. These findings demonstrate a novel speaking context within which speech motor adaptation is observed, suggesting that it can be robust to the effects of speech task and speech input from another speaker. The increased variability in the adaptation response however suggests that there may be speaker-specific effects of synchronous speech on the adaptation response. To better understand this variability, further studies are planned to test for the effects of the acoustic-phonetic properties of the accompanist voice on adaptation, and the effects of metronome-timed speech.



Adaptation responses in solo speech and synchronous speech groups. Grey arrows indicate individual adaptation responses, in the form of directions in F1/F2 space from a baseline block to the final block of adaptation. Group average responses are shown in red and blue for solo and synchronous speech groups respectively. The pale grey arrow at 315 degrees indicates the angle of the F1 up, F2 down perturbation applied to speech feedback.

References

Bradshaw, A. R., & McGettigan, C. (2021). Convergence in voice fundamental frequency during synchronous speech. *PLOS ONE, 16*(10), e0258747. `https://doi.org/10.1371/journal.pone.0258747`

Lametti, D. R., Smith, H. J., Watkins, K. E., & Shiller, D. M. (2018). Robust Sensorimotor Learning during Variable Sentence-Level Speech. *Current Biology, 28*(19), 3106-3113.e2.`https://doi.org/10.1016/j.cub.2018.07.030`

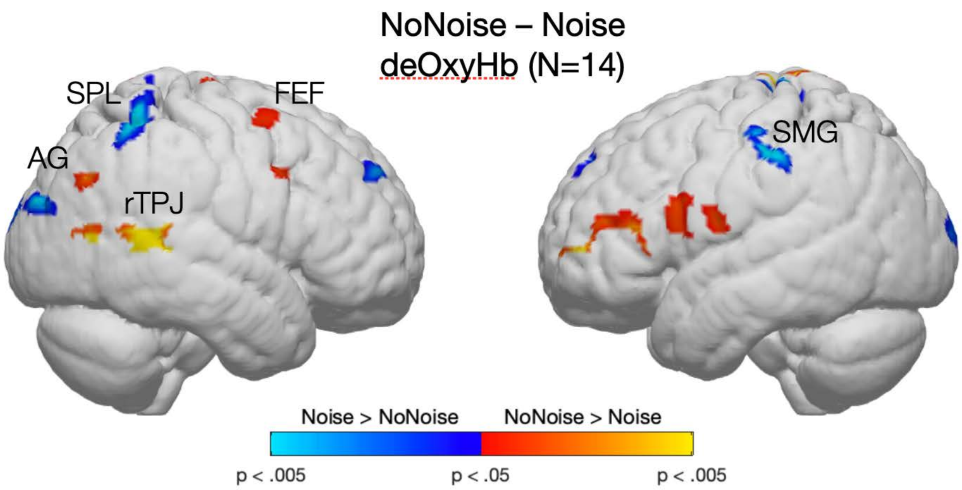
NEURAL AND BEHAVIORAL RESPONSES TO TALKING FACES IN COCKTAIL PARTY NOISE

Mark Tiede[1,2], Juliana Brenner[1], Adam Noah[1], Xian Zhang[1], Joy Hirsch[1,2,3]
[1]*Yale University, New Haven, United States of America*
[2]*Haskins Laboratories, New Haven, United States of America*
[3]*University College London, London, United Kingdom*

This work extends an earlier study of face-viewing eye contact behavior under noise (Vatikiotis-Bateson et al., 1998) through simultaneous observation of eye-tracking with functional near-infrared spectroscopy (fNIRS). The prior study (of two participants) showed that as noise increased the proportion of gaze directed at the mouth also increased, while transitions between the mouth and eyes decreased. Separately, previous work (Hirsch et al., 2017; Dravida et al., 2020; Noah et al., 2020) has identified neural correlates for eye contact cues associated with dyadic interaction. Eye contact is fundamental to social communication, especially for assessing pragmatic aspects of discourse ("is this person being sincere?"); however, under noisy conditions more resources must be allocated to tracking mouth movement for visual cues to help disambiguate acoustically confusable phonemes. This motivates two predictions for this ongoing joint eye-tracking / fNIRS study of face viewing under Noise relative to a NoNoise baseline: that without noise cortical areas associated with social interaction and eye contact will show more activity, while under noise areas associated with language processing and cue integration will show more activity.

To date 14 participants (7F, mean age 24.1) watched and listened to video clips of six speakers talking extemporaneously about various topics (e.g. "Small towns are better than big cities"). Participant eye movements (Tobii x3-120) and cortical hemodynamic response functions (Shimadzu LABNIRS; 134 acquisition channels) were recorded during six trials. Within each trial participants viewed six 15s video clips from the same speaker alternating with 15s rest intervals. To assess engagement after each trial participants were asked two Yes/No questions; the first addressed speaker sincerity, and the second topic comprehension. For half the trials in counterbalanced randomized order cocktail party noise (Fischer et al., 2020) was played over loudspeakers at 75dB. Hemodynamic response functions were subjected to spatial filtering (Zhang et al., 2016) and general linear modeling, then mapped to a topographic representation of the MNI cortex. Locations of speaker eyes and mouth were tracked using OpenFace (Baltrušaitis et al., 2018) and aligned with participant gaze patterns for eye/mouth target categorization.

Behavioral results confirm the pattern observed in (Vatikiotis-Bateson et al., 1998), showing a highly significant shift in fixations from the eyes to the mouth under noise. The neural effects were assessed by contrasting the deOxyHb signals between the NoNoise vs. Noise conditions and are illustrated in Figure 1. Under NoNoise as predicted these show greater activity in the right Temporoparietal Junction (associated with processing higher order social cues; e.g. theory of mind), as well as the right Angular Gyrus and Frontal Eye Fields (associated with face and eye processing); while under Noise greater activity is observed in the right Superior Parietal Lobule and the left Supramarginal Gyrus (associated with visual processing, salience detection and cue integration). We interpret these preliminary results as support for the hypothesis that under noise resource reallocation occurs, deprioritizing social cue evaluation. From an action-perception perspective these results suggest implications for production as well, in that hyperarticulation and Lombard effects known to occur under noise (Garnier et al., 2018) may come at the expense of coproducing facial social cues.



fNIRS results contrasting NoNoise vs. Noise conditions: red areas show greater activity under NoNoise; blue areas greater activity under noise.

References

Baltrušaitis, T., Zadeh, A., Lim, Y.C. & Morency, L.-P. (2018). Openface 2.0: Facial behavior analysis toolkit. 2018 *13th IEEE international conference on automatic face & gesture recognition* (FG 2018), IEEE59-66.

Dravida, S., Noah, J.A., Zhang, X. & Hirsch, J. (2020). Joint attention during live person-toperson contact activates rTPJ, including a sub-component associated with spontaneous eye-to-eye contact. *Frontiers in Human Neuroscience, 14*, 201.

Fischer, T., Caversaccio, M. & Wimmer, W. (2020). Multichannel acoustic source and image dataset for the cocktail party effect in hearing aid and implant users. *Scientific Data, 7*(1), 1-13.

Garnier, M., Ménard, L. & Alexandre, B. (2018). Hyper-articulation in Lombard speech: An active communicative strategy to enhance visible speech cues? *The Journal of the Acoustical Society of America, 144*(2), 1059-1074.

Hirsch, J., Zhang, X., Noah, J.A. & Ono, Y. (2017). Frontal temporal and parietal systems synchronize within and across brains during live eye-to-eye contact. *Neuroimage, 157*, 314-330.

Noah, J.A., Zhang, X., Dravida, S., Ono, Y., Naples, A., McPartland, J.C. & Hirsch, J. (2020). Real-time eye-to-eye contact is associated with cross-brain neural coupling in angular gyrus. *Frontiers in Human Neuroscience, 14*, 19.

Vatikiotis-Bateson, E., Eigsti, I.-M., Yano, S. & Munhall, K.G. (1998). Eye movement of perceivers during audiovisualspeech perception. *Perception & Psychophysics, 60*(6), 926-940.

Zhang, X., Noah, J.A. & Hirsch, J. (2016). Separation of the global and local components in functional near-infrared spectroscopy signals using principal component spatial filtering. *Neurophotonics, 3*(1), 015004.

REAL-TIME MAGNETIC RESONANCE IMAGING FOR PHONETIC RESEARCH: CURRENT STUDIES

Philip Hoole
Ludwig-Maximilians-University, Munich, Germany

The presentation will first briefly review developments in real-time MRI, including likely future developments and the potential for new kinds of experiments, and then focus on two case studies that illustrate the use of real-time MRI to analyse the articulatory manoeuvres of speech organs that have hitherto generally been restricted to studies with very few speakers, namely velum movement and vertical larynx movement. The analyses are taken from a large corpus of German data for over 30 speakers recorded at 50fps with a pixel size of 1.4 x 1.4mm. In addition, we will emphasize the flexibility of RT-MRI to extend consideration to additional articulators where required.

The overall aim of the analysis of velum kinematics is to improve our understanding of the phonetic forces that can lead diachronically to contrastive vowel nasalization. Specifically, we will look firstly at quite subtle temporal phenomena, namely differences in anticipatory nasal coarticulation related to voicing of the post-nasal consonant in VNC sequences. Secondly, we will consider spatial effects related to the influence of different vowel categories on the amount of velum opening in both nasal and non-nasal consonantal contexts, again showing the potential of the technique for revealing subtle articulatory differences, that in this particular case are readily explainable in terms of the balance of forces between levator palatini (for raising the velum) and downward tug from the connection to the tongue via the palatoglossus muscle.

The second study will consider larynx height in vowel production. We exploit the properties of the German vowel system in an attempt to give a more balanced picture of whether larynx height is more closely related to vowel height or rounding. Previous results in the literature are quite messy (particularly for vowel height) which in turn is related to the pervasive, but for larynx height particularly pronounced problem, of inter-speaker variability. We look briefly at whether the amount of vowel-specific modulation of larynx height can be related to speaker anatomy and speaker preferences for lip protrusion.

Finally, the increasing utility of real-time MRI for investigating connected speech processes can be illustrated by the fact that the procedures developed for investigating larynx height in German turned out to useful for analysing an ongoing sound-change involving ejectives that were richly – albeit unintentionally – represented in a parallel English corpus.

References

Carignan, C., Coretta, S., Frahm, J., Harrington, J., Hoole, P., Joseph, A., Kunay, E., & Voit, D. (2021). Planting the Seed for Sound Change: Evidence from Real-Time MRI of Velum Kinematics in German. *Language 97*(2), 333-364

Hoole, P., Coretta, S., Carignan, C., Kunay, E., Joseph, A., Voit, D., & Frahm, J. (2020). Control of larynx height in vowel production revisited: A real-time MRI study. *International Seminar on Speech Production*, `https://issp2020.yale.edu/S08/hoole_08_03_033_poster.pdf`

Kunay, E. (2021): *Vowel nasalization in German: a real-time MRI study*. Dissertation, LMU München: Faculty for Languages and Literatures. `https://doi.org/10.5282/edoc.29340`

SPEECH RESPIRATION: PAST RESEARCH, RECENT TRENDS, AND FUTURE DEVELOPMENTS

Susanne Fuchs
Leibniz-Zentrum Allgemeine Sprachwissenschaft, Berlin, Germany

In this review, I will briefly summarize past and present research on speech respiration. Speech production and breathing have an interwoven relationship that depends on communicative and physical constraints (Fuchs & Rochet-Capellan, 2021). Specifically, breathing is highly flexible and adaptable, since breathing pauses occur frequently at meaningful pauses, it adapts to cognitive load in perception and to communicative needs. It is, however, also stable within a given speaker within certain conditions, across days. Speech production may also adapt to respiration, i.e., the larynx can compensate for the loss of air. To better understand this interwoven relationship, the following interdisciplinary research efforts are needed:

1. integration of fast and slow rhythms in speech motor control,
2. investigations of respiration in coordination with cognitive and physical activities,
3. integrating ontogenetic and phylogenetic perspectives on breathing,
4. developing models of speech production that include respiration.

References

Fuchs, S., & Rochet-Capellan, A. (2021). The respiratory foundations of spoken language. *Annual Review of Linguistics, 7*(1), 13-30.

MEASURING GROWTH AND PREDICTING SPEECH OUTCOMES IN CHILDREN WITH AND WITHOUT DYSARTHRIA

Katherine Hustad
University of Wisconsin-Madison, Madison, Wisconsin, United States of America

This presentation will report findings from an ongoing longitudinal study on speech and language acquisition in children with cerebral palsy over 10 years of development and a related cross-sectional study on normative development of speech intelligibility in typical children. Results will highlight growth curves for both typical and dysarthric speech development as well as normative benchmarks for typical speech intelligibility development. Statistical models from early speech production data and the outcomes that they predict at 8 years of age in children with CP will also be discussed.

SPEECH PRODUCTION ERRORS IN AUSTRALIAN ENGLISH-DUTCH BILINGUAL CHILDREN

Hayo Terband[1,2], Ymke Rankenberg[2], Anniek van Doornik[2,3]
[1]*University of Iowa, Iowa City, Iowa, United States of America*
[2]*Utrecht University, Utrecht, The Netherlands*
[3]*HU University of Applied Sciences, Utrecht, The Netherlands*

Introduction Determining whether suspected speech 'abnormalities' in bilingual children are due to a pattern of bilingual language acquisition - and language dominance - or due to a speech sound disorder is a challenging task for speech-language pathologists (SLPs; Girolametto & Cleave, 2010; McLeod, Verdon, Bowen, & Speech, 2013). With the advantage of not directly involving linguistic knowledge, nonword repetition (NWR) is often used as a diagnostic task, in particular with respect to phonological and language disorders, (e.g., Boerma, et al., 2015; Dos Santos & Ferré, 2018; Ortiz, 2021; Schwob, et al., 2021). However, previous studies have shown that typically developing bilingual children can score significantly worse on these tasks compared to typically developing monolingual children (Engel de Abreu, 2011; Thordardottir & Brandeker, 2013; Windsor, Kohnert, Lobitz, & Pham, 2010). Furthermore, most studies reported only general outcome measures such as percentage of items or phonemes correct (Ortiz, 2021; Schwob, et al., 2021) and only few studies featured a detailed analysis of speech errors (i.e., Chiat & Roy, 2007; Kapalková, Polišenská & Vicenová, 2013). The aim of the present study was to investigate how NWR productions of English-Dutch bilingual children differ from norm data, how they relate to other speech tasks as to establish the potential role of NWR in diagnosing speech sound disorders in bilingual children.

Method & materials 64 Australian English-Dutch billingual children between 4 and 12 years of age (M = 7.96, SD = 2.33; 29 boys, 35 girls) from the Dutch school in Sydney participated in this study. All children completed the Dutch test battery *Computer Articulation Instrument* (CAI; Maassen et al., 2019), which includes picture naming, nonword repetition, consistency of word and nonword repetition and diadochokinesis. Data on language exposure was collected through parent/caregiver questionnaires. A group-level quantitative phonological error analysis compared performance across tasks and relative to norm data for Dutch. Possible associations between tasks were examined through a correlational analysis, both with z-scores and with the raw data.

Results & Discussion The English-Dutch bilingual children scored lower compared to the norm data on picture naming and consistency of word and nonword repetition. The phonological processes voicing and fronting were predominant. While the children showed a small delay on the other speech tasks, the scores on NWR were age appropriate. These results confirm NWR as (most) language-neutral assessment of speech production in bilingual children.

The correlation analysis revealed a negative correlation between NWR and consistency of repeating nonwords, while no other significant correlations were observed. The results on the consistency tasks (5 consecutive repetitions of words and nonwords) showed an interesting pattern of increased transfer of English features with each subsequent repetition. The memory trace of the acoustic model appears to fade with each repetition and the task thus slowly becomes a delayed imitation task, but more so for children with good (initial) imitation skills.

Conclusion These results confirm NWR as (most) language-neutral assessment of speech production in bilingual children. Voicing and fronting need attention of SLPs assessing English-Dutch bilingual children.

References

Boerma, T., Chiat, S., Leseman, P., Timmermeister, M., Wijnen, F., & Blom, E. (2015). A quasi-universal nonword repetition task as a diagnostic tool for bilingual children learning Dutch as a second language. *Journal of Speech, Language, and Hearing Research, 58*(6), 1747-1760.

Chiat, S. (2015). Nonword repetition. In S. Armon-Lotem, J. de Jong, & N. Meir (Eds.), *Methods for assessing multilingual children: Disentangling bilingualism from language impairment* (pp. 125–150). Bristol, United Kingdom: Multilingual Matters.

Chiat, S., & Roy, P. (2007). The preschool repetition test: An evaluation of performance in typically developing and clinically referred children. *Journal of Speech, Language, and Hearing Research, 50*(2), 429-443.

Dos Santos, C., & Ferré, S. (2018). A nonword repetition task to assess bilingual children's phonology. *Language Acquisition, 25*(1), 58-71.

Engel de Abreu, P. M. (2011). Working memory in multilingual children: Is there a bilingual effect? *Memory, 19*(5), 529-537.

Kapalková, S., Polišenská, K., & Vicenová, Z. (2013). Non-word repetition performance in Slovak-speaking children with and without SLI: novel scoring methods. *International Journal of Language & Communication Disorders, 48*(1), 78-89.

Maassen, B., van Haaften, L., Diepeveen, S., van den Engel-Hoek, L., Veenker, T., Terband, H., & De Swart, B. (2019). *Computer Articulatie-Instrument (CAI).* Amsterdam: Boom test uitgevers.

McLeod, S., Verdon, S., Baker, E., Ball, M. J., Ballard, E., David, A. B., ... & Zharkova, N. (2017). Tutorial: Speech assessment for multilingual children who do not speak the same language (s) as the speech-language pathologist. American *Journal of Speech-Language Pathology, 26*(3), 691-708.

Ortiz, J. A. (2021). Using nonword repetition to identify language impairment in bilingual children: A meta-analysis of diagnostic accuracy. *American Journal of Speech-Language Pathology, 30*(5), 2275-2295.

Schwob, S., Eddé, L., Jacquin, L., Leboulanger, M., Picard, M., Oliveira, P. R., & Skoruppa, K. (2021). Using nonword repetition to identify developmental language disorder in monolingual and bilingual children: A systematic review and meta-analysis. *Journal of Speech, Language, and Hearing Research, 64*(9), 3578-3593.

Thordardottir, E., & Brandeker, M. (2013). The effect of bilingual exposure versus language impairment on nonword repetition and sentence imitation scores. *Journal of Communication Disorders, 46*(1), 1-16.

Windsor, J., Kohnert, K., Lobitz, K. F., & Pham, G. T. (2010). Cross-Language Nonword Repetition by Bilingual and Monolingual Children. *American Journal of Speech-Language Pathology, 19*, 298-310.

SPEECH NATURALNESS AND INTELLIGIBILITY IN CHILDREN WITH DYSARTHRIA: RELATIONSHIPS WITH AUDITORY-PERCEPTUAL CHARACTERISTICS AND THE ISSUE OF SPEECH DEVELOPMENT

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

Background & Aims The consequences of dysarthria on everyday communication are complex and manifold. Symptoms of the speech disorder may for instance limit a speaker´s intelligibility and naturalness as perceived by communication partners. In adults with dysarthria, intelligibility and naturalness have been shown to be closely connected and determined by symptoms related to different speech subsystems (Klopfenstein, 2015; Legner & Ziegler, 2022; Schölderle et al., 2016). In children with dysarthria, communication-related research has predominantly focused on intelligibility, while naturalness has not been studied comprehensively so far, neither on its own, nor in its relationship to intelligibility. Relevant questions remain open, e.g., to what extent intelligibility and naturalness are predictable by the same speech symptoms in children as in adults and what role developmental influences play in this context. Earlier studies comparing children with dysarthria with typically developing children identified specific markers of childhood dysarthria (e.g., *hypernasality, strained-strangled voice, reduced articulatory precision*; see Schölderle et al., under review). It can be expected that these markers play a crucial role in listeners´ impressions of children´s speech, especially of their speech naturalness.

This study aimed (a) to analyze the connection between intelligibility and naturalness in children with neurological conditions, and (b) to predict both parameters by auditory-perceptual characteristics covering all speech subsystems. The discussion will account for previous studies providing a comparison to typically developing children (Schölderle et al., under review; Schölderle et al., 2020).

Methods 28 children with neurological conditions participated in the study (10 Females, 18 Males; range 5-9 years). Speech was assessed using BoDyS-KiD (Haas et al., 2021), a children's version of the Bogenhausen Dysarthria Scales (BoDyS, see Ziegler et al., 2017). This approach allows for systematic auditory-perceptual analyses of all speech subsystems. In an additional listening experiment, naïve listeners completed a sentence transcription task providing intelligibility scores, and rated the children´s speech naturalness on visual analogue scales.

Results & Discussion Intelligibility and naturalness showed a high, significant correlation. Similar to findings in adults, we documented dissociations in one direction, i.e., relatively high intelligibility with low naturalness ratings, but not in the opposite one (Schölderle et al., 2016). Intelligibility was predicted by the articulatory dimension of the BoDyS-KiD, particularly by the symptom reduced articulatory precision. Naturalness was determined by reduced articulatory precision as well, but also by reduced articulation rate. These findings, which demonstrate the relevance of articulatory function for intelligibility and of prosodic features for naturalness, are also in line with reports on adults with dysarthria (Lehner & Ziegler, 2022; Schölderle et al., 2016). While reduced articulatory precision is amongst the previously reported specific markers of childhood dysarthria, reduced articulation rate is not. In fact, a slower rate has been shown to

be among the most prevalent features in typically developing children (Schölderle et al., 2020). The fact that listeners associate not only specific symptoms, but also developmental speech features with unnatural speech must be considered a challenge for speech naturalness as a diagnostic marker in childhood dysarthria.

References

Haas, E., Ziegler, W., & Schölderle, T. (2021). Developmental courses in childhood dysarthria: Longitudinal analyses of auditory-perceptual parameters. *Journal of Speech, Language, and Hearing Research, 64*(5), 1421-1435.

Klopfenstein, M. (2015). Relationship between acoustic measures and speech naturalness ratings in Parkinson's disease: A within-speaker approach. *Clinical Linguistics & Phonetics, 29*(12), 938-954.

Lehner, K., & Ziegler, W. (2022). Indicators of Communication Limitation in Dysarthria and Their Relation to Auditory-Perceptual Speech Symptoms: Construct Validity of the KommPaS Web App. *Journal of Speech, Language, and Hearing Research, 65*(1), 22-42.

Schölderle, T., Staiger, A., Lampe, R., Strecker, K., & Ziegler, W. (2016). Dysarthria in adults with cerebral palsy: Clinical presentation and impacts on communication. *Journal of Speech, Language, and Hearing Research, 59*(2), 216-229.

Schölderle, T., Haas, E., & Ziegler, W. (under review). Childhood dysarthria – auditory-perceptual profiles against the background of typical speech motor development. *Journal of Speech, Language, and Hearing Research.*

Schölderle, T., Haas, E., & Ziegler, W. (2020). Age norms for auditory-perceptual neurophonetic parameters: A prerequisite for the assessment of childhood dysarthria. *Journal of Speech, Language, and Hearing Research, 63*(4), 1071-1082.

Ziegler, W., Staiger, A., Schölderle, T., & Vogel, M. (2017). Gauging the auditory dimensions of dysarthric impairment: Reliability and construct validity of the Bogenhausen Dysarthria Scales (BoDyS). *Journal of Speech, Language, and Hearing Research, 60*(6), 1516-1534.

AN EXPLORATORY ANALYSIS OF INDIVIDUAL-LEVEL PREDICTORS OF CAS TREATMENT RESPONSE

Molly Beiting[1], Jenya Iuzzini-Seigel[2], Edwin Maas[1]
[1] *Temple University, Philadelphia, United States of America*
[2] *Marquette University, Milwaukee, United States of America*

Background Treatment for childhood apraxia of speech (CAS) typically exceeds the intensity and duration required for other speech disorders. Irrespective of treatment approach, studies report wide variability in outcomes (e.g., Campbell, 1999). Treatment-level factors only partially account for the variability in treatment response (e.g., practice schedule, intensity [Edeal & Gildersleeve-Neumann, 2011; Namasivayam et al., 2015; Maas et al., 2019; Maas & Farinella, 2012], feedback frequency [Maas et al., 2012]). Individual-level factors such as disorder severity are also suspected to impact outcomes, but there have been few systematic investigations involving these factors (Murray et al., 2015), and none involving multiple individual-level factors. A better understanding of individual-level variability is needed to plan maximally efficient targeted treatments and generate accurate prognoses.

The present study examines the impact of four individual-level baseline factors on progress in CAS treatment: (1) sound accuracy (2) word inconsistency, (3) speech perception ability, and (4) age.

Methods Participants included 27 children (ages 4-9) involved in an intensive integral simulation-based treatment (Apraxia of Speech Systematic Integral Stimulation Treatment [ASSIST]; Maas et al., 2019). Treatment response was quantified as pre-to-post changes in speech accuracy on individually designed treatment targets. Sound accuracy was captured by Percentage of Consonants Correct (PCC) and Percentage of Vowels Correct (PVC; Shriberg et al., 1997), from single words on the Goldman-Fristoe Test of Articulation (Goldman & Fristoe, 2015). Inconsistency was captured by two measures: the inconsistency percentage from the DEAP Inconsistency subtest (Dodd et al., 2006), which is based on three separate trials of 25 words, and the Inconsistency Severity Percentage (ISP; Iuzzini & Forrest, 2010), which is a segmental measure of inconsistency. Perceptual ability was measured using an auditory discrimination task (Beiting et al., 2022) with averaged scores reported in AZM (Zhang & Mueller, 2005), a nonparametric measure of detection sensitivity that accounts for response bias by considering both hits and false alarms. Age was measured continuously, in months.

We conducted bivariate analyses to examine the correlation between each predictor variable and the outcome. We used multivariable linear regression to determine whether the variables of interest predicted response to treatment. Regression models were constructed using a bidirectional stepwise selection process. Analyses were conducted in R Studio.

Results & discussion Identification of behavioral predictors of CAS treatment outcomes is needed to provide clinicians with essential data to factor into treatment selection decisions, inform prognoses, and guide future development of maximally efficient, precision treatments. Results indicated that speech accuracy, inconsistency, and age were not predictive of CAS treatment response. Within a subgroup of 12 children for whom perceptual data were available, baseline speech perception skills were the sole significant predictor of treatment response. Children with greater perceptual ability demonstrated greater change in speech accuracy as a result of therapy. Further research is needed to replicate and extend these findings. Future CAS treatment research should include comprehensive assessments of speech accuracy, inconsistency, and perceptual ability to enable further predictive research among larger groups.

References

Beiting, M., Iuzzini-Seigel, J., & Maas, E. (2022, February). Speech perception abilities among children with childhood apraxia of speech: Preliminary findings [Poster presentation]. *21st Biennial Conference on Motor Speech*, Charleston, SC.

Campbell, T. F. (1999). Functional treatment outcomes in young children with motor speech disorders. In A. J. Caruso & E. A. Strand (Eds.), *Clinical Management of Motor Speech Disorders in Children.* Thieme.

Dodd, B., Hua, Z., Crosbie, S., Holm, A., & Ozanne, A. (2006). *Diagnostic Evaluation of Articulation and Phonology.* San Antonio, TX: PsychCorp.

Edeal, D. M., & Gildersleeve-Neumann, C. E. (2011). The importance of production frequency in therapy for childhood apraxia of speech. *American Journal of Speech-Language Pathology, 20*(2), 95-110.

Goldman, R., & Fristoe, M. (2015). *GFTA-3: Goldman Fristoe 3 Test of Articulation.* PsychCorp.

Iuzzini, J., & Forrest, K. (2010). Evaluation of a combined treatment approach for childhood apraxia of speech. *Clinical Linguistics & Phonetics, 24*(4-5), 335-345.

Maas, E., & Farinella, K. A. (2012). Random versus blocked practice in treatment for childhood apraxia of speech. *Journal of Speech, Language, & Hearing Research, 55*, 561-578.

Maas, E., Butalla, C. E., & Farinella, K. A. (2012). Feedback frequency in treatment for childhood apraxia of speech. *American Journal of Speech-Language Pathology, 21*, 239-257.

Maas, E., Gildersleeve-Neumann, C., Jakielski, K., Kovacs, N., Stoeckel, R., Vradelis, H., & Welsh, M. (2019). Bang for your buck: A single-case experimental design study of practice amount and distribution in treatment for childhood apraxia of speech. *Journal of Speech, Language, and Hearing Research, 62*, 3160-3182.

Murray, E., McCabe, P., & Ballard, K. J. (2015). A randomized controlled trial for children with childhood apraxia of speech comparing rapid syllable transition treatment and the Nuffield Dyspraxia Programme–Third Edition. *Journal of Speech, Language, and Hearing Research, 58*(3), 669-686.

Namasivayam, A. K., Pukonen, M., Goshulak, D., Hard, J., Rudzicz, F., Rietveld, T., ... & van Lieshout, P. (2015). Treatment intensity and childhood apraxia of speech. *International Journal of Language & Communication Disorders, 50*(4), 529-546.

Shriberg, L. D., Austin, D., Lewis, B., McSweeny, J. L., & Wilson, D. L. (1997). The percentage of consonants correct (PCC) metric: Extensions and reliability data. *Journal of Speech Language and Hearing Research, 40*, 708-722.

Zhang, J., & Mueller, S. T. (2005). A note on ROC analysis and non-parametric estimate of sensitivity. *Psychometrika, 70*, 145-154.

PEDIATRIC MOTOR SPEECH DISORDERS: A PANEL DISCUSSION ON A CONSENSUS DELPHI PROJECT

Edwin Maas[1], Theresa Schölderle[2], Pam Williams[3], Deborah Hayden[4], Jenya Iuzzini-Seigel[5], Ray Kent[6], Elizabeth Murray[7], Ignatius Nip[8], Hayo Terband[9]

[1] *Temple University, Philadelphia, United States of America*
[2] *Ludwig-Maximilians-University, Munich, Germany*
[3] *University College London Hospitals NHS Foundation Trust*
[4] *PROMPT Institute*
[5] *Marquette University, Milwaukee, United States of America*
[6] *University of Wisconsin-Madison, Madison, United States of America*
[7] *The University of Sydney, Sydney, Australia*
[8] *San Diego State University, San Diego, United States of America*
[9] *Department of Communication Sciences and Disorders, University of Iowa*

The evidence base about pediatric motor speech disorders (pMSD) is limited compared to that on MSD in adults. One obstacle to progress in virtually all aspects of clinical practice and research is the lack of guidelines or consensus on assessment protocols and related measures. Developing consensus on the information needed from an assessment is a critical step for achieving the long-term goal of advancing clinical practice and scientific understanding of pMSD.

The Academy of Neurologic Communication Disorders and Sciences (ANCDS) convened a committee to advance clinical practice and research related to pMSD. Long-term goals include creation of clinical practice guidelines, standardization of research practices, and creation of well-characterized resources for clinicians and scientists. To achieve these goals, the committee will first aim to establish consensus on the information needed to identify, characterize, and differentiate speech motor involvement in children suspected of having a MSD. An e-Delphi approach will be adopted, in which experts with a range of perspectives and backgrounds contribute anonymously via multiple iterative rounds of responses.

In anticipation of this e-Delphi project, we seek to increase awareness and solicit input from individuals and stakeholders with different perspectives. To ensure a broad and international representation, we propose a 1-hour session at the SMC conference. First, a representative of the ANCDS pMSD committee will introduce the project (10 minutes). Next, three individuals from different backgrounds will discuss their theoretical/research/clinical background, perspective on this project, and opportunities/barriers for characterizing and differentiating pMSD (15 minutes each). Finally, we will facilitate an open discussion among panelists and audience (20 minutes).

Panelists are:

Dr. Pam Williams (Speech and Language Therapist and Honorary Lecturer; University College London Hospitals NHS Foundation Trust). For over 30 years, Dr. Williams worked at the Nuffield Hearing and Speech Centre, London, known for expertise in CAS. Her research has focused on diadochokinetic skills of children with typically-developing and disordered speech. She is also a former trustee and chair of the Dyspraxia Foundation, a UK charity that supports individuals with coordination difficulties.

Dr. Theresa Schölderle (speech-language pathologist and scientist; Institute for Phonetics, Ludwig Maximilian University, Munich). As a postdoc at the Clinical Neuropsychology Research Group (Institute for Phonetics, LMU Munich), Dr. Schölderle specialized in childhood dysarthria. She currently leads a project developing an assessment approach, which takes developmental aspects into account systematically. She has worked clinically with children and adults with dysarthria in a specialized center and is co-author of a standardized German assessment tool for adults with dysarthria.

This special session will advance the discussion and provide an important launchpoint for improving diagnostic and treatment protocols, developing common procedures and terminology across clinics, and promoting implementation science.

APPLYING OUR EXPERTISE IN MOTOR SPEECH TO CLINICAL TRIALS FOR NEURODEGENERATIVE DISEASE

Adam Vogel
University of Melbourne, Melbourne, Australia
Redenlab Inc., Melbourne, Australia

Neurodegenerative disease can lead to changes in speech. These changes worsen as the disease progresses and can improve with effective treatment. Subtle changes can even occur prior to diagnosis in the case of autosomal dominant disorders. The phenotype varies across indications but can impact specific speech subsystems and lead to global changes in naturalness and intelligibility. Motor speech deficits frequently lead to reduced quality of life.

There is recognition within the pharmaceutical industry and regulatory bodies that clinical trials need to include meaningful outcomes to demonstrate treatment efficacy. The field of digital biomarkers has grown exponentially over the past decade. Inclusion of these technologies in trials has also grown, but an understanding of how to use the data they produce is limited. Similarly, the logistical and theoretical challenges facing trialists remain, including how and when to test, protocol design, and requirements for evidence supporting outcomes.

In this talk I will discuss how speech has been used as an exploratory and secondary outcome in industry run clinical trials across multiple disease groups.

SPEECH FOR HEALTH ANALYSIS: ON AI AND CHALLENGES

Björn W. Schuller[1,2,3]
[1]*Imperial College London, London, United Kingdom*
[2]*University of Augsburg, Augsburg, Germany*
[3]*audEERING*

Speech is increasingly considered "the new blood" when it comes to diagnosis of health. This comes, as computational analysis by artificial intelligence offers considerable potential for earlier and more accessible diagnosis and monitoring at scale for a broad range of health conditions. Such include next to psychological disorders, neurodevelopmental and neurodegenerative ones also diseases of the respiratory, nervous, or musculoskeletal and circulatory systems. While practitioners have always been listening to their patients, computers offer a range of advantages. Such include unshared attention without tiredness, the ability to perfectly memorise patients' health development, near perfect signal resolution, potential objectivity and anonymity, and the ability to learn from more examples than humans could experience in their lifetime if available to name but a few. Here, we shall discuss the latest advancement from the perspective of AI to empower such health monitoring. This includes hot-off the fire deep learning approaches for representation learning and modelling, but also explainability, efficiency, fairness, personalisation, and privacy aspects. From there, we shall move to the most pressing challenges to allow for rapid distribution of approaches ready for tomorrow's global health monitoring – safely, trustworthy, and reliably, and ideally with human experts in the loop.

DEVELOPING NEW SPEECH SIGNAL PROCESSING ALGORITHMS FOR BIOMEDICAL AND LIFE SCIENCES APPLICATIONS: PRINCIPLES, FINDINGS, CHALLENGES, AND A VIEW TO THE FUTURE

Athanasios Tsanas
Telescot, Usher Institute, Edinburgh Medical School, University of Edinburgh, Edinburgh, United Kingdom

I will briefly outline the main physiological principles of voice production and describe how these link to the key concepts for developing speech signal processing algorithms to characterize speech and extract potentially useful information. I will demonstrate the applicability and differences of speech signal processing algorithmic concepts across different applications, in combination with state of the art statistical machine learning techniques. Finally, I will touch on open questions, challenges, and upcoming problems as we develop robust, parsimonious, generalizable decision support tools mining speech signals across diverse biomedical and life sciences applications.

Poster Session I

SPEECH KINEMATICS AND COORDINATION MEASURED WITH A MEG NEUROIMAGING-COMPATIBLE SPEECH TRACKING SYSTEM

Ioanna Anastasopoulou[1], Pascal van Lieshout[2], Douglas Cheyne[2,3], Blake W Johnson[1]
[1]*Macquarie University, Sydney, New South Wales, Australia*
[2]*University of Toronto, Toronto, Canada*
[3]*Hospital for Sick Children Research Institute, Toronto, Canada*

Articulography and functional neuroimaging are two major tools for studying the neurobiology of speech production. Until now, however, it has generally not been feasible to use both in the same experimental setup because of technical incompatibilities between the two methodologies. Here we describe results from a novel articulography system dubbed Magneto-articulography for the Assessment of Speech Kinematics (MASK; Alves et al., 2016), used for the first time to obtain kinematic profiles of oro-facial movements during speech together with concurrent magnetoencephalographic (MEG) measurements of neuromotor brain activity.

MASK was used to characterise speech kinematics in six healthy adults, and the results were compared to measurements from a separate participant with a conventional Electromagnetic Articulography (EMA) system. Analyses targeted the gestural landmarks of reiterated utterances /ipa/ and /api/, produced at normal and faster rates (Van Lieshout, 2007). The results demonstrate that MASK reliably characterises key kinematic and movement coordination parameters of speech motor control. This new capability for measuring and characterising speech movement parameters, and the brain activities that control them, within the same experimental setup, paves the way for innovative cross-disciplinary studies of neuromotor control of human speech production, speech development, and speech motor disorders.

References

Alves, N., Jobst, C., Hotze, F., Ferrari, P., Lalancette, M., Chau, T., van Lieshout, P., & Cheyne, D. (2016). An MEG-Compatible Electromagnetic-Tracking System for Monitoring Orofacial Kinematics. *IEEE Transactions on Biomedical Engineering, 63*(8), 1709–1717. https://doi.org/10.1109/TBME.2015.2500102

Anastasopoulou, I., van Lieshout, P., Cheyne, D. O., & Johnson, B. W. (2022). Speech Kinematics and Coordination Measured With an MEG-Compatible Speech Tracking System. *Frontiers in Neurology, 13.*

Van Lieshout, P. H. H. M., Bose, A., Square, P. A., & Steele, C. M. (2007). Speech motor control in fluent and dysfluent speech production of an individual with apraxia of speech and Broca's aphasia. *Clinical Linguistics & Phonetics, 21*(3), 159–188. https://doi.org/10.1080/02699200600812331

COMPENSATORY MOVEMENT OF THE TONGUE FOR SPEECH PRODUCTION WITH OR WITHOUT MASKING NOISE

Morgane Bourhis[1], Pascal Perrier[1], Christophe Savariaux[1], Takayuki Ito[1,2]
[1] *Univ. Grenoble-Alpes, CNRS, Grenoble-INP, GIPSA-Lab, Grenoble, France*
[2] *Haskins Laboratories, New Haven, CT, United States*

Introduction In speech production, achieving a specific acoustical goal is fundamental for the interaction with listeners. Hence the assessment of the task achievement from auditory feedback certainly plays a major role in correcting and learning speech articulation (Guenther & Hickok, 2015). Somatosensory inputs have been also shown to be important in speech motor control (Patri et al., 2020), but the detailed nature of this sensory feedback and its role in speech movement correction is yet unknown. In a recent study (Ito et al, 2020) when the tongue was suddenly pulled forward due to a tongue perturbation, the induced change of tongue posture was rapidly corrected. The spectral characteristics of the perturbed speech sound were also corrected toward those of the original sound. Since this quick compensatory response was induced faster than any possible cortical auditory feedback loop ($<$ 140 ms for the movement change), it may rely on the sole use of somatosensory information, which would then be tuned toward efficient auditory correction.

Methodology In order to test this hypothesis, we examined whether and how the quick compensatory response of the tongue is induced in the absence of auditory feedback. We carried out the experiment using the same tongue perturbation as in Ito et al (2020), and compared the compensatory responses observed with and without auditory masking. The experiment consisted in 150 trials during which the subjects were asked to whisper vowel /e/ during a few seconds. In half of the trials a pink noise was presented to the subject through earphones so as to mask the auditory feedback. In the other half, normal auditory feedback was received. These two conditions were tested in pseudo-random order. We used a robotic device for the perturbation, in which a 1 Newton force was applied in the forward direction during 1 second. It was applied only in a limited number (15) of randomly selected trials in each auditory condition in order to avoid anticipation. We recorded articulatory movements using ElectroMagnetic Articulography (EMA), in synchrony with the /e/ sound. EMA sensors were attached to the tongue tip (TT), blade (TB), and dorsum (TD), the upper and lower lips and the lower jaw. The first and second formant frequencies were extracted for acoustical evaluation. All data were aligned at the perturbation onset and averaged across trials and subjects.

Results & discussion A quick compensatory movement of the tongue was induced in both auditory feedback conditions in response to the mechanical perturbation. The results including acoustic sound are currently processed and will be presented at the conference. We expect that the auditory feedback had no impact on the early phase of the compensatory movement. This would support our hypothesis that the observed quick feedback is somatosensory rather than auditory in nature. A speech specific auditory-somatosensory mapping for this compensation may be learned during the speech acquisition period, maintained throughout speech development and ontogeny and may be used to tune somatosensory feedback alone toward the preservation of auditory goals.

References

Guenther, F. H., & Hickok, G. (2015). Role of the auditory system in speech production. In Handbook of Clinical Neurology (1st ed., Vol. 129). Elsevier B.V. https://doi.org/10.1016/B978-0-444-62630-1.00009-3

Ito, T., Szabados, A., Caillet, J. L., & Perrier, P. (2020). Quick compensatory mechanisms for tongue posture stabilization during speech production. *Journal of Neurophysiology, 123*(6), 2491–2503. https://doi.org/10.1152/JN.00756.2019

Patri, J. F., Ostry, D. J., Diard, J., Schwartz, J. L., Trudeau-Fisette, P., Savariaux, C., & Perrier, P. (2020). Speakers are able to categorize vowels based on tongue somatosensation. *Proceedings of the National Academy of Sciences, 117*(11), 6255-6263.

USING TDCS TO PROMOTE TARGETED NEUROREHABILITATION OF CORTICAL SPEECH NETWORK IN ACQUIRED APRAXIA OF SPEECH

Adam Buchwald[1], Juhi Kidwai[1], E. Susan Duncan[2]
[1]*New York University, New York, United States of America*
[2]*Louisiana State University, Baton Rouge, United States of America*

Across domains, motor skill acquisition is associated with changes in motor cortex based on practice-dependent plasticity. When combined with motor learning protocols, transcranial direct current stimulation (tDCS) can promote practice-dependent plasticity in targeted neural tissue (Awosika & Cohen, 2019). Thus, tDCS has the potential to enhance rehabilitation of motor speech disorders that respond to speech motor learning intervention. This may be particularly useful for stroke rehabilitation in speech/language, for which engaging cortical regions in the left hemisphere has been associated with stronger behavioral outcomes (Turkeltaub et al., 2011). Despite this promise, the underlying mechanism of tDCS remains incompletely understood, and the results of studies using tDCS to enhance recovery have been equivocal. Here, we examine speakers with acquired apraxia of speech (AOS) subsequent to a single left-hemisphere stroke. We specifically test whether tDCS can enhance effectiveness of speech motor learning treatment if we target the left ventral premotor and motor cortices during treatment in order to (re)strengthen their role in the cortical speech network.

Three participants with acquired AOS completed a four-month multiple-baseline multiple probe crossover single-subject treatment study comparing treatment outcomes during treatment with targeted tDCS vs. sham tDCS. Treatment phases included 12 behavioral sessions (3x/week) focused on specific speech sounds, with the order of active vs. sham tDCS counterbalanced across participants. Electrode placement was personalized for each individual based on their lesion, using an approach that maximized current that targeted our regions of interest – nonlesioned tissue in the ventral pre-motor and motor cortices. We obtained 1mm3 structural scans at baseline and resting-state fMRI (12 mins/session) at baseline as well as after each treatment phase to identify changes in the cortical speech network. We identified regions using the Brainnetome atlas (Fan et al., 2016) which has extensive differentiation among regions in the speech network.

In our primary analysis, we examined changes in a network of regions associated with speech production, including pars opercularis, ventral portions of the precentral gyrus and post-central gyrus, and the anterior insula, building on previous work examining functional connectivity in individuals with acquired AOS (New et al., 2015). We specifically looked at intra- and inter-hemispheric connectivity within the network. These analyses revealed a significant increase in LH connectivity following targeted tDCS in our two stronger behavioral responders (but not our mild responder), consistent with tDCS promoting left hemisphere activation during the speech motor learning task. We are following this analysis by asking whether stronger outcomes are associated with connectivity patterns that more closely resemble control participants. To achieve this, we are defining a cortical speech network based on seed-based analyses of resting-state fMRI data from 65 older individuals with normal cognition in a freely available database (LaMontagne et al., 2019). Specifically, we are identifying the network of regions functionally connected to left ventral premotor/motor cortices. We will then compare examine our AOS participants' changes within this network, to consider whether better treatment outcomes are associated with connectivity patterns similar to unimpaired speakers. Taken together, these analyses will improve our understanding of the mechanism through which tDCS may be used to enhance neurorehabilitation in individuals with acquired deficits in speech motor control.

References

Awosika, O.O., & L.G. Cohen. (2019). Transcranial Direct Current Stimulation in Stroke Rehabilitation: Present and Future. In *Practical Guide to Transcranial Direct Current Stimulation*, 509-539.

Fan, L., et al. (2016). The Human Brainnetome Atlas: A New Brain Atlas Based on Connectional Architecture. *Cereb Cortex, 26*(8): p.3508-26.

LaMontagne, P.J., et al. (2019). OASIS-3: Longitudinal Neuroimaging, Clinical, and Cognitive Dataset for Normal Aging and Alzheimer Disease. *medRxiv*, 2019.12.13.19014902.

New, A.B., et al. (2015). Altered resting-state network connectivity in stroke patients with and without apraxia of speech. *NeuroImage. Clinical, 8*, p. 429-439.

Turkeltaub, P.E., et al. (2011). Are networks for residual language function and recovery consistent across aphasic patients? *Neurology, 76*(20), p. 1726-34.

CONTROL AND REGULATION OF SUBGLOTTAL PRESSURE IN SPEECH

Didier Demolin
LPP CNRS, Grenoble, France

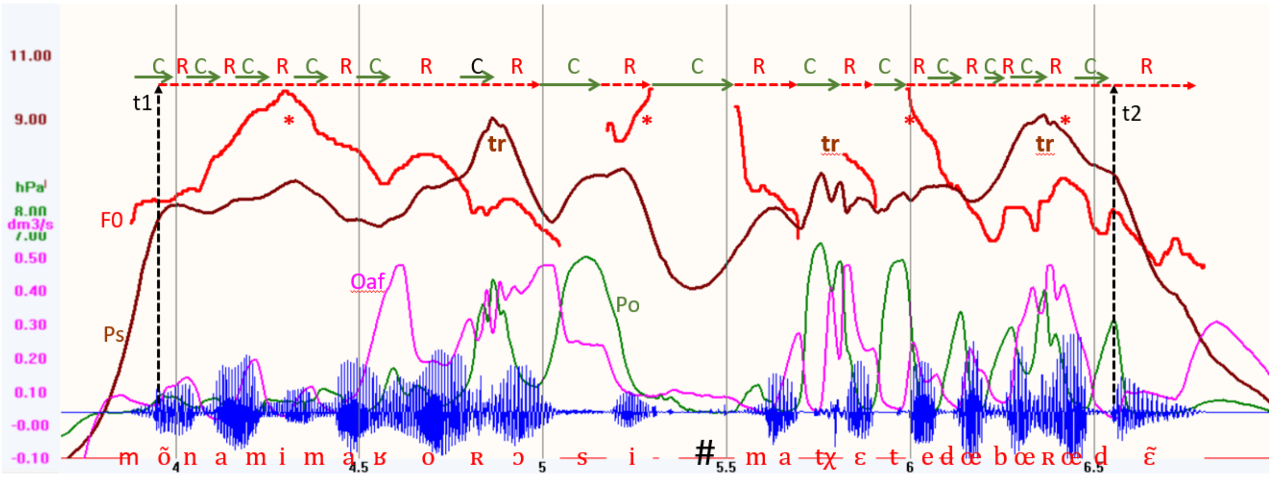
Subglottal pressure (Ps) is maintained at a quasi constant level with some slight fluctuations during speech sentences while the volume of air in the lungs diminishes (Ladefoged, 1967; Hixon et al., 2020) (Figure 1). During speech, i.e in the expiration phase of respiration, the elastic recoil forces trigger a rapid reduction of the pulmonic volume but Ps (sometimes called alveolar pressure) does not fall strongly after the initial rising and before the final falling phase (Figure 1).

There are several factors contributing to maintain this fairly constant Ps. It has been shown that respiratory muscles involved in expiration (internal intercostals, external oblique & rectus abdominis) come gradually into action depending on the length and intensity of sentences (Ladefoged, 1967). However for short and average length sentences (1 to 5 seconds), there doesn't seem to occur a substantial contribution of expiratory muscles to maintain Ps level but from a gradual increase of internal intercostals (Ladefoged, 1967). In addition to the activity of respiratory muscles, the succession of glottal impedance (R) in the larynx and constrictions in the vocal tract (C) also contribute prevent the rapid diminution of Ps (Ladefoged, 1967). Indeed (R) and (C) create an impendance to the passage of airflow coming out of the lungs and contribute to maintain the observed level of Ps. This explantion is sustained by patterns observed during the production of different speech units in sentences. Vowels or diphthongs making a glottal resistance show a gradual fall of Ps but in stressed syllables where Ps rises. Consonants making a complete or partial vocal tract closure show a gradual Ps rise or a flat contour. Specific features such as a positive VOT or aspiration of stops during which the vocal folds are open after the closure release show a fall of Ps. The magnitude of these rises and falls of Ps are of the order of a 2 hPa increase or decrease for stressed syllables and trills but for emphatic stress where it can be higher (Benuerel, 1970).

Data examined for this paper in our database (Demolin et al., 2019) suggest that Ps quasi constant level in speech sentences is also regulated by successive glottal (R) and vocal tract (C) resistances in short and average length sentences.

Material and method Data for this paper come a set of sentences pronounced by 2 native English male speakers (1 English and 1 American) and 4 native French speakers (2 female and 2 male) (Demolin et al., 2019). Ps was measured with a needle (ID 2 mm) inserted in the trachea. Oral airflow (Oaf) with a flexible silicon mask, both synchronized with the audio signal. The microphone was at a quasi constant distance of the lips. The audio signal, oral airflow and subglottal pressure (Ps) were recorded simultaneously with the Physiologia workstation (Teston 1983). The audio signal was digitized at 16 kHz and the physiological data at 2 kHz. Accuracy of measurements are 1 mbar for pressure and 1 ms for time.

The procedure preserved the rights and welfare of human research subjects, in respect of the ethical committee's rules (`https://www.erasme.ulb.ac.be/fr/ethique`).



*Ps and Po (intra oral pressure) in hecto Pascal (hPa); Oaf (oral airflow) in dcm^3/s and F0 curves of the French sentence 'Mon ami Mario Rossi m'a traité de bredin'; t1 indicates the end of the initial Ps rising phase, t2 the start of the Ps falling phase. R = glottal resistance and C = VT constriction; Tr show the increase of Ps for trills. * indicates stressed syllables. # show a pause between the 2 sentence phrases.*

References

Benguerel, A.-P. (1970). Some physiological aspects of stress in French. PhD University of Michigan, Ann Arbor.

Demolin, D., Hassid, S., Ponchard, C., Yu, S., & Trouville, R. (2019). Speech aerodynamic database. *LPP, ILPGA*, ArtSpeech.

Hixon, T.J. Weismer, G. & Holt, J.D. (2020). *Preclinical Speech Science. Anatomy, Physiology, Acoustics, and Perception.* San Diego. Plural Publishing.

Ladefoged, P. (1967). *The areas of experimental phonetics.* London. Oxford University Press.

ACOUSTIC AND AERODYNAMIC CONSEQUENCES OF THE NASAL POLYPOSIS PATHOLOGY

Amélie Elmerich[1], Angélique Amelot[1], Lise Crevier-Buchman[1], Shinji Maeda[1], Jean-François Papon[2]
[1]*CNRS, Université Sorbonne-Nouvelle, Paris, France*
[2]*Hôpital Bicêtre, Service d'Oto-Rhino-Laryngologie, Le Kremlin Bicêtre, France*

Background Nasal cavities take part of the resonant system of human to produce nasal phonemes. Nasal polyposis, which is an inflammation of the sino nasal mucosal, interferes with the configuration and airflow passage within the cavities, affecting not only the resonance but also the quality of nasal sounds (Hong et al., 1997). Depending on the degree of severity and in response to ineffective drug treatment, sinus surgery may be necessary.

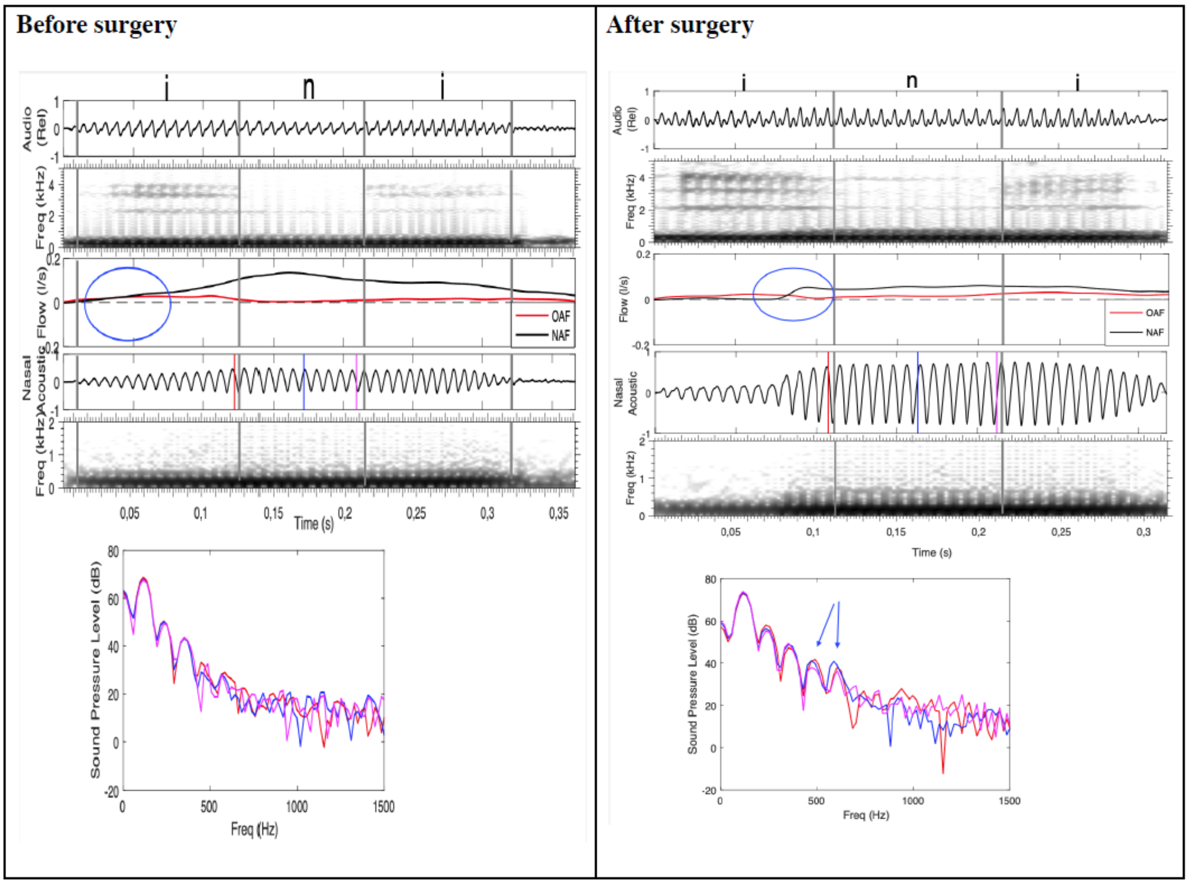
Objectives The goal of the current study is to evaluate the acoustic and aerodynamic effects of the nasal polyposis pathology from preoperative and postoperative data.

Method Four male patients with nasal polyposis participated. They read a text with VNV and VCV sequences (V = [i, a, u, y, e, o]; N = [m, n], C = [p, b, t, d, s, z], n = 912). Nasal (NAF) and oral (OAF) airflow in liter/second (l/s) were recorded using a pneumotachograph mask (3M9710E) developed in our laboratory (Elmerich et al., 2019), twice, preoperatively and at the third month postoperatively. Audio signals were captured with a microphone (AKG C520L). Also, acoustic signals were derived from the pressure sensor in the nasal compartment of the mask. Those three signals are simultaneously recorded with the sampling frequency of 20kHz using a PC operated DataTransaltion acquisition card. Spectral slices were calculated in the nasal acoustic signals at the end of the preceding vowel, at the center and at the end of the nasal consonant. Average airflow was calculated.

Results For this report we focus on one patient (PM03). Aerodynamically, in preoperative, in the [ini] sequence (figure below), there is nasal airflow gradually and continuously, starting at the beginning of the first vowel. The amount of NAF decreases from 0.02 to -0.0001 l/s for the first vowel in postoperative. The opening of the velopharyngeal port is anticipated. The crossing point (marked by the blue circle) between NAF and OAF in preoperative is gradual and less salient than in postoperative. This point comes at the end of the first vowel in postoperative; the opening of the velopharyngeal port appears to be punctually coordinated with the formation of the palatal closure.

Acoustically, in the spectral slice, postoperatively, we observe more intensity below 500 Hz and a double peak at 500 Hz (blue arrows in the graphics) probably due to the coupling between nasal cavities and sinuses (Maeda, 1982). This coupling can be better in postoperative because of the absence of polyps and the enlargement of the communication between nasal cavities and sinuses. The preoperative spectra appear to be highly dumped, while the postoperative ones to be resonant.

Conclusion Nasal polyposis perturbs aerodynamic and acoustics. The surgery can improve it.



Acoustic and airflow records of [ini] utter by a patient before and after operation. From top to bottom row, (1) audio signal captured with a head-set microphone, (2) its spectrogram, (3) the nasal airflow (NAF in black) and oral airflow (OAF in red), (4) nasal acoustic signal, (5) its spectrogram, and (6) spectral slices sampled at the i-n boundary, n-center, and n-release (n-i boundary) are illustrated. Note that nasal acoustic signal is captured by the pressure sensor placed inside the nasal compartment of the airflow mask. Components from DC to a low frequency, say 40 Hz are considered to related to airflow and those higher than 40 Hz to acoustic waves, i.e. sounds.

References

Elmerich, A., Amelot, A., Maeda, S., Laprie, Y., Papon, J. F., & Crevier-Buchman, L. (2020). F1 and F2 measurements for French oral vowel with a new pneumotachograph mask. In *ISSP 2020-12th International Seminar on Speech Production.*

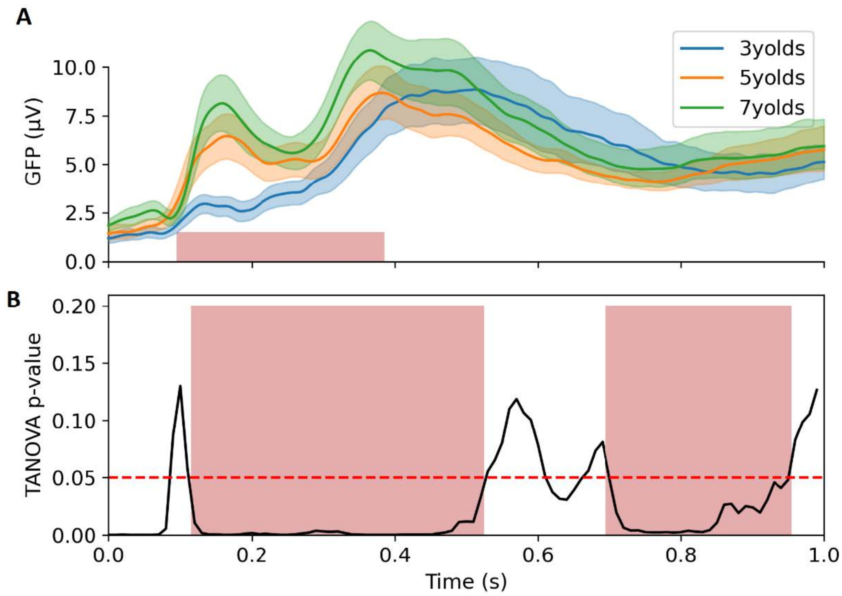
Hong, K. H., Kwon, S. H. & Jung, S. S. (1997). The Assessment of Nasality with a Nasometer and Sound Spectrography in Patients with Nasal Polyposis. *Otolaryngology-Head and Neck Surgery,117*(4), 343-348. https://doi.org/10.1016/S0194-5998(97)70124-4

Maeda, S. (1982). The role of the sinus cavities in the production of nasal vowels. In ICASSP'82. *IEEE International Conference on Acoustics, Speech, and Signal Processing, 7*, 911-914.IEEE.

NEUROPHYSIOLOGICAL PROCESSES UNDERPINNING SPEECH PRODUCTION IN 3-, 5-, AND 7-YEAR-OLD CHILDREN

Doreen Hansmann[1], Reza Shoorangiz[1,3], Catherine Theys[1]
[1] *University of Canterbury, Christchurch, New Zealand*
[3] *New Zealand Brain Research Institute, Christchurch, New Zealand*

Speech production is a complex process, involving multiple processing steps such as word selection, retrieval of the relevant sounds and their articulation (Levelt et al., 1999). Over the last decades EEG research has explored the temporal and topographic dynamics of these underlying processes in school-aged children and adults, revealing spatio-temporal changes of brain activity during development (Indefrey & Levelt, 2004; Indefrey, 2011; Laganaro et al., 2015). With increasing age, a reduction in processing times was observed. In addition, differences in topographic patterns have been reported during the early visual and linguistic processing stages (Laganaro et al., 2015). These findings show the importance of identifying maturational changes to allow accurate interpretation of the neural correlates underlying speech processing in children. However, previous data focused on primary school children (>7-year-olds) and we currently do not know if similar changes can be expected in preschoolers. We therefore aimed to compare the neurophysiological processes associated with speech production in 3-, 5- and 7-year-old typically developing children. EEG data was collected during picture naming in twenty 3-, 5- and 7-year-olds using a 32-channel BioSemi ActiveTwo-system. During offline analysis, the PREP pipeline was used to identify artefactual electrodes and minimize line noise (Bigdely-Shamlo et al., 2015). Visual inspection of independent components was done to minimize stereotypical artefacts (Delorme et al., 2007). EEG was bandpass filtered between 0.2 Hz and 30 Hz and segmented from 200ms pre-stimulus to 1s post-stimulus. To further improve the quality of data and reject noisy epochs, Autoreject was applied (Jas et al., 2017). Finally, Ragu was used to analyse event-related potentials and global field power (Habermann et al., 2018). To avoid finding map differences due to scaling of EEG, data was normalized prior to performing topographic ANOVA (TANOVA) (Habermann et al., 2018; Murray et al., 2008). Results were corrected for multiple comparisons using global duration statistics. Our stimulus-locked global field power results (Fig. 1) show a main effect of group from 100-380ms ($\eta\hat{2}=0.26$, $p<0.01$; 3yo < 5&7yo). TANOVA results indicated a main effect of group between 120-520ms ($\eta\hat{2}=0.09$, $p<0.01$) and 700-950ms ($\eta\hat{2}=0.08$, $p<0.01$). Post-hoc comparisons showed that during 120-520ms, there was a difference between 3yo and 5yo ($p=0.01$), 3yo and 7yo ($p<0.01$), and 5yo and 7yo ($p<0.01$) children. During the 700-950ms period, we found a difference between 3yo and 5yo ($p=0.02$) and 3yo and 7yo ($p<0.01$) children, but no difference between 5yo and 7yo ($p>0.3$) children. Our results showed neurophysiological differences in 3-, 5- and 7-year-old children during picture naming. Temporal and topographic dynamics of electrical brain activity differed across all 3 age groups, illustrating developmental differences in the early visual and linguistic processing stages with increasing age. Microstate analysis will further refine our results, including those in the later motor execution phase, by identifying spatio-temporal dynamics across our 3 age groups. These results extend previous findings in older children (Laganaro et al., 2015) by providing essential information on neurophysiological changes associated with speech production during the preschool years, a period of rapid cognitive development.



A. *Global field power (GFP) for stimulus-aligned event-related potentials (ERPs) for 3-, 5-, and 7-year-old children, significant portion is highlighted in red.* ***B.*** *Periods of significant differences in the TANOVA analysis throughout the stimulus-aligned ERPs are displayed in red.*

References

Bigdely-Shamlo, N., Mullen, T., Kothe, C., Su, K. M., & Robbins, K. A. (2015). The PREP pipeline: standardized preprocessing for large-scale EEG analysis. *Frontiers in neuroinformatics, 9*, 16.

Delorme, A., Sejnowski, T., & Makeig, S. (2007). Enhanced detection of artifacts in EEG data using higher-order statistics and independent component analysis. *Neuroimage, 34*(4), 1443-1449.

Habermann, M., Weusmann, D., Stein, M., & Koenig, T. (2018). A student's guide to randomization statistics for multichannel event-related potentials using Ragu. *Frontiers in neuroscience, 12*, 355.

Indefrey, P. (2011). The spatial and temporal signatures of word production components: a critical update. *Frontiers in psychology, 2*, 255.

Indefrey, P., & Levelt, W. J. (2004). The spatial and temporal signatures of word production components. *Cognition, 92*(1-2), 101-144.

Jas, M., Engemann, D. A., Bekhti, Y., Raimondo, F., & Gramfort, A. (2017). Autoreject: Automated artifact rejection for MEG and EEG data. *NeuroImage, 159*, 417-429.

Laganaro, M., Tzieropoulos, H., Frauenfelder, U. H., & Zesiger, P. (2015). Functional and time-course changes in single word production from childhood to adulthood. *NeuroImage, 111*, 204-214.

Levelt, W. J., Roelofs, A., & Meyer, A. S. (1999). A theory of lexical access in speech production. *Behavioral and brain sciences, 22*(1), 1-38.

Murray, M. M., Brunet, D., & Michel, C. M. (2008). Topographic ERP analyses: a step-by-step tutorial review. *Brain topography, 20*(4), 249-264.

TONGUE STRETCH REFLEX FOR SPEECH MOTOR CONTROL

Takayuki Ito[1,2], Mohamed Bouguerra[1], Morgane Bourhis[1], Pascal Perrier[1]
[1] *Univ. Grenoble Alpes, CNRS, Grenoble INP, GIPSA-lab, Grenoble, France*
[2] *Haskins Laboratories, New Haven, CT, United States*

Stretch reflex is known as one mechanism that contributes to the stability of human sensorimotor control. When a muscle is stretched by unknown disturbance, the muscle makes a quick contraction to maintain the posture (and then keep its length constant). Although this quick reflex has been quite extensively investigated in limb system, the extent to which it contributes to the motor control of the tongue remains unknown. While a previous study failed to induce stretch reflex in the tongue (Neilson et al., 1979), we have found behavioral evidence that the tongue shows relatively quick reaction of motion in response to external mechanical perturbation (Ito et al., 2020). Although the latency of this behavioral response was in the range of reflex delay, the involvement of reflex mechanism is still uncertain due to lack of neurophysiological evidence. This study aims at clarifying this issue with the recording of the muscle activity induced by a quick stretch of the tongue during speaking. In the test, the tongue stretch was produced using a precisely-controlled robotic device. Electromyographic response (EMG) was recorded from the anterior part of the mouth floor using a uni-polar surface Ag-AgCl electrode as done in Ishiwata et al., (1997). The recorded EMG signal contains muscle activation in the Genioglossus and of the Geniohyoid muscles. These muscles are involved in the production of the vowel /i/ (Miyawaki et al., 1975; Baer et al., 1988; Buchaillard et al., 2009). Since our previous study (Ito et al., 2020) also found a clear compensatory response in the production of vowel /i/, this study focuses on the production of vowel /i/.

Participants were asked to sustain this vowel for 3 s. In addition to this speech condition, we also carried out a non-speech condition in which participants were asked to produce the same magnitude of muscle activation, but without articulating any speech sound, and a rest condition associated with little or no muscle activation. For comparison with the latency of voluntary reaction, we also carried out a reaction task, in which the subjects were asked to produce a strong muscle activation once they perceived the perturbation force during the production of /i/. Our preliminary result showed that the increase of muscle activation in response to the perturbation force was induced in the periods corresponding to the latency of compensatory response observed in our previous behavioral study. The magnitude of the response was larger in the speech condition than in the non-speech condition. The response was also induced earlier than the voluntary reaction task. These preliminary results are consistent with the hypothesis that the observed response is driven by reflex mechanism. Since the latency of the observed response was comparable with the cortical reflex for lip compensation in speech (Ito et al., 2005), the current reflex may also be mediated within the cortical loop.

References

Baer T, Alfonso PJ, Honda K (1988) Electromyography of the tongue muscles during vowels in /əpVp/ environment. *Annual Bulletin RILP 22*:7–19.

Buchaillard S, Perrier P, Payan Y (2009) A biomechanical model of cardinal vowel production: muscle activations and the impact of gravity on tongue positioning. *J Acoust Soc Am 126*:2033–2051.

Ishiwata Y, Hiyama S, Igarashi K, Ono T, Kuroda T (1997) Human jaw-tongue reflex as revealed by intraoral surface recording. *J Oral Rehabil 24*:857–862.

Ito T, Kimura T, Gomi H (2005) The motor cortex is involved in reflexive compensatory adjustment of speech articulation. *Neuroreport 16*:1791–1794.

Ito T, Szabados A, Caillet J-L, Perrier P (2020) Quick compensatory mechanisms for tongue posture stabilization during speech production. *J Neurophysiol 123*:2491–2503.

Miyawaki K, Hirose H, Ushijima T, Sawashima M (1975) A preliminary report on the electromyographic study of the activity of the lingual muscles. *Ann Bull RILP 9*:91–106.

Neilson PD, Andrews G, Guitar BE, Quinn PT (1979) Tonic stretch reflexes in lip, tongue and jaw muscles. *Brain Res 178*:311–327.

WITHIN-TALKER STABILITY OF INTER-ARTICULATORY STRATEGIES IN RESPONSE TO CUED SPEECH MODIFICATIONS: AN ANALYSIS ACROSS MULTIPLE TIME POINTS

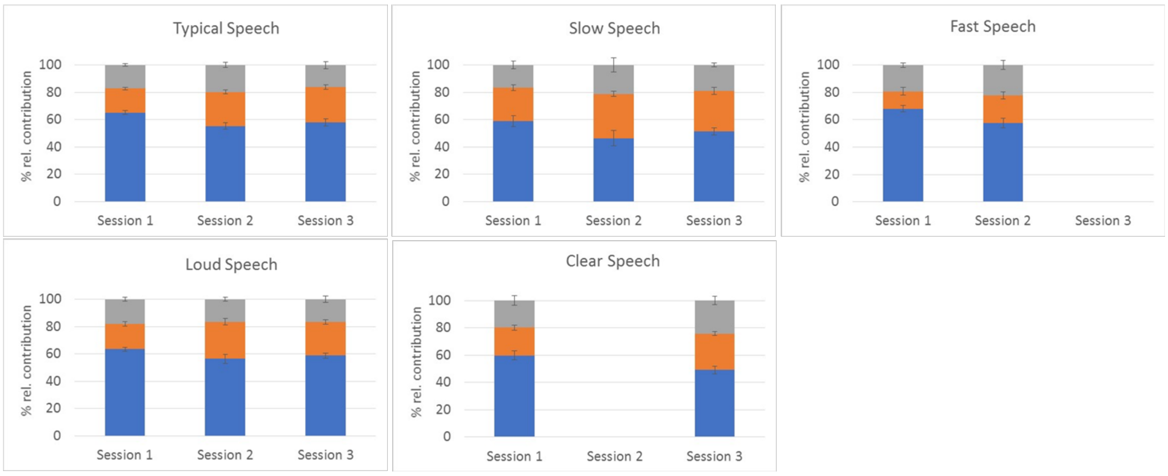
Daniel Kim, Antje S. Mefferd
Vanderbilt University Medical Center, Nashville, United States of America

Rationale Motor equivalence allows talkers to use various inter-articulatory strategies to achieve the same auditory-perceptual outcome (e.g., Abbs & Cole, 1987; Hughes & Abbs, 1976). Such articulatory flexibility is thought to explain, at least in part, kinematic performance variability commonly observed across typical talkers. Inter-talker variability has also been well-documented for inter-articulatory strategies in response to various cued speech modifications (e.g., Mefferd et al., 2019). For example, in response to cued slow speech some talkers decrease their relative contribution of the jaw and increase their relative contribution of the lower lip while others implement the opposite articulatory strategy or do not change their relative contributions at all (e.g., Hertrich & Ackermann, 2000). However, it is currently unclear how stable such inter-articulatory strategies are within the same talker.

Based on the motor schema theory (Schmidt, 1975), motor variability is presumably low when talkers repeat an utterance immediately back-to-back in a blocked fashion. That is because talkers are thought to draw from the same general motor program (GMP). Yet, it remains unclear if and to what extent inter-articulatory strategies in response to cued speech modifications vary within the same talker when utterances are performed at different time points (e.g., days apart). Such knowledge can inform current theories on speech motor control and provide an important context for future studies on talkers with dysarthria, apraxia of speech, and stuttering. Therefore, this study sought to determine the within-talker stability of inter-articulatory strategies in response to cued speech modifications across multiple time points.

Methods Ten typical talkers completed three data recording sessions which were one week apart from another. During each session, talkers generated sets of 15 repetitions of "Buy Bobby a puppy" under five speech conditions: typical, fast, slow, clear, and loud speech. Kinematic data was collected using six infrared motion capture cameras (Motion Analysis, Ltd.). Speech movements were tracked with small reflective markers placed on the midline of the upper lip, lower lip, and jaw. Four markers were attached to the forehead and served as reference markers. The Euclidean distance signals between one reference marker and each articulator-based marker were used to extract upper lip, lower lip, and jaw displacement associated with the closing movement during "Bob." Lower lip movements were decoupled from the jaw using linear subtraction. The % relative contribution of each articulator to lip closure was calculated. The resulting dataset was submitted to a linear mixed model. Segment duration and vocal intensity were extracted and submitted to regression analyses to determine if session-to-session variability in task performance (rate, loudness) accounts for variance in inter-articulatory strategies within talkers.

Results & Discussion Preliminary findings of one talker are presented in Figure 1. Significant session effects were observed for all speech conditions across all articulators. Only for loud and slow speech variability in task performance (segment duration) explained in part variance in inter-articulatory strategies ($R^2 \leq .20$). If the full dataset reveals similar findings, it suggests that inter-articulatory strategies are temporarily stable but can vary significantly across sessions. Theoretical implications will be discussed.



Preliminary findings based on the articulatory performance of one talker. The relative contribution of the jaw, lower lip, and upper lip towards lip closure is presented across three data recording sessions. Blue = jaw, orange = independent lower lip, grey = upper lip. Each condition consists of 15 repetitions. Error bars = SD across repetitions. Clear and fast speech conditions have missing data for session 2 and 3, respectively.

References

Abbs, J., & Cole, K.J. (1987). The neural mechanisms of motor equivalence and goal achievement. In S.P. Wise (Ed.), *Higher brain functions: Recent explorations of the brain's emergent properties* (pp. 15-43). New York Wiley.

Hertrich, I., & Ackermann, H. (2000). Lip-jaw and tongue-jaw coordination during rate-controlled syllable repetitions. *Journal of the Acoustical Society of America, 107*(4). 2236- 2247.

Hughes, O. & Abbs, J. (1976). Labial-mandibular coordination in the production of speech: Implications for the operation of motor equivalence. *Phonetica. 33*(3). 199-221.

Schmidt, R. (1975). A schema theory of discrete motor skill learning. *Psychological Review, 82*, 225-260.

Mefferd, A.S., Efionayi, L. & Mouros, S. (2019). Tongue- and jaw-specific response patterns to speaking rate manipulations. In Sasha Calhoun, Paola Escudero, Marija Tabain & Paul Warren (eds.) *Proceedings of the 19th International Congress of Phonetic Sciences*, Melbourne, Australia 2019 (pp. 3706-2710).

BREATHING AND SPEECH ADAPTATION: DO INTERLOCUTORS ADAPT THEIR SPEECH TOWARDS A SPEAKER TALKING UNDER PHYSICAL EFFORT?

Tom Offrede[1], Susanne Fuchs[2], Christine Mooshammer[1]
[1]*Humboldt-Universität zu Berlin, Berlin, Germany*
[2]*Leibniz-Centre General Linguistics (ZAS), Berlin, Germany*

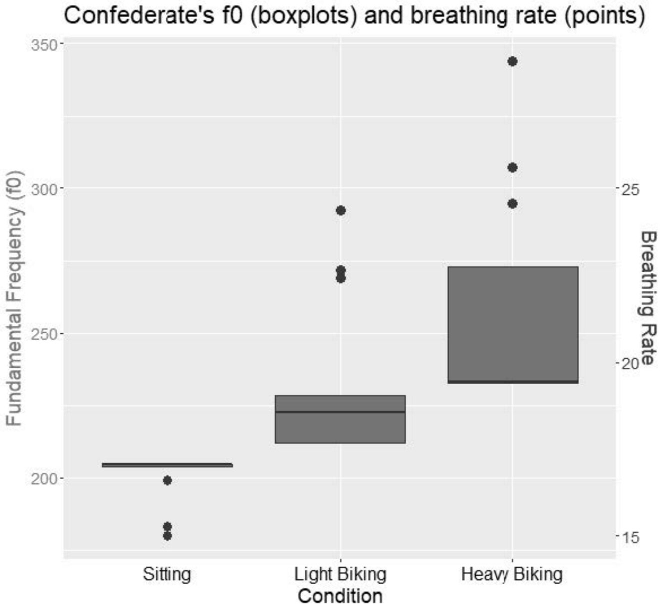
Respiration in humans does not only depend on metabolic needs, but is also sensitive to cognitive and social phenomena. Paccalin and Jeannerod (2000), for instance, reported that individuals ob-serving an actor performing physical activity of varied levels of effort increase their own breathing rate. Respiratory behavior may thus be involved in the action-perception loop or reflect, for example, empathy of the observer. One process with which breathing is intrinsically related is speech (Fuchs & Rochet-Capellan, 2021). Physical exercise (which affects breathing) has been demonstrated, for example, to increase fundamental frequency (f0), intensity, speech rate, jitter, and shimmer in speech (Fuchs et al., 2015; Primov-Fever et al., 2013), although the mechanisms through which those changes happen are not yet clear (Weston et al., 2020). Finally, breathing modulations of speech may be even more complex during dialogue. In this case, one's speech is also affected by that of one's interlocutor (e.g., Pardo, 2006)—a phenomenon sensitive to social dynamics (e.g., Gallois et al., 2005).

The current study investigated whether interlocutors adapt their breathing and speech characteris-tics to a confederate who is performing physical activity at different levels of intensity. Following Paccalin and Jeannerod (2000), we would expect breathing rate to change in the direction of the con-federate's, even if the interlocutor is physically inactive. This might in turn also affect their speech acoustics.

We recorded the speech and breathing activity of 22 native speakers of German (all assigned female at birth, with a mean age of 27.05, SD = 7.63). They performed three speech tasks—free speech about a prompted topic, synchronous reading, and reading alone (only the first two are discussed here)—while in 'interaction' with a confederate who appeared on a pre-recorded video projected on the wall. During the filming of the confederate's video, her acoustic and respiration data were also recorded. There were three experimental conditions: the confederate was (a) sitting, (b) biking with light effort, or (c) biking with heavier effort. The participants were exposed to all conditions in counterbalanced order.

As expected, during free speech, the confederate's breathing rate and f0 increased with physical effort (see Figure 1). The participants' acoustic and breathing data showed large interindividual vari-ability. Overall, during synchronous read speech, the participants tended to produce higher f0s and faster breathing rates with increase in the confederate's physical effort (although the latter with a small effect size). In contrast, during free speech, they did not seem to adapt their f0 or speech rate to the confederate. Their breathing rate, in turn, tended to decrease with increase in confederate's physical effort, although this pattern had a small effect size. This trend is opposite to our hypothesis. Finally, when the participants were watching the confederate while both were silent, their breathing rates did not seem to be affected directly; however, the amplitude of their inhalations got lower with rise in confederate's effort.

Our findings indicate that while participants may adapt their respiration and speech acoustics to a confederate when reading synchronously, this does not seem to extend to spontaneous speech. The absence of breathing adaptation may be due to the participants' respiration being modulated primarily by their speech, and not as strongly by their observation of the confederate's action. The absence of speech adaptation, on the other hand, could be explained by there not being enough turn changes for the confederate's speech acoustics to feed the participants' perception–production system.



The confederate's f0 (boxplots) refers to the left-side y axis; her breathing rate (points) refers to the right-side y axis.

References

Fuchs, S., Reichel, U. D., & Rochet-Capellan, A. (2015). Changes in speech and breathing rate while speaking and biking. `https://doi.org/10.5282/ubm/epub.25254`

Fuchs, S., & Rochet-Capellan, A. (2021). The respiratory foundations of spoken language. *Annual Review of Linguistics, 7*, 13–30.

Gallois, C., Ogay, T., & Giles, H. (2005). Communication accommodation theory: A look back and a look ahead. *Theorizing about intercultural communication* (pp. 121–148). Thousand Oaks: Sage.

Paccalin, C., & Jeannerod, M. (2000). Changes in breathing during observation of effortful actions. *Brain Research, 862*(1-2), 194–200. `https://doi.org/10.1016/S0006-8993(00)02145-4`

Pardo, J. S. (2006). On phonetic convergence during conversational interaction. *The Journal of the Acoustical Society of America, 119*(4), 2382–2393. `https://doi.org/10.1121/1.2178720`

Primov-Fever, A., Lidor, R., Meckel, Y., & Amir, O. (2013). The effect of physical effort on voice characteristics. *Folia Phoniatrica et Logopaedica, 65*(6), 288–293. `https://doi.org/10.1159/000361047`

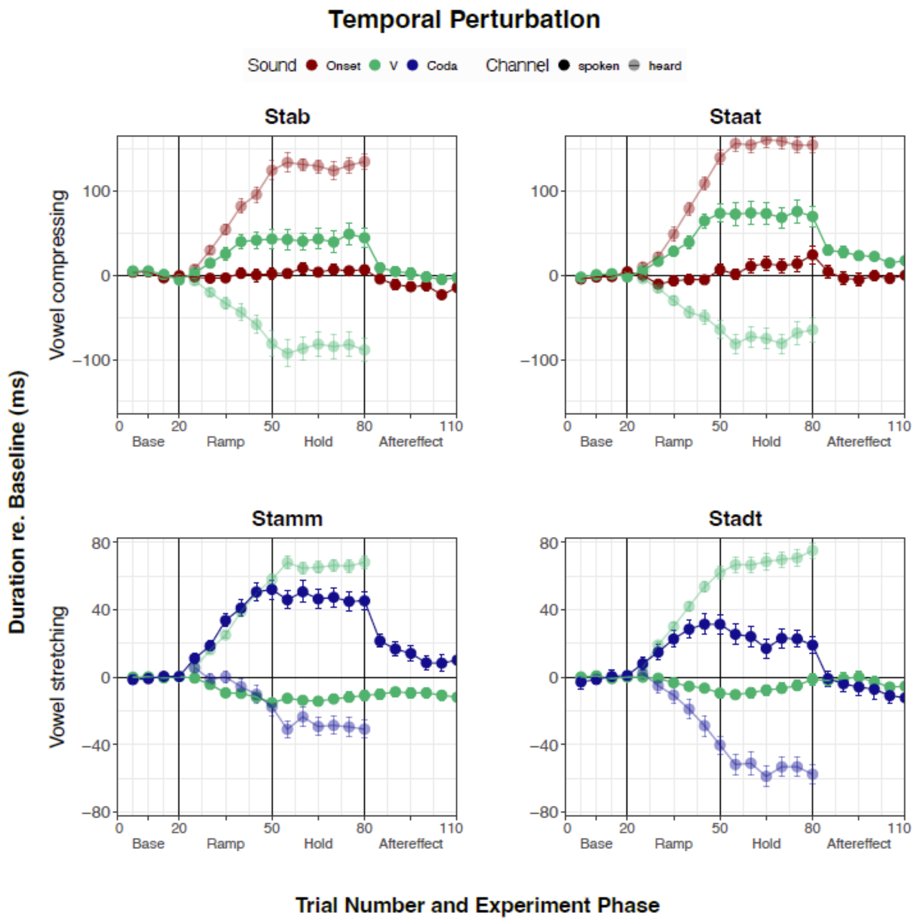
Weston, H., Fuchs, S., & Rochet-Capellan, A. (2020). Speech during light physical activity: Effect on f0 and intensity. *Proceedings of the 12th International Seminar on Speech Production.*

TEMPORAL PERTURBATION OF QUANTITY CONTRASTS BETWEEN AND WITHIN A LEXICAL CATEGORY

Miriam Oschkinat[1], Eva Reinisch[2], Phil Hoole[1]
[1]*Ludwig-Maximilians-University, Munich, Germany*
[2]*Austrian Academy of Sciences, Vienna, Austria*

Research about perceptual identification of words showed that speakers tend to hear words rather than non-words in tokens ambiguous along a word/non-word continuum, known as the *Ganong effect* (Ganong, 1980). This effect provided evidence that in speech perception, lexical decision making does not rely on purely auditory cues but also on knowledge about the linguistic system. The current study tests the effect of lexical categorization when the speaker's own auditory feedback is altered, causing a change in lexical category. With real-time auditory feedback perturbation, we examine the perceptual categorization of ambiguous words and assess the consequences of the shift in production. Real-time auditory feedback perturbations typically trigger compensatory responses to an applied shift with productions opposed to the direction of the manipulation (see Caudrelier and Rochet-Capellan, 2019 for an overview). While this effect has been extensively studied for spectral alterations (e.g., first or second formant frequencies in vowels), recent studies also found compensatory responses in manipulations of temporal parameters in speech (Floegel et al., 2020; Oschkinat and Hoole, 2020; Karlin et al., 2021; Oschkinat and Hoole, 2022). In the current study, the duration of the German vowels /a/ and /a:/, a phoneme contrast realized in quantity without strong additional spectral cues, was manipulated in real-time. Participants produced the German words *Stab* (/ʃtap/ "pole") and *Staat* (/ʃtat/ "state") whereby the onset /ʃt/ was stretched and the vowel /a:/ compressed (vowel compression condition). In a second condition, the words *Stamm* (/ʃtam/ "trunk") and *Stadt* (/ʃtat/ "city") were manipulated by stretching the vowel /a/ and compressing the coda /t/ or /m/ (vowel stretching condition). While *Staat* and *Stadt* form a minimal pair in German, *Stamm* and *Stab* do not have lexical neighbors with respect to vowel quantity. We expect compensatory responses to be more pronounced in words with a lexical neighbor (*Staat*, *Stadt*) due to the speaker's aim to perceive a stimulus that lies within the intended lexical category. Figure 1 visualizes the duration of the segments of interest throughout the experiment for each word. The produced signal is visualized in solid points/lines, the perturbed (perceived) signal in high transparency. Results showed compensatory responses in production in the opposite direction to the vowel manipulation in all words (hold phase compared to base productions, Figure 1). We found larger effects for *Staat* than *Stab* as expected, but greater reactions for *Stamm* than for *Stadt* (contrary to our hypothesis). The former findings are in line with the study by Bourguignon et al. (2014) who found more compensation for spectral shifts that changed the lexical category to a/another real word rather than to a pseudo-word.

For *Stadt* and *Stamm*, results indicate that the vowel manipulation as well as the perceivability of manipulations in the coda segment shape the reactions. These findings provide new insights into the link between perception and production in the online control of the temporal structure of speech, and suggest that the perceptual Ganong effect could be extended to speech production (Bourguignon et al., 2014).



Durations relative to the baseline mean per segment binned per 5 trials over the course of the experiment. Perturbation was applied in phases (Base(line): no perturbation, Ramp: increasing perturbation, Hold: maximum perturbation, aftereffect: no perturbation). Results refer to the durations in the Hold phase compared to the baseline. Solid lines indicate the spoken signal, transparent lines the received feedback. The vowel compression condition in the upper panels, the vowel stretching condition in the lower panels.

References

Bourguignon, N.J., Baum, S.R., & Shiller, D.M. (2014). Lexical-perceptual integration influences sensorimotor adaptation in speech. *Frontiers in human neuroscience 8*, 208.

Caudrelier, T., and Rochet-Capellan, A. (2019). "Changes in speech production in response to formant perturbations: An overview of two decades of research," in *Speech production and perception: Learning and memory*, eds. Susanne Fuchs, Joanne Cleland & A. Rochet- Capellan. (Berlin: Peter Lang), 15-75.

Floegel, M., Fuchs, S., & Kell, C.A. (2020). Differential contributions of the two cerebral hemispheres to temporal and spectral speech feedback control. *Nature Communications 11*:2839, 1-12.

Ganong, W.F. (1980). Phonetic categorization in auditory word perception. Journal of experimental psychology: *Human perception and performance 6*(1), 110.

Karlin, R., Naber, C., and Parrell, B. (2021). Auditory Feedback Is Used for Adaptation and Compensation in Speech Timing. *Journal of Speech, Language, and Hearing*

Research. https://doi.org/10.1044/2021_JSLHR-21-00021

Oschkinat, M., & Hoole, P. (2020). Compensation to real-time temporal auditory feedback perturbation depends on syllable position. *The Journal of the Acoustical Society of America 148*(3), 1478-1495. https://doi.org/10.1121/10.0001765

Oschkinat, M., & Hoole, P. (2022). Reactive feedback control and adaptation to perturbed speech timing in stressed and unstressed syllables. *Journal of Phonetics 91*, 101133. https://doi.org/10.1016/j.wocn.2022.101133

INVESTIGATING FEEDBACK AND FEEDFORWARD CONTROL DURING VOWEL PRODUCTION BY DUTCH ADULT SPEAKERS: INSIGHTS FROM AUDITORY FEEDBACK PERTURBATION TASKS

Teja Rebernik[1,2], Jidde Jacobi[1,3], Roel Jonkers[1], Aude Noiray[4,5], Sanne Oud[1], Aliene Reinders[1], Martijn Wieling[1,5]

[1] *University of Groningen, Groningen, The Netherlands*

[2] *Netherlands Cancer Institute, Amsterdam, The Netherlands*

[3] *Macquarie University, Sydney, Australia*

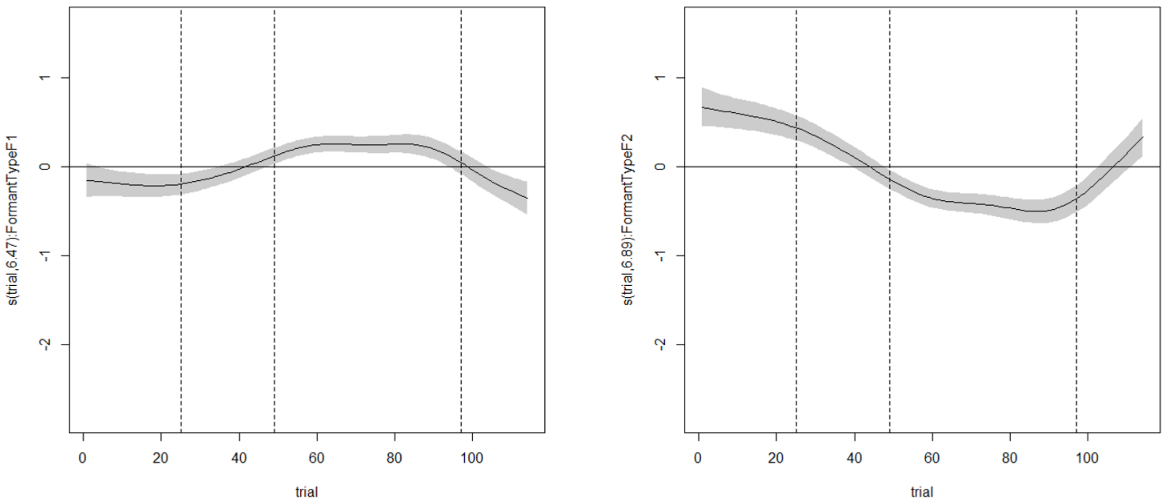
[4] *Laboratoire Dynamique du Langage (DDL), Lyon, France*

[5] *Haskins laboratories, New Haven, CT, United States of America*

Introduction In this study, we investigated how Dutch adult speakers respond to changes in their auditory feedback. Specifically, participants produced speech while their formants were shifted upwards or downwards in either an unexpected or a sustained manner: an unexpected perturbation taps into a speaker's reflexive responses and the functioning of their feedback system, while a sustained perturbation assesses a speaker's adaptive responses and the ability of their feedforward system to learn new motor commands. Some studies have previously investigated adaptive responses in Dutch speakers (e.g., van Brenk & Terband, 2020) but to our knowledge none have tested Dutch speakers' reflexive responses nor used the target word combinations of this study. In line with prior studies on other languages (predominantly English), we expected participants to respond to the formant perturbation by shifting their own formant production in the opposite direction (e.g., Houde & Jordan, 2002).

Methods In total 59 speakers (34 female; age range 22 to 59) completed a sustained perturbations task, an unexpected perturbations task, or both. Speakers were instructed to read several high frequency monosyllabic Dutch words containing the vowel /ɛ/ while wearing headphones. During production, vowel formants were shifted using Audapter software (Cai et al., 2008). During the sustained perturbations task, formants were gradually shifted with each successive production, until reaching the maximum perturbation of 20% decrease in F1 and 15% increase in F2 (resulting in a vowel closer to /ɪ/). There were altogether 114 trials, of which 48 were produced at the maximum perturbation (i.e., the STAY phase). During the unexpected perturbations task, the perturbation was smaller than in the sustained perturbations task, namely a 10% decrease in F1 and 5% increase in F2 for an upward trial (resulting in /ɪ/), and a 10% increase in F1 and 5% decrease in F2 for a downward trial (resulting in /ɑ/). There were altogether 120 trials: two thirds of the trials were randomly perturbed either upwards or downwards, the rest were control catch trials. Most shifts in both experimental tasks resulted in a real Dutch word being produced. For example, the target word /bɛt/ ('bed') could be shifted towards /bɪb/ ('bridle bit') or /bɛt/ ('bath'). After completing the experiment, the participants filled out a questionnaire, asking them whether they had noticed the perturbation.

Results & discussion We used generalized additive mixed models to analyse participants' reflexive and adaptive responses. For both tasks, participants showed expected responses: when vowels were perturbed upwards, the speakers shifted their productions downwards and vice versa (see Figure 1 for plotted response to the sustained perturbations task). This adaptation occurred even though, according to their questionnaire responses, few speakers consciously noticed the perturbation.



Change in F1 (left) and F2 (right) in response to a sustained perturbation. Dotted vertical lines denote the four phases of the experiment (START, RAMP, STAY, END).

References

Cai, S., Boucek, M., Ghosh, S.S., Guenther, F.H., & Perkell, J.S. (2008). A system for online dynamic perturbation of formant frequencies and results from perturbation of the Mandarin triphthong /iau/. In *Proceedings of the 8th International Seminar on Speech Production*, Strasbourg, France, Dec. 8 - 12, 2008, pp. 65-68.

Houde, J. F., & Jordan, M. I. (2002). Sensorimotor Adaptation of Speech I: Compensation and Adaptation. *Journal of Speech, Language, and Hearing Research, 45*(2), p. 295-310. https://doi.org/10.1044/1092-4388(2002/023)

van Brenk, F., & Terband, H. (2020). Compensatory and adaptive responses to real-time formant shifts in adults and children. *The Journal of the Acoustical Society of America, 147*(4). https://doi.org/10.1121/10.0001018

LESIONS CAUSING ACQUIRED NEUROGENIC STUTTERING CONNECT TO COMMON BRAIN AREAS

Catherine Theys[1], Elina Jaakkola[2], Tracy Melzer[3], Michael D. Fox[4,5], Juho Joutsa[2,6]
[1] *University of Canterbury, Christchurch, New Zealand*
[2] *University of Turku, Turku, Finland*
[3] *New Zealand Brain Research Institute, Christchurch, New Zealand*
[4] *Brigham and Women's Hospital, Boston, MA, USA*
[5] *Harvard Medical School, Boston, MA, USA*
[6] *Turku University Hospital, Turku, Finland*

Stuttering can be developmental in origin, but it can also develop secondary to acquired neurological conditions. Imaging studies have linked stuttering with differences in brain structure and function, but its neural origin remains unknown (Etchell et al., 2018). As investigating the relationship between focal brain lesions and associated symptoms can reveal causal relationships (Damasio & Damasio, 1989), we aimed to identify brain areas causally linked with stuttering by investigating two independent datasets of participants with stroke-induced neurogenic stuttering.

The first dataset was created following a systematic literature review and consisted of 20 cases with lesion-induced acquired neurogenic stuttering (7 females, 16-77 years). The second dataset consisted of 20 prospectively identified individuals with acquired neurogenic stuttering following stroke (7 females, 45-87 years). This dataset also included a control group of 17 stroke patients without stuttering but matched for presence of other speech, language and cognitive problems (6 females, 50-83 years) (Theys et al., 2013). In the two neurogenic stuttering datasets, stuttering was associated with heterogeneous lesion locations. To identify if these heterogeneous locations were linked to common brain areas, we applied a relatively new technique, lesion network mapping, to both datasets (Fox, 2018).

As a first step, for every participant whole-brain lesion connectivity maps were created based on their focal lesion location, and thresholded to identify brain regions significantly connected with the lesion location (lesion networks). Next, lesion network maps of all participants were overlaid to identify areas of overlap between their lesion networks (Boes et al., 2015; Darby et al., 2018). For the first dataset, overlapping network regions included the bilateral basal ganglia, amygdala, thalamus, insula and Heschl's gyrus, left-sided inferior frontal gyrus and right-sided cerebellum. In the more homogeneous prospective dataset, lesion network mapping results converged with those from the first dataset.

To assess if the identified brain areas were specific to stuttering, we compared lesion connectivity of the 20 prospective participants with neurogenic stuttering with those from the matching control group of 17 participants without acquired stuttering, Voxelwise T-tests between the stuttering and control group, thresholded at FWE corrected $P<0.05$ were conducted within the areas identified in the first dataset. Connections specifically associated with stuttering onset were in the lentiform nuclei and amygdala.

These findings show that focal lesions causing stuttering are linked to a network of brain areas consistent with the speech production network (Bohland & Guenther, 2006), and extending into the amygdala. However, comparison of lesion connectivity networks between participants who developed stuttering and controls allowed for identification of more specific brain areas, including the bilateral lentiform nuclei.

References

Etchell, A. C., Civier, O., Ballard, K. J. & Sowman, P. F. (2018). A systematic literature review of neuroimaging research on developmental stuttering between 1995 and 2016. *Journal of Fluency Disorders 55*, 6-45. https://doi.org/10.1016/j.jfludis.2017.03.007

Damasio, H. & Damasio, A. R. (1989). *Lesion analysis in neuropsychology.* Oxford University Press.

Theys, C., De Nil, L., Thijs, V., van Wieringen, A. & Sunaert, S. (2013). A crucial role for the cortico-striato-cortical loop in the pathogenesis of stroke-related neurogenic stuttering. *Human Brain Mapping 34*, 2103-2112. https://doi.org/10.1002/hbm.22052

Fox, M. D. (2018). Mapping Symptoms to Brain Networks with the Human Connectome. *The New England Journal of Medicine 379*, 2237-2245. https://doi.org/10.1056/NEJMra1706158

Boes, A. D. et al. (2015). Network localization of neurological symptoms from focal brain lesions. Brain 138, 3061-3075. https://doi.org/10.1093/brain/awv228

Darby, R. R., Horn, A., Cushman, F. & Fox, M. D. (2018). Lesion network localization of criminal behavior. P*roceedings of the National Academy of Sciences - PNAS 115*, 601-606. https://doi.org/10.1073/pnas.1706587115

Bohland, W. B. & Guenther, F. H. (2006). An fMRI investigation of syllable sequence production. *NeuroImage 32*, 821-841.

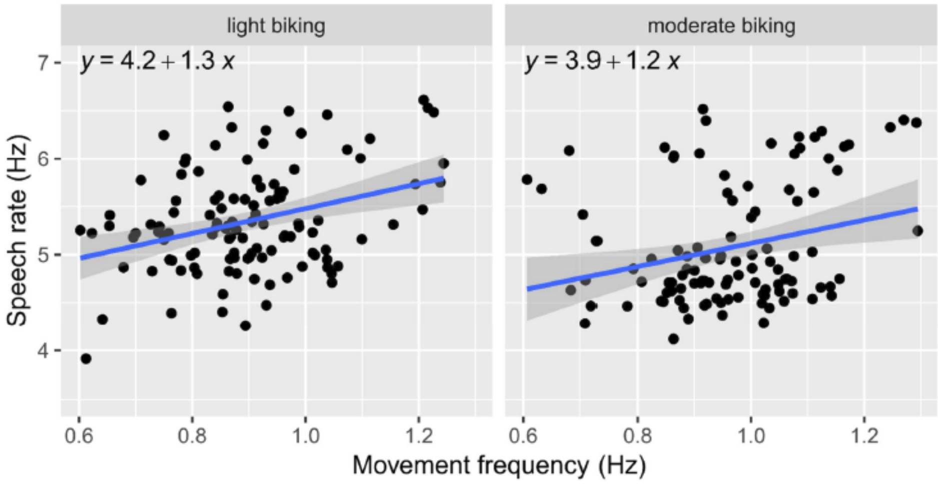
ON THE RELATION BETWEEN MOTION RATE AND SPEECH RATE WITH INCREASING PHYSICAL WORKLOAD

Heather Weston[1,2], Wim Pouw[3,4], Susanne Fuchs[1]
[1]*Leibniz-Centre General Linguistics (ZAS), Berlin, Germany*
[2]*Humboldt-Universität zu Berlin, Berlin, Germany*
[3]*Radboud University, Nijmegen, Netherlands*
[4]*Max Planck Institute for Psycholinguistics, Netherlands*

Introduction Evidence for temporal coupling between spoken language and manual gestures has been found in different domains, including language acquisition (Iverson et al., 2007) and laboratory experiments examining speech during finger tapping (e.g., Parrell et al., 2014). However, it is not clear whether these findings generalize to adults engaging in real-world activities, or if coupling exists with whole-body movements. In a study on spontaneous speech produced while cycling, Fuchs et al. (2015) found a small increase in speech rate with increased workload, suggesting a potential influence of pedaling rate on speech. But spontaneous speech contains hesitation phenomena that slow speech rate. It is also possible that movements are not in a temporal relation with syllables, but rather stressed syllables, as found by Rochet-Capellan et al. (2008). The present study investigates the production of connected speech during everyday physical activity, asking 1) whether speech rate changes at different intensities of rhythmic aerobic exercise and 2) whether there is a correlation between speech rate and movement rate. It extends previous work by using a reading task to minimize the hesitancy phenomena characteristic of spontaneous speech and recording pedal rate, as well as by looking at phonetic aspects of speech that span multiple time scales (Tilsen & Arvaniti, 2013).

Method In a within-subjects design, 48 female German speakers (age: 19–34; $\bar{x} = 23.6$) read a 126-word passage three times in each of three conditions: sitting still (CONTROL), and during LIGHT- and MODERATE-intensity cycling on a stationary bicycle. Exercise intensity – the physical effort required to perform a task – was defined as a percentage of maximal heart rate and calculated using the Karvonen formula, a standard method that incorporates an individual's age and resting pulse to reflect physical fitness (Tanaka et al., 2002). The physical effort required in the experimental conditions is thus comparable across participants. Pedaling motions were recorded using an OptiTrack motion capture system and pedaling rate was calculated for each reading trial using a peak detection algorithm. Speech rate was similarly calculated using a phonetic-syllable proxy, by detecting the interval between local maxima of a smoothed amplitude envelope (excluding intervals that spanned pauses) (He & Dellwo, 2017).

Results & Discussion There were three preliminary findings at group level. First, there was an effect of exercise intensity on pedaling rate: as intensity – physical workload – increased, so did the rate of movement. Second, there was an effect of exercise intensity on speech rate, which decreased in the moderate condition, though there was some interspeaker variation in this condition. Third, an overall correlation was found between pedaling rate and speech rate (Figure 1): speakers who pedaled faster tended to speak faster. These findings suggest that whole-body situational influences may also have implications of for speech production. The full paper will investigate whether there are stable temporal relation between stressed syllables and movement rate and whether different speakers show distinct timing relations.



Correlation between speech rate and movement rate during light- and moderate-intensity cycling

References

Fuchs, S., Reichel, U. D., & Rochet-Capellan, R. (2015). Changes in speech and breathing rate while speaking and biking. *18th International Congress of Phonetic Sciences*, ICPhS 2015, Glasgow, UK, August 10-14, 2015.

He, L. & Dellwo, V. (2017). Amplitude envelope kinematics of speech: parameter extraction and applications. *The Journal of the Acoustical Society of America, 141*, 3582–3582.

Iverson, J., Hall, A., Nickel, L., & Wozniak, R. (2007). The relationship between reduplicated babble onset and laterality biases in infant rhythmic arm movements. *Brain and Language, 101*, 198–207.

Parrell, B., Goldstein, L., Lee, S., & Byrd, D. (2014). Spatiotemporal coupling between speech and manual motor actions. *Journal of Phonetics, 42*, 1–11.

Rochet-Capellan, A., Laboissière, R., Galván, A., & Schwartz, J.-L. (2008). The speech focus position effect on jaw–finger coordination in a pointing task. *Journal of Speech, Language, and Hearing Research, 51*, 1507–1521.

Tanaka, H., Monahan, K.D., & Seals, D.R. (2001). Age-predicted maximal heart rate revisited. *Journal of the American College of Cardiology, 37*(1), 153–156.

Tilsen, S., & Arvaniti, A. (2013). Speech rhythm analysis with decomposition of amplitude envelope: characterizing rhythmic patterns within and across languages. *The Journal of the Acoustical Society of America, 134*(1), 628–639.

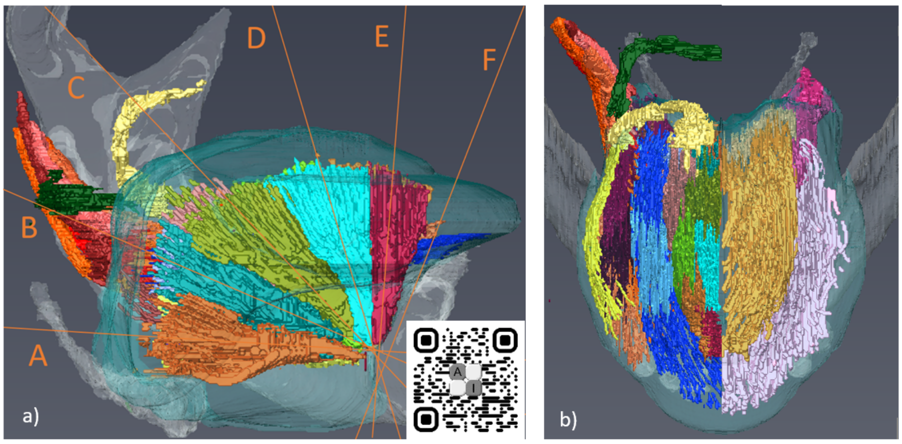
THE COMPARTMENTAL TONGUE: EVIDENCE FOR INDEPENDENT NEUROMUSCULAR CONTROL OF SIX SECTORS OF THE OROPHARYNGEAL CAVITY.

Alan A. Wrench
Queen Margaret University, Edinburgh, United Kingdom

Sleep apnoea researchers divide the genioglossus into two separately stimulable horizontal and oblique neuromuscular compartments (NMCs). Biomechanical modellers find this is insufficient to control tongue shapes for speech and propose at least three functional compartments (Honda et al., 2013). How many independently controlled compartments are there? This paper reviews evidence to determine the answer to that question. It collates observations by anatomists and histologists (Barnwell, 1977; Miyawaki, 1975; Mu & Sanders, 1999; Mu & Sanders, 2010; Saito & Itoh, 2003; Sakamoto, 2017), to try to identify and enumerate NMCs of each tongue muscle. Supporting evidence is provided by a novel detailed 3D atlas constructed from fibre-level segmentation of the Visible Human Female (Ackerman, 1998) (Figure 1).

Studies show the hyoglossus consists of at least three compartments (Sakamoto, 2017) two of which (chondroglossus and ceratoglossus) merge with the superior longitudinal muscle (Figure 1b). The styloglossus comprises four to six compartments (Barnwell, 1977; Sakamoto, 2017) leading from the styloid process to the base of the tongue. One or two of these compartments enter the root (Honda et al., 2013; Saito & Itoh, 2007). The anterior part of the styloglossus extending along the lateral margin from the root to the tip is separately innervated having one or two lateral longitudinal NMCs (Mu & Sanders, 2010; Saito & Itoh, 2007). The inferior longitudinal muscle has two parts (Mu & Sanders, 2010), a longitudinal compartment which courses to the tip and an oblique compartment which terminates in the dorsum near the anterior genioglossus(Figure 1c). Fibre-level segmentation presented here and innervation evidence (Mu & Sanders, 1999) indicates each genioglossus muscle has ten compartments arranged in two rows from root to blade (Figure 1a A-E & b). They divide the tongue body into five sectors[3]. The superior longitudinal muscle consists of short in-series fibres covering the dorsum of the tongue (Mu & Sanders, 2010) and capable of localised contraction. Transversus and verticalis fibres lie in alternating laminae with separate innervation (Mu & Sanders, 2010). It is likely that groups of laminae form NMCs (Mu & Sanders, 2010).

This fine neuromuscular structure, rather than complicating motor control, may simplify it. Compartments of genioglossus, verticalis and transversus muscles may work together in localised groups to control oropharyngeal cavity size in six sectors (Figure 1a A-F). A biomechanical model demonstrates this sector-based functionality (Wrench & Balch, 2015). Musculotopic organisation of the hypoglossal nucleus places motoneurons for these three muscles in close proximity (McClung & Goldbert, 2000). The genioglossus NMCs in each of its five sectors contract to form a localised medial lingual groove. The verticalis NMC in the corresponding sector widens the groove. The transversus muscle acts locally in opposition to both the genioglossus and verticalis, stretching them and compressing them width-wise in the transverse axis to eliminate the groove. In the longitudinal axis the oblique inferior longitudinal NMC (Figure 1b) constrains rostro-caudal expansion of the tongue body, constricting the velar region of the oropharyngeal cavity while not restricting movement and shape of the tongue blade/tip. The blade is independently controlled. A 6th verticalis NMC flattens it and an opposing transversus NMC narrows it. Inferior, lateral and superior longitudinal NMCs independently control ventroflexion, retroflexion and left-right position of the tip. In conclusion, the most likely number of genioglossus NMCs is ten, arranged in medial/lateral pairs in five sectors. Each pair controlling degree of constriction in that sector. Sequential activation of these sectors may be used to propel the bolus during swallowing (Yeung et al., 2022).



a) Five medial NMCs of genioglossus A-E; blade/tip F; Glossopharyngeal part of superior constrictor (green), Styloglossus NMCs (red/orange), Palatoglossus (yellow). b) Transverse view additionally showing inferior longitudinal (blue) and inferior oblique (pale blue), hyoglossus (maroon=basioglossus NMC; cerise=ceratoglossus), superior longitudinal (pale pink and gold), chondroglossus (beige).

References

Ackerman, M.J. (1998). The visible human project. *Proceedings of the IEEE, 86*(3), p. 504-511.

Barnwell, Y.M. (1977). The morphology of musculus styloglossus in fifteen-week human fetuses. *International Journal of Orofacial Myology and Myofunctional Therapy, 3*(2), p. 8-46.

Honda, K., et al. (2013). Anatomical considerations on the extrinsic tongue muscles for articulatory modeling. In *Proceedings of Meetings on Acoustics ICA2013*.

McClung, J.R. & S.J. Goldberg. (2000). Functional anatomy of the hypoglossal innervated muscles of the rat tongue: a model for elongation and protrusion of the mammalian tongue. *The Anatomical Record, 260*(4), p. 378-386.

Miyawaki, K. (1975). A preliminary report on the electromyographic study of the activity of lingual muscles. *Annual Bulletin RILP, 9*, p. 91-106.

Mu, L. & I. Sanders. (2010). Human tongue neuroanatomy: Nerve supply and motor endplates. *Clinical anatomy (New York, N.Y.), 23*(7), p. 777-791.

Mu, L. & I. Sanders. (1999). Neuromuscular organization of the canine tongue. *The Anatomical Record: An Official Publication of the American Association of Anatomists, 256*(4), p. 412-424.

Saito, H. & I. Itoh. (2007). The three-dimensional architecture of the human styloglossus especially its posterior muscle bundles. *Annals of Anatomy, 189*(3), p. 261-267.

Saito, H. & Itoh, I. (2003). Three-dimensional architecture of the intrinsic tongue muscles, particularly the longitudinal muscle, by the chemical-maceration method. *Anatomical science international, 78*(3), p. 168-176.

Sakamoto, Y. (2017). Configuration of the extrinsic muscles of the tongue and their spatial interrelationships. *Surgical and Radiologic Anatomy, 39*(5), p. 497-506.

Wrench, A.A. & Balch, P. (2015). Towards a 3D tongue model for parameterising ultrasound data. *Proceedings of the 18th ICPhS*, Glasgow.

Yeung, J., et al. (2022). Task-dependent neural control of regions within human genioglossus. *Journal of applied physiology, 132*(2), p. 527-540.

MULTI-PARAMETRIC ANALYSIS OF THE RESPIRATORY ACTIVITY IN SPEECH PRODUCTION

Shi Yu, Didier Demolin
CNRS, Grenoble, France
Sorbonne-Nouvelle, Paris, France

Background Speech production requires fine-tuned coordination of articulators in interaction with the respiratory support. To maintain subglottal pressure (Ps), the driving force of phonation, a complex mechanism acts, involving intrinsic properties of the lungs, such as their volume (capacity) and elastic recoil, and extrinsic support, comprising various inspiratory and expiratory muscles. At the same time, to communicate linguistic content, multiple acoustic-perceptual parameters including intensity, stress, and intonation, need to be modulated in accordance with the respiratory support, which requires various active and passive regulatory mechanisms. The objectives of this study are to: (1) Verify previous results on respiratory activity in speech production with up-to-date instrumentation and method. (2) Examine respiratory support activities associated with prosodic variations such as that in expressions of different modalities, and stress.

Method A multi-parametric setup including acoustic signal, EGG (electroglottography), aerodynamics (intra-oral pressure (Pio) and oral airflow), EMG (activities of scalene, external intercostal, abdominal, and abdominal external oblique muscles), kinematics (jaw, thorax, and abdomen displacement), and EEG (electroencephalography), was built to record synchronously activities related to speech production and respiration. In addition, vital capacity (VC) was measured using a spirometer. Estimation of lung volume change during speech production was derived from recorded measures. Linguistic materials consist of repetitions of the syllable [pa] with sustained speech. Ps variations can be estimated from Pio when full closure of the vocal tract is achieved. One French and one Mandarin Chinese subjects were investigated with the same setup, each condition is recorded with 40 repetitions.

Results The control of subglottal pressure (Ps) during speech production is supported by various combinations respiratory effectors, we identified several markers that indicate the interplay of respiratory activities and variations of linguistic content. During sustained speech, abdominal muscles are activated when 80%-90% of the vital capacity of the lungs is used, which maintain a sufficient Ps level for phonation but reveal the instability of coordination in such situation. In experiments with modality expressions (statement, question, exclamation/emphasis), subject uses typically 30% of the VC to complete the sentence content, and supplementary respiratory activity is required to realise prominences in prosodic variations such as the emphasis or final rising intonation in questions.

Discussion The present study enhances the Edinburgh study of speech breathing by applying up-to-date instrumentation and analysis techniques. The coordination of respiratory muscles suggests that the respiratory activities may contribute to linguistic content of speech production in a finer manner than solely support and maintain a relatively stable Ps level.

COMMUNICATION IS MOVEMENT: THE TEMPORAL COORDINATION OF PITCH, ARTICULATOR MOVEMENTS, AND MANUAL GESTURES IN AUTISTIC VERSUS NEUROTYPICAL ADULTS

Claudia I. Abbiati, Shelley L. Velleman, Kim R. Bauerly
University of Vermont, Burlington, Vermont, United States of America

Autism is a complex neurodevelopmental condition identified in one of 44 children (CDC, 2022). While autism is characterized by persistent impairments in social interaction and by restricted and repetitive behaviors (APA, 2013), movement disturbances are increasingly recognized as additional core symptoms (Torred & Donnellan, 2015; Fournier et al., 2010). Movement impairments are implicated in the development of social communication skills (Hirata et al., 2014; Dziuk et al., 2007); they are linked to concurrent and prospective social communication deficits experienced by autistics (McCleery et al., 2013; Moody et al., 2017). The temporal coordination of prosody and manual gestures impacts the development and use of effective communication skills (Ejiri, 1998; Iverson & Fagan, 2004; Wagner et al., 2014). For instance, gestures are timed with prominent aspects of prosody: emphatic speech stress (Wagner et al., 2014; Shattuck-Hufnagel & Prieto, 2019), pitch accents (Loehr, 2012; Mendoza-Denton & Jannedy, 2011), and peaks of fundamental frequency (Leonard & Cummins, 2011). Some studies have shown that the temporal coordination between prosody and manual gestures is "weaker" in autistics versus neurotypicals (de Marchena & Eigsti, 2010), but such findings are based on subjective coding and do not consider articulator movement coordination. The purpose of this study is to use motion tracking technology to determine whether the temporal coordination among pitch, articulator movements, and manual gestures differs in age- and sex-matched autistic versus neurotypical adults in a structured speaking-gesture task. Participants were monolingual English-speaking young adults (18-35 years). To date, twenty autistic and nineteen neurotypical participants have been tested. Participants produced ten isolated repetitions of four different declarative sentences varying in length and complexity (i.e., short-simple, long-complex). Adhering to previous research, sentences were loaded with sounds requiring lip movement (MacPherson & Smith, 2013). Northern Digital's electromagnetic articulography Wave system (Northern Digital Inc.) measured lip movements. Participants were fitted with a 3X3mm sensor on the midline borders of their upper and lower lips using Glustitch PeriAcryl90 (Gluestitch, Inc.) and a head-mounted 5DOF sensor to track head motion. A labile microphone captured acoustic samples of speech. A 19-camera Vicon motion tracking system measured manual gestures. For this measurement, reflective markers were placed along the dominant upper extremity (e.g., shoulder, elbow, index finger). Participants were instructed to point to a picture and contrastively stress a target word in each sentence to correct misspoken sentences. Data analysis is currently underway.

Three-dimensional lip movements, corrected for head movements and smoothed using a 15Hz low-pass filter, are being analyzed for temporal aspects of pitch and articulator movements using SMASH (Green et al., 2013). We are analyzing the production of the first syllable in the stressed target word ("bo" [bɑ] from "Bobby") for each production of a sentence, specifically time to peak displacement of the lips (LA) and time to peak fundamental frequency (F0). Nexus software (Vicon) is being used to examine time to peak extension of the index finger (IF) while pointing. Time is measured in milliseconds. Differences in time will be computed between F0-LA, F0-IF, and LA-IF between groups for each sentence type.

Preliminary results show that autistic young adults have a less coordinated pattern (i.e., greater differences in timing) between F0, LA, and IF. Findings will provide a strong foundation for future clinically meaningful motor research.

References

American Psychiatric Association. (2013). Diagnostic and statistical manual of mental disorders. 5th ed. Arlington, VA: American Psychiatric Association.

Data and statistics on autism spectrum disorder. (2022) Center for Disease Control and Prevention. Updated December 3, 2021. Accessed March 7, 2022. `https://dx.doi.org/10.15585/mmwr.ss7011a1`

de Marchena A, Eigsti IM. (2010). Conversational gestures in autism spectrum disorders: Asynchrony but not decreased frequency. *Autism Research, 3*(6):311-322.

Dziuk MA, Gidley Larson JC, Apostu A, Mahone EM, Denckla MB, Mostofsky SH. (2007). Dyspraxia in autism: Association with motor, social, and communicative deficits. *Developmental Medicine and Child Neurology, 49*:734-739.

Ejiri K. (1998). Relationship between rhythmic behavior and canonical babbling in infant vocal development. *Phonetica, 55*:226–237.

Fournier KA, Hass CJ, Naik SK, Lodha N, Cauraugh JH. (2010). Motor coordination in autism spectrum disorders: A synthesis and meta-analysis. *J Autism Dev Disord., 40*:1227-1240.

Green JR, Wang J, Wilson DL. SMASH: A tool for articulatory data processing and analysis. In: Bimbot F, Cerisara C, Fougeron C, Gravier G, Lamel L, Pellegrino F, Perrier P eds. (2013). *Interspeech. International Speech Communication Association,*, 1331–1335.

Hirata S, Okuzumi H, Kitajima Y, Hosobuchi T, Nakai A, Kokubun M. (2014). Relationship between motor skill and social impairment in children with autism spectrum disorders. *International Journal of Developmental Disabilities, 61*:251-256.

Iverson JM, Fagan MK. (2004). Infant vocal–motor coordination: Precursor to the gesture–speech system? *Child Development, 75*, 1053–1066.

Leonard T, Cummins F. (2011). The temporal relation between beat gestures and speech. *Language and Cognitive Processes, 26*:1457–1471.

Loehr DP. Temporal, structural, and pragmatic synchrony between intonation and gesture. Laboratory Phonology. 2012;3:71–89.

MacPherson MK, Smith A. (2013). Influences of sentence length and syntactic complexity on the speech motor control of children who stutter. *Journal of Speech, Language, and Hearing Research, 56*(1):89-102.

McCleery JP, Elliot NA, Sampanis DS, Stefanidou CA. (2013). Motor development and motor resonance difficulties in autism: Relevance to early intervention for language and communication skills. *Frontiers in Integrative Neuroscience*, 7:30.

Mendoza-Denton N, Jannedy S. (2011). Semiotic layering through gesture and intonation: A case study of complementary and supplementary multimodality in political speech. *Journal of English Linguistics, 39*:265–299.

Mody M, Shui AM, Nowinski LA, Golas SB, Ferrone C, O'Rourke AO, McDougle CJ. (2017). Communication deficits and the motor system: Exploring patterns of associations in autism spectrum disorders (ASD). *Journal of Autism and Developmental Disorders, 47*:155-162.

Shattuck-Hufnagel S, Prieto P. Dimensionalizing co-speech gestures. Proceedings of the

International Congress of Phonetic Sciences. 2019.

Torres EB, Donnellan AM. (2015). *Autism: The movement perspective.* (n.p.): Frontiers Media SA.

Wagner P, Malisz Z, Kopp S. (2014). Gesture and speech in interaction: An overview. *Speech Communication, 57*:209–232.

LONGITUDINAL CHANGES TO SPEECH IN PARKINSON'S DISEASE

Defne Abur, Amanda Mount, Cara Stepp
Boston University, Boston, Massachusetts, United States of America

Purpose The manifestation of speech impairments in Parkinson's disease (PD) is highly variable and the years since PD diagnosis are not consistently associated with speech decline (Holmes et al., 2000; Miller et al., 2011). Specifically, measures of articulation show decline as PD progresses (Skodda et al., 2012; Skodda et al., 2011; Ho et al., 1999), but measures of voice do not show associations with the time spent living with PD (Holmes et al., 2000; Skodda et al., 2009). It is not clear if these findings are due to different symptom progression by speech subsystem or high individual variability in PD. Thus, this proposal examines how longitudinal changes in speech from two time points relate to disease progression within the same person with PD. We hypothesized that changes in acoustic measures of articulation would better predict the time between sessions compared to changes in acoustic measures of voice in PD.

Methods Twenty-six persons with PD (10 female, 16 male) on medication participated at two time points ranging 3 – 73 months apart. Speakers completed four speech tasks: (1) prolonged corner vowels; (2) five words with vocalic nuclei requiring large changes to the vocal tract shape and unvoiced consonants; (3) The Rainbow Passage (Fairbanks, 1960); (4) a 1-minute conversational speech sample. Ten acoustic measures of voice and articulation were calculated (see Table 1).

Task	Voice	Justification	Articulation	Justification
1	harmonics-to-noise ratio (HNR)	HNR an objective measure of breathy voice (Shrivastav & Sapienza, 2003), which is a symptom of PD (Broadfoot et al., 2019).	vowel space area (VSA), vowel articulation index (VAI)	VSA relates to speech intelligibility and VAI is sensitive to PD disease stage (Skodda et al., 2012; Weismer et al., 2001).
2	relative fundamental frequency (RFF)	RFF is an estimate of laryngeal tension (Stepp et al., 2011; Eadie & Stepp, 2013; Lien et al., 2015) and sensitive to PD disease stage (Stepp, 2013).	2nd formant ($F2$) slope	F2 slopes are associated with speech intelligibility (Kim et al., 2011; Weismer et al., 1992).
3, 4	Mean and standard deviation of fundamental frequency (f_0 and f_0SD), cepstral peak prominence (CPP)	Mean f_0 differs in PD over time (Skodda et al., 2009) and can impact speech perception (Holt et al., 2001; Whalen et al., 1993; Faulkner & Rosen, 1999). f_0SD is reduced in PD (Holmes et al., 2000; Bowen et al., 2013; Flint et al., 1992) and sensitive to disease stage (Harel et al., 2004). CPP is a correlate of speech severity (Goberman & Elmer, 2005).	articulatory-acoustic vowel space (AAVS), articulation rate (AR)	AAVS is sensitive to changes in speech clarity in PD (Whitfield & Goberman, 2013). AR reflects speech rate, which is affected in PD (Goberman & Elmer, 2005; Ludlow et al., 1987; Ackerman et al., 1997).

Results and Discussion A stepwise linear regression was performed on the time between sessions using the change in all acoustic measures for the 26 speakers. Two measures statistically predicted ($p < 0.05$) the time between sessions (VAI: F = 4.69, p = 0.04 and F_2 slope: F = 5.47, p = 0.02). Extending prior findings of articulatory decline between early-stage and later-stage PD, the current work identifies acoustic measures of articulation (i.e., VAI and F_2 slope) that are sensitive to the time progression of PD at an individual-level. These results also support the notion that the decline in vocal features of speech does not relate to PD progression.

References

Ackermann, H., J. Konczak, & I. Hertrich. (1997). The temporal control of repetitive articulatory movements in Parkinson's disease. *Brain and language, 56*(2): p. 312-319.

Bowen, L.K., et al. (2013). Effects of Parkinson's disease on fundamental frequency variability in running speech. *Journal of medical speech-language pathology, 21*(3): p. 235.

Broadfoot, C., et al. (2019). Based Updates in Swallowing and Communication Dysfunction in Parkinson Disease: Implications for Evaluation and Management. *Perspectives of the ASHA Special Interest Groups, 4*(5): p. 825-841.

Eadie, T.L. & C.E. Stepp (2013). Acoustic correlate of vocal effort in spasmodic dysphonia. *Annals of Otology, Rhinology & Laryngology, 122*(3): p. 169-176.

Faulkner, A. & S. Rosen. (1999). Contributions of temporal encodings of voicing, voicelessness, fundamental frequency, and amplitude variation to audio-visual and auditory speech perception. *The Journal of the Acoustical Society of America, 106*(4): p. 2063-2073.

Flint, A.J., et al. (1992). Acoustic analysis in the differentiation of Parkinson's disease and major depression. *Journal of Psycholinguistic Research, 21*(5): p. 383-399.

Goberman, A.M. & Elmer. (2005). Acoustic analysis of clear versus conversational speech in individuals with Parkinson disease. *J Commun Disord, 38*(3): p. 215-230.

Harel, B., M. Cannizzaro, & P.J. Snyder. (2004). Variability in fundamental frequency during speech in prodromal and incipient Parkinson's disease: A longitudinal case study. *Brain and cognition, 56*(1): p. 24-29.

Heman-Ackah, Y.D., D.D. Michael, & G.S. Goding Jr. (2002). The relationship between cepstral peak prominence and selected parameters of dysphonia. *Journal of Voice, 16*(1): p. 20-27.

Ho, A.K., et al. (1999). Speech impairment in a large sample of patients with Parkinson's disease. *Behavioural Neurology, 11*(3): p. 131-137.

Holmes, et al. (2000). Voice characteristics in the progression of Parkinson's disease. *International Journal of Language & Communication Disorders, 35*(3): p. 407-418.

Holt, L.L., A.J. Lotto, & K.R. Kluender. (2001). Influence of fundamental frequency on stop-consonant voicing perception: A case of learned covariation or auditory enhancement? *The Journal of the Acoustical Society of America, 109*(2): p. 764-774.

Kim, Y., R.D. Kent, & G. Weismer. (2011). An acoustic study of the relationships among neurologic disease, dysarthria type, and severity of dysarthria. *Journal of Speech, Language, and Hearing Research.*

Lien, Y.-A.S., et al. (2015). Voice relative fundamental frequency via neck-skin acceleration in individuals with voice disorders. *Journal of Speech, Language, and Hearing Research, 58*(5): p. 1482-1487.

Ludlow, C.L., N.P. Connor, & C.J. Bassich. (1987). Speech timing in Parkinson's and Huntington's disease. *Brain and Language, 32*(2): p. 195-214.

Miller, N., et al. (2011). Changing perceptions of self as a communicator in Parkinson's disease: A longitudinal follow-up study. *Disability and rehabilitation, 33*(3): p. 204-210.

Shrivastav, R. & Sapienza, C.M. (2003). Objective measures of breathy voice quality obtained using an auditory model. *The Journal of the Acoustical Society of America, 114*(4): p. 2217-2224.

Skodda, S., A. Flasskamp, & U. Schlegel (2011). Instability of syllable repetition as a marker of disease progression in Parkinson's disease: a longitudinal study. *Movement disorders, 26*(1): p. 59-64.

Skodda, S., H. Rinsche, & U. Schlegel (2009). Progression of dysprosody in Parkinson's disease over time—a longitudinal study. *Movement disorders, 24*(5): p. 716-722.

Skodda, S., W. Grönheit, & U. Schlegel. (2012). Impairment of vowel articulation as a possible marker of disease progression in Parkinson's disease. *Public Library of Science ONE, 7*(2).

Stepp, C.E., et al. (2011). Effects of voice therapy on relative fundamental frequency during voicing offset and onset in patients with vocal hyperfunction. *Journal of Speech, Language, and Hearing Research.*

Stepp, C.E., (2013). Relative fundamental frequency during vocal onset and offset in older speakers with and without Parkinson's disease. *The Journal of the Acoustical Society of America, 133*(3): p. 1637-1643.

Weismer, G., et al. (2001). Acoustic and intelligibility characteristics of sentence production in neurogenic speech disorders. *Folia Phoniatrica et Logopaedica, 53*(1): p. 1-18.

Weismer, G., et al. (1992). Formant trajectory characteristics of males with amyotrophic lateral sclerosis. *The Journal of the Acoustical Society of America, 91*(2): p. 1085-1098.

Whalen, D.H., et al. (1993). 0 gives voicing information even with unambiguous voice onset times. *The Journal of the Acoustical Society of America, 93*(4): p. 2152-2159.

Whitfield, J.A. & A.M. Goberman. (2014). Articulatory–acoustic vowel space: Application to clear speech in individuals with Parkinson's disease. *Journal of communication disorders, 51*: p. 19-28.

INVESTIGATING ATTENTIONAL FOCUS AS A MEDIATING FACTOR TO THE CHANGES IN SPEECH-MOTOR CONTROL DURING SOCIAL STRESS IN PEOPLE WHO STUTTER

Kim R. Bauerly
University of Vermont, Burlington, Vermont, United States of America

Evidence suggests social stress has a stronger destabilizing effect on speech-motor control in adults who stutter (AWS) compared to adults who do not stutter (ANS)(Van Lieshout et al., 2014; Jackson et al., 2016). However, the nature of this interaction is not fully understood. One reason may be due to effects from mediating factors such as attention or consciousness in control which have gone largely ignored. Consciously controlling movement (Wulf et al., 2001) that is normally carried out automatically may lead to reduced flexibility (Lohse et al., 2010) and adaptability and has been shown to disrupt speech performance (Freedman e tal., 2007). Evidence shows highly anxious speakers shift away from the listener when under social evaluative stress (SET)(Chen et al., 2002) and considering high levels of social anxiety are also reported in AWS (Craig & Tran, 2014), it can be hypothesized that AWS also shift away from listener's reactions when speaking under SET. The effects of these attentional shifts on behavioral, autonomic and articulatory parameters have not been investigated.

Some of our earlier work (Bauerly, 2021) provided evidence that the attentional biases in AWS when speaking under SET is similar to what is reported in high anxious ANS6. This presentation will provide results from a study investigating the effects of attentional biases on speech motor control in AWS during SET. For this study, participants repeated sentences under low SET (i.e. one audience member) vs. high (i.e. nine audience members) SET conditions while the duration and variability of lip aperture assessed speech motor control. Under low SET, AWS showed a significant decrease in lip aperture variability when speaking under the external-cued versus no-cue condition. Under high SET conditions (i.e. audience), a significant Group x Condition interaction was found. Post-hoc comparisons indicated that the significant interaction was due to decreases in the AWS' lip aperture variability when cued to speak under external cued conditions (i.e. focus on the audience) compared to no cue, Audience conditions. These findings lend important insights into the relationship between attentional focus and speech motor control in AWS when speaking under social stress.

This presentation also includes preliminary data investigating the effects of attention on speech motor control in high vs. low socially anxious AWS and ANS using eye gaze to measure attentional focus and a virtual reality environment (withVR; Walkom, 2022) to elicit social stress. In this study, participants repeated sentences and produced a monologue to an audience (n=9) in a VR cafe setting while eye gaze, autonomic activity and articulatory movement was measured. Preliminary data (5 AWS, 7 ANS) suggested high anxious AWS and ANS exhibited an increase in emotional reactivity along with an attentional bias towards background information. High anxious AWS also showed an increase in stuttering.

A discussion will focus on the importance of understanding the effects of stress-related attentional shifts on speech motor control in PWS during SET conditions and its implications on fluency levels.

References

Bauerly, K.R. (2021). Attentional biases in adults who stutter before and following social threat induction. *Folia Phoniatrica et Logopaedica*, 34614498.

Chen, Y.P., Ehlers, A., Clark, D.M., & Mansell, W. (2002). Patients with generalized social phobia direct their attention away from faces. *Behavior Research and Therapy, 40*, 677-687.

Craig, A. & Tran, Y. (2014). Trait and social anxiety in adults with chronic stuttering: Conclusions from following meta-analysis. Journal of Fluency Disorders, 40, 35-43.

Freedman, S., Maas, E., Caligiuri, M., Wulf, G. & Robin, D. (2007). Internal versus external: Oral-motor performance as a function of attentional focus. *Journal of Speech, Language, and Hearing Research, 50*, 131-136.

Jackson, E.S., Tiede, M., Whalen, D.H. (2016). The impact of social-cognitive stress on speech variability, determinism, and stability in adults who do and do not stutter. *Journal of Speech Language Hearing Research, 59*, 1295-1314.

Lohse, K.R., Sherwood, D.E., Healy, A.F. (2010). How changing the focus of attention affects performance, kinematics, and electromyography in dart throwing. Human Movement Science, 29, 542-555.

Van Lieshout, P., Ben-David, B., Lipski, M. & Namasivayam, A. (2014). The impact of threat and cognitive stress on speech motor control in people who stutter. *Journal of Fluency Disorder, 40*, 93-109.

Wulf, G., McNevin, N., & Shea, C. (2001). The automaticity of complex motor skill learning as a function of attentional focus. *The Quarterly Journal of Experimental Psychology, 54A*, 1143–1154.

Walkom, G. (2022). withVR. [Computer software].

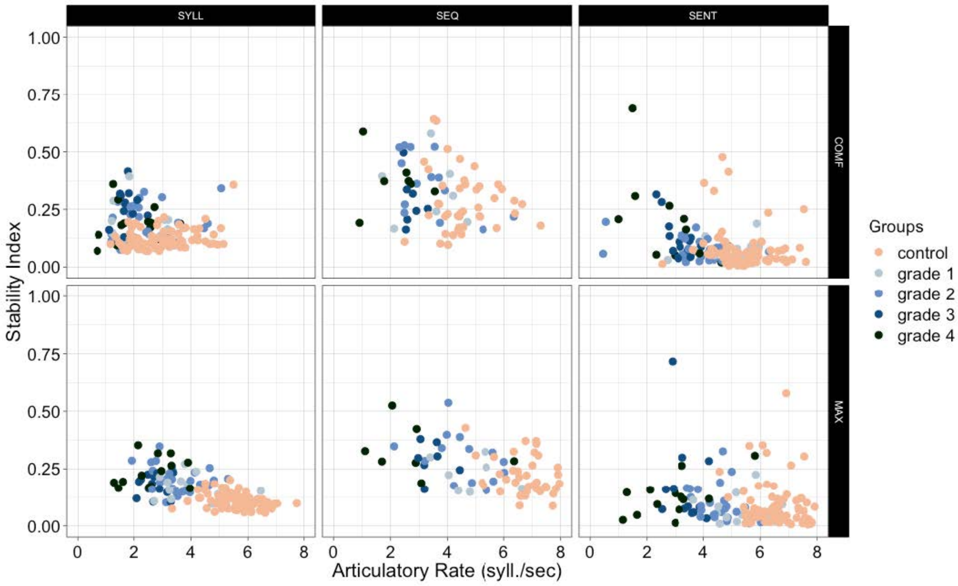
EFFECTS OF INSTRUCTION AND CONTENT ON REPETITION PERFORMANCES OF ATAXIC DYSARTHRIC AND HEALTHY SPEAKERS

Angélina Bourbon, Cécile Fougeron, Lise Crevier-Buchman
CNRS, Grenoble, France
Université Sorbonne Nouvelle Paris-3, Paris, France

Speakers with ataxic dysarthria associated with a cerebellar syndrome present multidimensional impairment, including a slowing of speech (e.g., Duffy, 2013). This characteristic is often accompanied by isochronic syllables in English, observable in repetition paradigms (e.g., Ackermann & Hertrich, 1994). However, this impairment in temporal organization has been shown to depend on both language (e.g., Ikui et al., 2012) and task (Ziegler & Wessel, 1996; Ziegler, 2002; Staiger et al., 2017). This task-dependent effect has been explained by patient heterogeneity in cerebellar diseases with impairments either at a low motoric level or at the level of the encoding of the speech plans. According to the locus of the breakdown, effects on temporal organization of speech could surface differently according to the speech/speech-like content of the material (syllable, sentences) (Maas, 2016) and the performance instructions (Kent et al., 1987). Our aim in this study is to question further the effect of the speech-likeliness of the content to utter and the instructions given in repetition tasks on articulation rate and its stability in healthy and ataxic French speakers.

35 ataxic speakers from different genetic origins (including 30 spinocerebellar ataxia), split into 4 group of dysarthric severity (4 = severe) and 41 matched healthy control, participated in 6 repetition tasks. The material to repeat were either single syllables (/ba/ and /go/), or sequence of syllables (/badego/), or two short meaningful sentences. These items had to be repeated continuously for 15 seconds in two instruction conditions: first, with a DDK performance instructions, i.e. as fast and as precisely as possible (maximum condition), then at a self-determined comfortable rate. Articulation rate in syllables/sec. and acceleration were then computed. The stability of the temporal organization during repetition was assessed via an index computing normalized duration differences between two consecutive items (syllable or sentence, excluding pauses) as $(|Item_i - Item_{i-1}$ / $((Item_i + Item_{i-1}/2)$. For the sentence repetition task, only the first 3 repetitions were considered in order to have an equivalent sample size for the two groups.

Results are illustrated in Fig1. An effect of the repeated item is found for control speakers in instruction condition, with a faster articulation rate and more temporal stability in the sentence when compared to syllable repetition. Sequence repetition is slower than sentence repetition only at comfortable rate with speaker-dependent stability. Ataxic speakers present the same item-dependent effect: in the sentence repetition task, they speak faster and with more stability from one repetition to the next than in the syllable or sequence repetition, in both instruction conditions. For the sequence repetition, stability is also highly dependent on the speaker. Nonetheless, ataxic speakers are slower than the control speakers in all tasks. Interestingly, after a certain grade of severity (grade 3 or 4), dysarthric speakers fail to accelerate in the Max condition for the sequence and sentence repetitions but do so for the less speech-like item: the syllable repetition.



Articulation rate and temporal stability index from one item to the next in the 6 repetition tasks: syllable (SYLL), sequence (SEQ), sentence (SENT) repetitions in both instruction conditions (COMF: comfortable rate, MAX: as fast and as accurate). Ataxic patients are represented in blue colors according to the severity of the dysarthria (1=mild and 4=severe) and the control healthy speakers in orange.

References

Ackermann, H., & Hertrich, I., (1994). Speech rate and rhythm in cerebellar dysarthria: an acoustic analysis of syllabic timing. *Folia Phoniatrica et Logopaedica, 46*, 70–78.

Duffy, J. R. (2013). *Motor speech disorders: Substrates, differential diagnosis, and management.* St. Louis, MO: Elsevier Health Sciences.

Ikui, Y., Tsukuda, M., Kuroiwa, Y., Koyano, S., Hirose, H., & Taguchi, T. (2012). Acoustic characteristics of ataxic speech in Japanese patients with spinocerebellar degeneration (SCD). International *Journal of Language and Communication Disorders, 47*(1), 84–94.

Kent, R. D., Kent, J. F., & Rosenbek, J. C. (1987). Maximum performance tests of speech production. *Journal of Speech and Hearing Disorders, 52*, 367–387.

Maas, E. (2016) Speech and nonspeech: What are we talking about? *International Journal of Speech-Language Pathology, 19*, 345–359.

Staiger, A., Schölderle, T., Brendel, B., Bötzel, K., and Ziegler, W. (2017). Oral Motor Abilities Are Task Dependent: A Factor Analytic Approach to Performance Rate. *Journal of Motor Behavior, 49*(5), 482–493.

Ziegler, W., & Wessel, K. (1996). Speech timing in ataxic disorders: Sentence production and rapid repetitive articulation. *Neurology, 47*, 208–214.

Ziegler, W. (2002). Task-related factors in oral motor control: Speech and oral diadochokinesis in dysarthria and apraxia of speech. *Brain and Language, 80*, 556–575.

MOTOR SPEECH PROGRAMMING AND UTTERING CONDITIONS: A PRELIMINARY STUDY IN SPEAKERS WITH MOTOR SPEECH DISORDERS

Marion Bourqui[1], Marion M. C. Corre[1,2], Cécile Fougeron[3], Marina Laganaro[1]
[1]University of Geneva, Geneva, Switzerland
[2]Geneva University Hospital, Geneva, Switzerland
[3]CNRS, Grenoble, France

Introduction There has been a long debate in the literature about speech production models for several years. Some authors propose a one-step model between phonological encoding and articulation (e.g. "phonetic encoding" in Levelt., 1989) while others include two processes allowing the transformation of a linguistic code into a motor program, (Guenther, 2016; Van der Merwe, 2020) sometimes called "motor speech planning" and "motor speech programming". The latter models are based on observations of the pathology. Indeed, a broad consensus has emerged in the literature that apraxia of speech (AoS) involves impaired ability to retrieve and/or assemble the different elements of the phonetic plans (Blumstein, 1990; Code, 2021; Ziegler, 2009), and the impairment has been located at the motor speech planning processing stage. A different locus has been attributed to dysarthria, which underlying impairment has been located at the motor speech programming processing stage. There is however very limited empirical evidence in favor of two distinct processing stages transforming a linguistic code into articulation. One approach that can be used to differentiate these two processing stages could be to contrast production in speakers with impaired motor speech planning (AoS) versus motor speech programming (dysarthria) in normal versus modified uttering conditions of speech. As speaking in different uttering conditions likely involves parametrization of motor speech programs (Jovicic & Saric, 2006), we expect a differential impact on speakers with AoS and speakers with dysarthria.

Methods

Participants: So far, we enrolled 6 participants suffering from AoS following a left hemisphere stroke; 6 participants suffering from hypokinetic dysarthria (Parkinson's disease - PD) and 12 matched healthy controls.

Material and procedure: Stimuli consist of 54 bisyllabic pseudo-words varying on the first syllable structure (CV versus CCV). A delayed production task was used to separate linguistic from motor speech encoding. The participant had to produce the target stimuli as fast and accurately as possible under two uttering conditions: normally or whispering in a counterbalanced order across participants.

Results Accuracy was coded by two independent raters (inter-rater agreement between .84 and .94, almost perfect agreement (Kappa statistics, Landis & Koch, 1977)). We found a task effect in participants with dysarthria who present lower performances in whispered speech, but not in participants with AoS. The preliminary acoustic analyses on syllables and pseudowords durations (using the Praat software, Boersma & Weenink, 2022), show a lengthening of the speech sequences in the whispered speech condition in all participants, but less pronounced in participants with dysarthria.

Conclusions Our preliminary results seem to indicate that whispering involves lengthening of speech sequences as previously shown for clear speech (Goberman & Elmer, 2005), but also that participants with dysarthria are more impacted at the level of performance when they are asked to produce speech under different uttering conditions.

Acknowledgments This research was supported by Swiss National Science Foundation grants no. CRSII5_173711.

References

Blumstein, S. (1990). Phonological deficits in aphasia: Theoretical perspectives. In A. Caramazza (Ed.), *Cognitive neuropsychology and neurolinguistics* (pp. 33-53). Hillsdalle: Lawrence Erlbaum.

Code, C. (2021). Contemporary issues in apraxia of speech. *Aphasiology, 35*(4), 391-396.

Goberman, A. M. & Elmer, L. W. (2005). Acoustic analysis of clear versus conversational speech in individuals with Parkinson disease. *Journal of Communication Disorders, 38*, 215-230.

Guenther, F. H. (2016). *Neural control of speech*: Mit Press.

Jovicic, S. T. & Saric, Z. (2006). Acoustic Analysis of Consonants in Whispered Speech. *Journal of Voice, 22*(3), 263-274.

Levelt, W. J. (1989). Speaking: From intention to articulation (Vol. 1): MIT press.

Van Der Merwe, A. (2020). New perspectives on speech motor planning and programming in the context of the four- level model and its implications for understanding the pathophysiology underlying apraxia of speech and other motor speech disorders. *Aphasiology, 34*, 1-27.

Ziegler, W. (2009). Modelling the architecture of phonetic plans: Evidence from apraxia of speech. *Language and Cognitive Processes, 24*(5), 631–661.

VALIDATION OF AN ALGORITHM FOR AUTOMATIC PAUSE AND SPEECH TIMING ANALYSES IN FRENCH SPEAKERS WITH ALS

Liziane Bouvier[1,2], Scotia McKinley[1,2], Yana Yunusova[2,3]
[1]*Sunnybrook Research Institute, Toronto, Canada*
[2]*University of Toronto, Toronto, Canada*
[3]*University Health Network—Toronto Rehabilitation Institute, Toronto, Canada*

Bulbar dysfunction is present in most patients with Amyotrophic Lateral Sclerosis (ALS) either from the onset or during the progression of the disease (Makkonen, 2018). Bulbar signs include impairments in speech (dysarthria) and swallowing (dysphagia) and are associated with faster disease progression, shorter survival, and lower quality of life (Chio et al., 2011). The development of effective measurement tools that improve early detection and monitoring of bulbar signs and progression monitoring is among the highest priorities in ALS research across the globe (Makkonen, 2018; Tena et al., 2021; Vieira et al., 2019; Yunusova et al., 2019).

Temporal measures of speech such as speaking rate, speech (phrase) and pause durations, percent pause time, mean phrase duration, etc., are of great interest in ALS. In English, various temporal measures of speech obtained during passage reading tasks have been identified as valid and reliable biomarker of bulbar dysfunction (Barnett et al., 2020) and showed significant sensitivity to different stages of bulbar disease (Barnett et al., 2020; Ball et al., 2002; Allison et al., 2018; Yunusova et al., 2016). Manual measurements of temporal features of speech are both time consuming and vulnerable to measurement error. Thus, automatic tools for pause detection and temporal measurements are preferable. Such tools (e.g., the Speech-Pause Analysis algorithm; Green et al., 2004) have been successfully validated and used in English but are lacking for French. Considering the linguistic differences in prosody between English and French, cross-linguistic validation of the measurement tools is required.

The aim of this study was to evaluate the validity of an algorithm designed to automatically extract pauses and speech timing information from connected speech samples, the Speech-Pause Analysis algorithm (10), in Canadian French. Speech samples were obtained from 37 Canadian French speakers with amyotrophic lateral sclerosis (ALS) and 20 control speakers. The speech samples consisted of digitally recorded recitations of the French translation of the Bamboo passage that was initially developed for use with SPA. Passage recordings were analysed manually using Praat and algorithmically using SPA.

The manual and algorithmic methods showed excellent within and inter-method reliability. In addition, the algorithmic method was faster to perform than the manual method. These results suggested that the algorithm provided an efficient and valid method for extracting pause and speech timing information from the translated Bamboo passage in Canadian French speakers. Further studies should investigate the best pause detection parameters in French speakers by comparing different parameters such as the minimal length of pauses with different populations and speaking rates. Validation with other clinical populations is also important.

References

Allison KM, Yunusova Y, Campbell TF, Wang J, Berry JD, Green JR. (2018) The diagnostic utility of patient-report and speech-language pathologists' ratings for detecting the early onset of bulbar symptoms due to ALS. *Physiol Behav.,* 18(5–6):358–66.

Ball LJ, Beukelman DR, Pattee GL. (2002). Timing of Speech Deterioration in People with Amyotrophic Lateral Sclerosis. *J Med Speech Lang Pathol, 10*(4):231–5.

Barnett C, Green JR, Marzouqah R, Stipancic KL, Berry JD, Korngut L, et al. (2020). Reliability and Validity of Speech & Pause Measures during Passage Reading in ALS. *Amyotroph Lateral Scler Front Degener., 21*(1–2):42–50.

Chiò A, Calvo A, Moglia C, Mazzini L, Mora G. (2011). Phenotypic heterogeneity of amyotrophic lateral sclerosis: a population based study. *J Neurol Neurosurg Psychiatry, 82*(7):740–6.

Green JR, Beukelman DR, Ball LJ. (2004). Algorithmic estimation of pauses in extended speech samples of dysarthric and typical speech. *Journal of Medical Speech-Language Pathology*, p. 149–54.

Makkonen T, Ruottinen H, Puhto R, Helminen M, Palmio J. (2018). Speech deterioration in amyotrophic lateral sclerosis (ALS) after manifestation of bulbar symptoms. *Int J Lang Commun Disord, 53*(2):385–92.

Tena A, Claria F, Solsona F, Meister E, Povedano M. (2021). Detection of Bulbar Involvement in Patients With Amyotrophic Lateral Sclerosis by Machine Learning Voice Analysis: Diagnostic Decision Support Development Study. *JMIR Med Informatics, 9*(3).

Vieira H, Costa N, Sousa T, Reis S, Coelho L. (2019). Voice-Based Classification of Amyotrophic Lateral Sclerosis: Where Are We and Where Are We Going? A Systematic Review. *Neurodegener Dis, 19*:163–70.

Yunusova Y, Graham NL, Shellikeri S, Phuong K, Kulkarni M, Rochon E, et al. (2016). Profiling speech and pausing in amyotrophic lateral sclerosis (ALS) and frontotemporal dementia (FTD). *PLoS One, 11*(1):1–18.

Yunusova Y, Plowman EK, Green JR, Barnett C, Bede P. (2019). Clinical measures of bulbar dysfunction in ALS. *Front Neurol, 10*: 106.

TEST-RETEST STABILITY OF WORD SYLLABLE DURATION FOR SPEAKERS WITH ACQUIRED APRAXIA OF SPEECH

Lisa Bunker[1], Dallin J. Bailey[2], Elaine Poss[3], Shannon Mauszycki[4], Julie L. Wambaugh[4]
[1] *Johns Hopkins University School of Medicine, Baltimore, Maryland, United States of America*
[2] *Auburn University, Auburn, Alabama, United States of America*
[3] *Renown South Meadows Medical Center, Reno, Nevada, United States of America*
[4] *University of Utah, Salt Lake City, United States of America*

Introduction Considerable effort has gone toward clarifying diagnostic criteria of acquired apraxia of speech (AOS) (McNeil et al., 2009). Despite improved criteria, diagnosis remains challenging. Recently, efforts have focused on identifying acoustic measures that characterize AOS for both improved diagnostic accuracy and differentiation from aphasia with phonemic paraphasia (APP) (Ballard et al., 2016; Cunningham et al., 2016; Haley & Jacks, 2019; Haley et al., 2012; Haley et al., 2021; Haley et al., 2017). For differential diagnosis, acoustic analyses often target speaking rate (i.e., speaking rate is reduced in AOS but [near] normal in APP). Word Syllable Duration (WSD) is a measure of speech rate potentially able to differentiate between AOS and APP (Haley & Jacks, 2019; Haley et al., 2012; Haley et al., 2021; Haley et al., 2017). Little is known about the reliability/stability of WSD over time. This investigation examined reliability of WSD at 1- and 4-week intervals for individuals with AOS and aphasia.

Methods Twenty-nine participants with chronic AOS and aphasia (i.e., >12 months-post-onset) secondary to stroke/neurological trauma were included in this study (11 Female, *mean* [SD] age = 54.3[12.7] years). Participants completed multisyllabic word repetition tasks across three sampling times (T1, T2, and T3) with 1-week and 4-week intervals between T1 to T2 and T1 to T3, respectively. Repetition tasks included lists comprised of ten 3-, 4-, or 5-syllable words each. Token selection and measurement procedures are detailed in Haley et al. (2017). We also removed any responses with fewer than three syllables for our analysis. Participants contributed an average of 9. 7 tokens per list (range = 5-10) per sampling occasion.

Total word duration was measured using Praat (Boersma & Weenink, 2001) and WSD was derived by dividing word duration by number of syllables produced. At each timepoint (T1, T2, and T3), WSDs were averaged across each set (e.g., across all ten 3-syllable words) as well as across all tokens combined. Due to skewed distributions/autocorrelation, a non-parametric Friedman's test (Fr, r referring to rank) was used to calculate test-retest reliability.

Results Results are reported in Table 1. Friedman's tests were not significant for 3-syllable words (Fr[2] = 1.10, p = 0.58), 4-syllable words (Fr[2] = 4.34, p = 0.11), 5-syllable words (Fr[2] = 1.93, p = 0.38) and all tokens combined (Fr[2] = 1.93, p = 0.38).

Discussion Non-significant Friedman's tests indicate that WSD was stable across sampling times (i.e., individuals with chronic AOS and aphasia repeated words with statistically equal speech rates across T1-T3). Demonstration of stability over 1- and 4-week intervals is important for WSD to be used for both diagnosis and measuring outcomes. As noted in Table 1, average WSD for a few participants varied nearly 200ms or more between T1-T3 for a single word list (i.e., 10 tokens or less), but the average max difference across timepoints decreased for combined lists.

Differences of even 100-200ms could be significant for diagnosis, thus more than 10 tokens or a combination of multisyllabic word lengths may be more preferable for reliability. Additional analysis is still needed to establish the reliability of WSD, but current results do suggest that WSD is stable with repeated assessments.

Word Length	F_r	*p*	T1 *M* (*SD*)	T1 Range	T2 *M* (*SD*)	T2 Range	T3 *M* (*SD*)	T3 Range
3 Syl.	1.10	.58	419 (142)	212-775	423 (161)	208-827	430 (162)	222-857
4 Syl.	4.34	.11	465 (201)	203-1049	455 (183)	213-973	445 (175)	212-946
5 Syl.	1.93	.38	480 (208)	212-1127	463 (181)	209-946	460 (196)	191-941
Combined (3-5 Syl.)	1.93	.38	454 (180)	209-968	447 (172)	215-885	445 (175)	208-879

Max difference in WSD across sample times:	3 Syl.	4 Syl.	5 Syl.	Combined
Mean	44	62	72	40
Range	4-112	10-214	12-294	1-128

Note. All word syllable duration (WSD) times in ms. T1 = time 1; T2 = time 2; T3 = time 3; Syll = syllable.

Friedman's Test Results for 3-, 4-, 5-, and Combined-Syllable Word List

References

Ballard KJ, Azizi L, Duffy JR, et al. (2016). A predictive model for diagnosing stroke-related apraxia of speech. *Neuropsychologia, 81*, pp. 129-139. https://doi.org/10.1016/j.neuropsychologia.2015.12.010

Boersma P, & Weenink D. (2001). Praat, a system for doing phonetics by computer. *Glot International, 5*(9/10), 341-345.

Cunningham KT, Haley KL, & Jacks A. (2016). Speech sound distortions in aphasia and apraxia of speech: reliability and diagnostic significance. *Aphasiology, 30*(4), pp. 396-413. https://doi.org/10.1080/02687038.2015.1065470

Haley KL, & Jacks A. (2019). Word-level prosodic measures and the differential diagnosis of apraxia of speech. *Clinical Linguistics and Phonetics, 33*(5), pp. 479-495.

Haley KL, Jacks A, de Riesthal M, Abou-Khalil R, & Roth HL. (2012). Toward a quantitative basis for assessment and diagnosis of apraxia of speech. *Journal of Speech, Language, and Hearing Research, 55*(5). https://doi.org/10.1044/1092-4388(2012/11-0318)

Haley KL, Jacks A, Jarrett J, et al. (2021). Speech metrics and samples that differentiate between nonfluent/ agrammatic and logopenic variants of primary progressive aphasia. *Journal of Speech, Language, and Hearing Research, 64*(3), pp. 754-775. https://doi.org/10.1044/2020_JSLHR-20-00445

Haley KL, Jacks A, Richardson JD, & Wambaugh JL. (2017). Perceptually salient sound distortions and apraxia of speech: A performance continuum. *American Journal of Speech-Language Pathology, 26*, pp. 631-640. https://doi.org/10.1044/2017_AJSLP-16-0103

McNeil MR, Robin DA, & Schmidt RA. (2009). Apraxia of speech: Definition and differential diagnosis. In: M. R. McNeil (Ed.), *Clinical Management of Sensorimotor Speech Disorders*, 2nd ed. (pp. 249-268). Thieme.

AN FMRI STUDY OF OVERT SPEECH PREPARATION AND PRODUCTION IN CHILDREN WHO STUTTER

Ho Ming Chow[1], Emily Garnett[2], Nan Ratner[3], Soo-Eun Chang[2]
[1] *University of Delaware, Newark, Delaware, United States of America*
[2] *University of Michigan, Ann Arbor, Michigan, United States of America*
[3] *University of Maryland, College Park, Maryland, United States of America*

Recent neuroimaging studies of children who stutter (CWS) have indicated that developmental stuttering is associated with functional and structural anomalies in the cortical and subcortical areas classically associated with speech motor control (Neef et al., 2015). FMRI is arguably the best non-invasive tool for the investigation of cortical and subcortical motor area activity during speech production. However, a major obstacle that hinders the use of fMRI to study speech production is that speech-related movements are a major source of contamination in fMRI. Thus, most of the previous fMRI studies employed relatively simple speaking tasks (e.g., single-word production) or non-speaking contexts (e.g., wakeful rest) to examine speech motor functions. However, these procedures do not optimally inform the neural correlates of childhood stuttering because moments of stuttering typically do not occur during simple speaking tasks (Usler & Walsh, 2018). Spontaneous, continuous speech, on the other hand, is ecologically-valid for studying the disorder because symptoms of stuttering are most apparent in this context. However, continuous speech production is also more likely to induce severe movement-related artifacts that contaminate the fMRI signal, especially in children. To overcome this obstacle, we used a technique, based on spatial independent component analysis (sICA), to remove fMRI artifacts during continuous speech production (Xu et al., 2014). This technique decomposes individual functional images into a number of components. Components with spatial biologically implausible patterns could be identified automatically and removed from the functional images.

Using this denoising technique, we examined two aspects of speech production known to induce dysfluencies in CWS, formulation and preparation of speech. To examine the effect of speech formulation, we compared brain activation during continuous (propositional) and automatic (non-propositional) speech production in CWS with their fluent peers. For the effect of speech preparation, we compared brain activation during the period shortly before (2 seconds) the initiation of speech in CWS and controls. We also examined age-related changes in brain activity patterns during speech preparation and production. In total, our final analysis included 28 CWS (13 boys and 15 girls) and 24 controls (11 boys and 13 girls) in the 5-12.5 age range.

Comparing continuous and automatic speech conditions, the control group exhibited increased left-lateralized activation in the inferior frontal gyrus, premotor cortex and pre-supplementary motor area as well as the bilateral posterior middle/inferior temporal gyri. The direct group comparison in the basal-ganglia-thalamocortical (BGTC) loop showed that CWS exhibited decreased activation in the left premotor cortex (PMC). Activation in the bilateral PMC during speech production furthermore increased with age in controls but decreased with age in CWS. Age-related activation decreases were observed in the putamen and the thalamus for CWS during speech preparation. Taken together, the results support the notion that stuttering is associated with aberrant development and functional deficits in the BGTC loop affecting the preparation and production of spontaneous speech.

References

Neef, N. E., Anwander, A., & Friederici, A. D. (2015). The neurobiological grounding of persistent stuttering: from structure to function. *Current Neurology and Neuroscience Reports, 15*(9), 63. https://doi.org/10.1007/s11910-015-0579-4

Usler, E. R., & Walsh, B. (2018). The effects of syntactic complexity and sentence length on the speech motor control of school-age children who stutter. *Journal of Speech, Language, and Hearing Research, 61*(9), 2157–2167. https://doi.org/10.1044/2018_JSLHR-S-17-0435

Xu, Y., Tong, Y., Liu, S., Chow, H. M., AbdulSabur, N. Y., Mattay, G. S., & Braun, A. R. (2014). Denoising the speaking brain: toward a robust technique for correcting artifact-contaminated fMRI data under severe motion. NeuroImage, 103, 33–47. https://doi.org/10.1016/j.neuroimage.2014.09.013

LANGUAGE LATERALISATION IN PEOPLE WHO STUTTER ACROSS DIFFERENT SPEECH & LANGUAGE TASKS

Birtan Demirel, Kate Watkins
University of Oxford, Oxford, United Kingdom

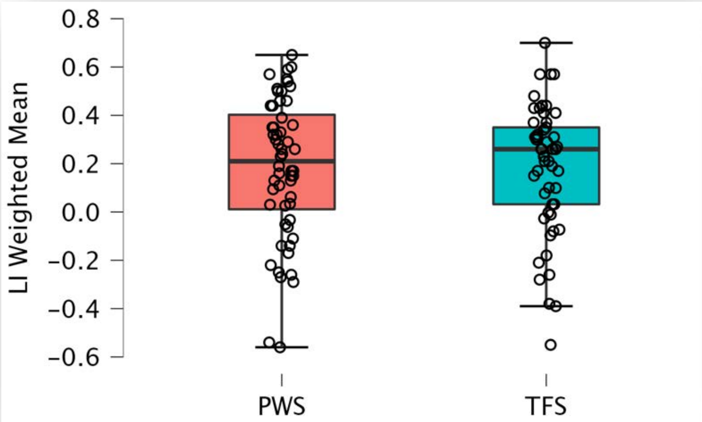
One of the oldest neurobiological explanations of stuttering is the incomplete cerebral dominance theory, which suggests altered patterns of hemispheric specialisation and competition between hemispheres for “dominance” over handedness and speech (Orton, 1927). Brain imaging findings in people who stutter of increased activity in the right hemisphere during speech production (Brown et al., 2005) or of shifts in activity from right to left when fluency is increased (de Nil et al., 2003) led to renewed interest in these ideas (Sato et al., 2011).

Here, we revisited this theory using functional MRI data obtained in children and adults who stutter. Laterality indices (LIs) were calculated for the frontal lobes using the LI toolbox (Wilke & Lidzba, 2007) running in Statistical Parametric Mapping where positive values indicate left- and negative ones right- lateralisation, and values between $-0.2 \leq \text{LI} \leq 0.2$ indicates bilateral representation (Wilke et al., 2006).

LIs were compared for overt sentence reading in 56 people who stutter (PWS) (10 female and 46 male, 7 left handed, Mage = 28 years, SDage = 9 years) and 53 typically fluent speakers (TFS) (13 female and 40 male, 2 left handed, Mage = 28 years, SDage = 10 years). The average LI in PWS was 0.18 (SD = 0.29) and in TFS it was 0.18 (SD = 0.26). A Bayesian independent samples t-test (two-sided) revealed a Bayes factor of 4.92 indicating moderate evidence in support of the null hypothesis (LIs are similar between PWS and TFS) rather than the alternative hypothesis (LIs differ between PWS and TFS). Chi-squared statistics confirmed that the number of individuals in each group who showed the typical pattern of leftwards laterality compared with atypical (right or bilateral LIs) did not differ (PWS 28 typical, 28 atypical; TFS 31 and 22; $\chi = 0.791$, $p < 0.374$).

In different subsets of the sample, we also analysed LIs for a covert auditory naming task (12 PWS, 16 TFS; the mean LIs in PWS 0.38 [SD = 0.29] and in TFS 0.41 [SD = 0.17]), covert sentence reading task (12 PWS, 12 TFS; 0.57 [SD = 0.09], 0.62 [SD = 0.10]) and overt picture description task (16 PWS, 18 TFS; 0.21 [SD = 0.27], 0.14 [SD = 0.25]). Bayesian independent samples t-test analyses revealed bayes factors of 2.69, 1.55, and 2.41, respectively, indicating evidence in support of the null hypothesis (no group differences for each task). Chi-squared analyses also confirmed that the groups did not differ in terms of the number of typically or atypically lateralised individuals in the covert auditory naming task (PWS 9 typical and 3 atypical; TFS 13 and 3; $\chi^2 = 0.159$, $p < 0.69$), covert sentence reading task (all participants were left lateralised), and overt picture description task (PWS 9 and 7; TFS 8 and 10; $\chi^2 = .47$, $p < 0.49$).

Our analyses found no support for the theory that laterality is reduced or differs in PWS compared with TFS. Further analyses will explore the laterality whether it differs according to lobe.



Overt Sentence Reading Task

References

Brown, S., Ingham, R. J., Ingham, J. C., Laird, A. R., & Fox, P. T. (2005). Stuttered and fluent speech production: an ALE meta-analysis of functional neuroimaging studies. *Human Brain Mapping, 25*(1), 105–117. https://doi.org/10.1002/hbm.20140

de Nil, L. F., Kroll, R. M., Lafaille, S. J., & Houle, S. (2003). A positron emission tomography study of short- and long-term treatment effects on functional brain activation in adults who stutter. *Journal of Fluency Disorders, 28*(4), 357–380. https://doi.org/https://doi.org/10.1016/j.jfludis.2003.07.002

Orton, S. T. (1927). Studies in Stuttering: Introduction. *Archives of Neurology & Psychiatry, 18*(5), 671–672.

Sato, Y., Mori, K., Koizumi, T., Minagawa-Kawai, Y., Tanaka, A., Ozawa, E., Wakaba, Y., & Mazuka, R. (2011). Functional lateralization of speech processing in adults and children who stutter. *Frontiers in Psychology, 2*, 70. https://doi.org/10.3389/fpsyg.2011.00070

Wilke, M., & Lidzba, K. (2007). LI-tool: A new toolbox to assess lateralization in functional MR-data. *Journal of Neuroscience Methods, 163*(1), 128–136. https://doi.org/https://doi.org/10.1016/j.jneumeth.2007.01.026

Wilke, M., Lidzba, K., Staudt, M., Buchenau, K., Grodd, W., & Krägeloh-Mann, I. (2006). An fMRI task battery for assessing hemispheric language dominance in children. *NeuroImage, 32*(1), 400–410. https://doi.org/https://doi.org/10.1016/j.neuroimage.2006.03.012

FINDING A NEEDLE IN A HAYSTACK: STUDYING STUTTERING ARTICULATORY TRAJECTORIES USING AUTOMATIC ANALYSIS ON LIMITED DATA

Ivana Didirková[1], Sébastien Le Maguer[2]
[1] *Université Paris 8 Vincennes-Saint-Denis, Paris, France*
[2] *ADAPT Centre/Trinity College Dublin, Dublin, Ireland*

Stuttering is a speech motor control disorder with neurological and genetic origins (Guitar, 2019). It is mainly characterized by frequent speech disfluencies (blocks, repetitions, prolongations), which are considered different from non-stuttered disfluencies (Lickley, 2017). Differentiating parameters include specific supraglottic articulatory patterns, such as a substantial amount of movements (Didirková et al., 2021), which are not systematically well converted into acoustic data. In parallel, recent studies initiating machine learning technologies to detect and label stuttering speech automatically are based on acoustic features and do not yet investigate how their prediction relates to disfluency. Furthermore, the standard machine learning models require a huge amount of data incompatible with articulatory-based analysis. For these reasons, we bring together articulatory phonetics and machine learning to investigate new ways of describing the trajectories of supraglottic speech articulators during Stuttering-Like Disfluencies (SLD) produced by Persons Who Stutter (PWS) and Other Disfluencies (OD) produced both by PWS and Persons Who do Not Stutter (PWNS). Due to spasmodic movements (Busan et al., 2021), we hypothesize that supraglottic articulatory trajectories should exhibit more deviation from their primary trajectory (i.e., the expected up-/downward or forward/backward movement) compared to OD. To test our hypothesis, we approximate the primary trajectory by computing a cubic regression for each configuration (i.e., a couple (axis, articulator)) and each disfluency. We then calculate a set of distances between the data points and the approximation of the primary trajectory to measure two key aspects relative to the primary trajectory: the stability and the deviation magnitude. The stability, quantified using the Zero Crossing (ZC), measures how the realized movement follows the primary trajectory. The closer to 0 is the ZC, the more the realized movement follows the primary trajectory. The deviation magnitude, quantified using the Mean Absolute Error (MAE) and the Root Mean Square Error (RMSE), indicates how far the realized movement is from the primary trajectory. The closer to 0 the MAE and the RMSE, the less the realized movement diverges from the primary trajectory. Our analyses concern four PWS and four PWNS, native French speakers, producing about one hour of semi-spontaneous speech with no disfluency elicitation technique. The corpus was captured using the Electromagnetic Articulograph (EMA) device Carstens AG-501 and manually annotated for SLD and OD. A full description of the corpus is available in Didirková et al (2021). Our results show that fluent speech is close to the primary trajectory for both PWS and PWNS. During OD, PWS produced movements closer to the primary trajectory than PWNS. While this result requires a more in-depth investigation, a possible explanation is that PWS subconsciously tried to have a more controlled movement. Nevertheless, as hypothesized, SLD diverged more from the primary trajectory than OD produced by PWS. These results demonstrate that it is possible to distinguish SLD from OD even with limited data. Based on these results, we are now experimenting the use of machine learning algorithms, such as autoencoders (Tschannen et al., 2018), to estimate what we can learn from these models using limited data. The contribution of our observations to speech production in PWS and automatic detection of stuttering events will be discussed during the conference.

References

Guitar, B. (2019). *Stuttering. An integrated approach to its nature and treatment*, 5th ed. Philadelphia: Wolters Kluwer.

Lickley, R. (2017). Disfluency in typical and stuttered speech. *Fattori sociali e biologici nella variazione fonetica-Social and biological factors in speech variation.*

Didirková, I., Le Maguer, S., & Hirsch, F. (2021). An articulatory study of differences and similarities between stuttered disfluencies and non-pathological disfluencies. *Clinical Linguistics & Phonetics, 35*(3), pp. 201–221.

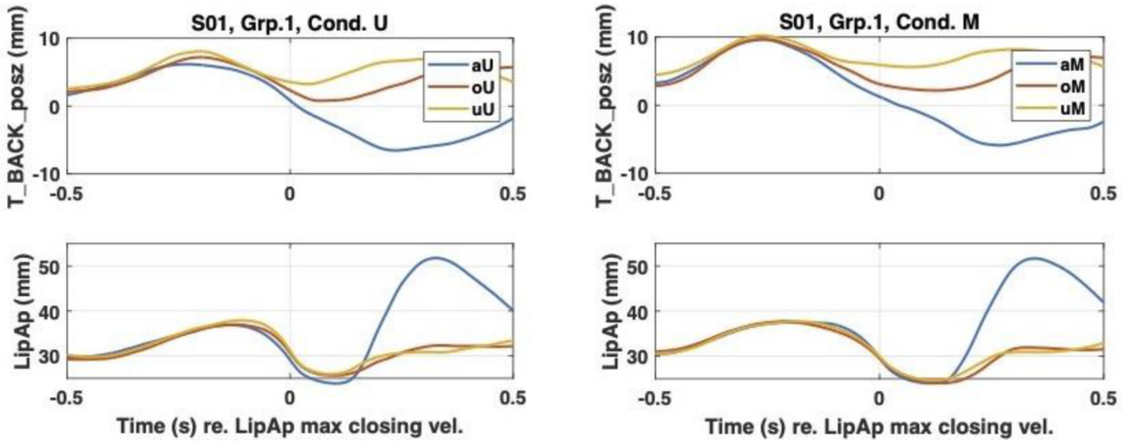
Busan, P., Moret, B., Masina, F., Del Ben, G., & Campana, G. (2021). Speech fluency improvement in developmental stuttering using non-invasive brain stimulation: insights from available evidence. *Frontiers in Human Neuroscience*, p. 437, 2021.

Tschannen, M., Bachem, O. F., & Lučić. (2018). Recent advances in autoencoder-based representation learning. *Bayesian Deep Learning Workshop, NeurIPS*

THE EFFECT OF RHYTHM ON INTER-GESTURAL COUPLING OF ONSET AND VOWEL GESTURES IN ADULTS WHO STUTTER

Mona Franke[1,2,3], Philip Hoole[1], Simone Falk[2,3]
[1]*Ludwig-Maximilians-University, Munich, Germany*
[2]*University of Montréal, Montréal, Canada*
[3]*BRAMS, Montréal, Canada*

In persons who stutter (PWS), dynamics in speech production processes are disrupted and even the perceptually fluent speech of PWS is temporally more variable than in persons who do not stutter. On an articulatory basis, PWS, for example, show higher inter-articulator variability over repeated productions (Smith et al., 2010) and longer movement durations of lip- and jaw-closing movements during fluent speech (Max et al., 2003). The speech motor system of PWS becomes more stable in rhythmic fluency-evoking conditions, e.g., speaking synchronously with a metronome (Namasivayam & van Lieshout, 2011), while an increase in complexity such as synchronizing speech and tapping with a rhythmic tone leads to more timing variability (Hulstijn et al., 1992). However, it is still unclear how exactly rhythm affects articulatory movements in stuttering. Based on the "gestural computational model" (Browman & Goldstein, 1991) we examine if there are differences in the coupling of articulatory gestures in adults who stutter (AWS) vs. adults who do not stutter (AWNS) in different rhythmic conditions. We expect to find 1) potential differences in the coupling of onset (C) and vowel (V) gestures between AWS and AWNS and 2) effects of verbal and non-verbal rhythms on the relative timing between onset and vowel movements. The rhythmic conditions include rhythmic motor pacing (tapping with produced words) in addition to the more traditional rhythmic auditory pacing (articulating with the rhythm of a metronome). The inclusion of motor pacing can inform us about the interrelation between manual motor control and articulatory movements (Parrell et al., 2014), and thereby, elucidate more general timing and rhythmic processes in PWS (e.g., Falk et al., 2015). We hypothesize that AWS have deficits in generating correct inter-gestural timing but that an external rhythm (auditory pacing) will reduce the deficits. Moreover, we hypothesize that motor pacing will lead to more stable articulatory timing in AWNS but not in AWS if the latter indeed have a more general timing deficit. Finally, we expect auditory-motor pacing to lead to greater variability in the gestural coupling in AWS compared to AWNS. We are currently conducting an Electromagnetic Articulography (EMA) experiment with AWS and AWNS, producing German mono- and disyllabic words, embedded in the carrier phrase ['zeːə _____'an] (Look at _____). Onsets of the target words are the bilabial consonants /b/, /m/ and the following long vowels are /a/, /o/, /u/. The experiment contains 4 different conditions; unpaced (self-paced reading), motor pacing (tapping with each word, self-paced), auditory pacing (one metronome sound per word, 90bpm), and auditory-motor pacing (tapping with the metronome during reading, 90bpm). Overall, there are 9 target words that participants produce 4 times in each condition in a randomized order. Analyses of the relative timing of CV gestures (including CV lag from onset to onset, as well as target to target attainment), and for CV gestures in different rhythmic conditions (see Figure 1, displaying the unpaced and auditory pacing condition) for the first speakers have already begun. We expect to have about 5 each AWS and AWNS ready for presentation by the summer.



Example for the temporal evolution of the vowel gesture (mean of all 12 vowel productions of the target words produced by one speaker who stutters (only perceptually fluent speech is analyzed)). The x-axis displays the time, lining up with the maximum velocity of the lip closure movement (= 0). The y-axis displays the vertical movement in mm of the tongue back sensor in the upper row and the Lip Aperture in the lower row. Left column: Unpaced condition. Right column: Auditory pacing (metronome) condition. The legend displays the colors of the different vowels (/a/ = blue, /o/ = red, /u/ = yellow) and the condition (U = unpaced, M = metronome). Note the earlier divergence of vowel-specific movements (relative to the consonantal line-up point) in the metronome condition.

References

Browman, C. & Goldstein, L. (1991). Gestural structures: distinctiveness, phonological processes, and historical change. In I. Mattingly and M. Studdert-Kennedy (Eds.), *Modularity and the Motor Theory of Speech Perception*, 313-338. Erlbaum: New Jersey

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

Hulstijn, W., Summers, J. J., van Lieshout, P. H. M, Peters, H. F. M. (1992). Timing in finger tapping and speech: A comparison between stutterers and fluent speakers. *Human Movement Science, 11*(1-2), 113-124. https://doi.org/10.1016/0167-9457(92)90054-F

Max, L., Caruso, A. J., & Gracco, V. L. (2003). Kinematic analyses of speech, orofacial nonspeech, and finger movements in stuttering and nonstuttering individuals. *Journal of Speech, Language, and Hearing Research, 46*, 215-232

Namasivayam, A. K, van Lieshout P. (2011). Speech motor skill and stuttering. *Journal of Motor Behavior, 43*(6), 477-89. https://doi.org/10.1080/00222895.2011.628347

Parrell, B., Goldstein, L., Lee, S., & Byrd, D. (2014). Spatiotemporal coupling between speech and manual motor actions. *Journal of phonetics, 42*, 1–11. https://doi.org/10.1016/j.wocn.2013.11.002

Smith, A., Sadagopan, N., Walsh, B., & Weber-Fox, C. (2010). Increasing phonological complexity reveals heightened instability in inter-articulatory coordination in adults who stutter. *Journal of Fluency Disorders, 35*(1) 1-18

SPEECH SYNCHRONIZATION ABILITIES IN PEOPLE WHO STUTTER

Maëva Garnier, Anneke Slis, Christophe Savariaux, Pascal Perrier
CNRS, Grenoble, France
GIPSA-Lab, Grenoble, France

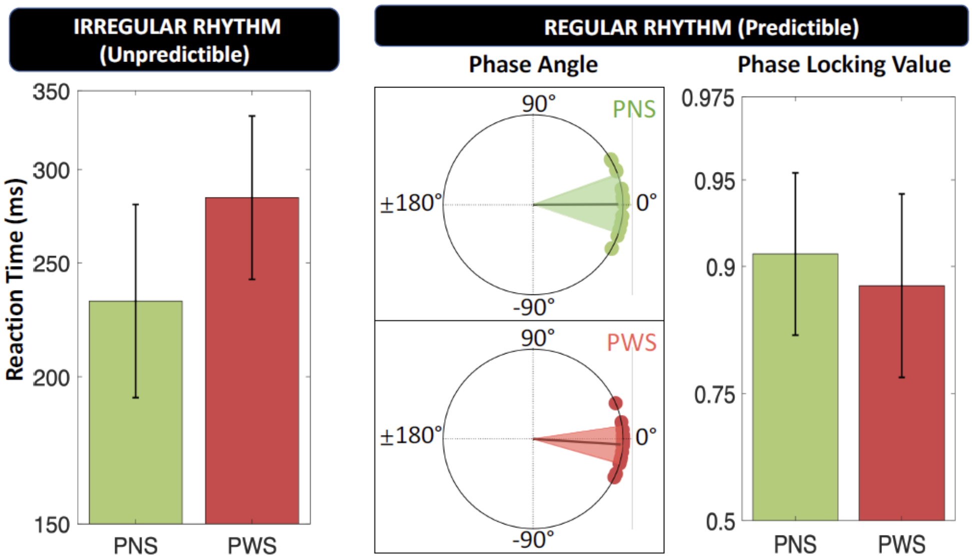
Introduction Stuttering is a speech fluency disorder, resulting in repetitions, prolongations or blockages of sounds. Although the etiology is still not well understood, it has been suggested that that individuals who stutter face motor difficulties, in particular with the initiation of motor sequences (Alm, 2004; Civier et al. 2013; Starkweather et al. 1984), and may present a deficit in predicting and perceiving sensorimotor events (Etchell et al., 2014; Olander et al. 2010; Falk et al. 2015; Sares et al. 2019; Slis et al. 2020).

Objectives The purpose of this study is to test these two aspects, by examining the reactivity of adults who stutter to an unpredictable event, in comparison to typical adult speakers, and to what extent this also affects, or not, their ability to synchronize with a predictable event.

Methodology Sixteen adults who stutter (PWS) and sixteen typical adults who do not stutter (PNS), matched in age, gender, and musical experience, produced two sentences ("Pattie passa la pagaie" and "Bali bannit la bagarre"), in two experimental conditions: synchronizing with a predictable auditory stimulus (metronome every 500ms) or following an unpredictable auditory stimulus (randomly appearing every 200-800ms). During the first synchronization task, the asynchrony (Phase Angle (PA)) between the stimulus and the syllable p-center was measured, as well as the stability of this asynchrony (Phase Locking Value (PLV) (Sares et al. 2019)). From the second task, reaction times (RT) were measured between the stimulus and the syllable onset.

Results and Discussion Compared to PNS, PWS showed significantly longer reaction times when asked to produce speech following an unpredictable external stimulus (ΔRT = 51±15ms, p=0.002, see Figure 1). This is consistent with difficulties in movement initiation. When following a predictable external stimulus, PWS no longer showed such a delay between the stimulus and their syllable onset. This suggest that PWS were able to internalize a pulse and to predict the reappearance of the stimulus, and to use it to compensate for their difficulties in initiating speech gestures. In the synchronization task, PWS aligned their syllable's p-center and the auditory stimuli with a synchronization accuracy and consistency comparable to PNS (ΔPA = -2.3 +/- 5.2 °, HPD=[-12.5 8.2] (using the R package bpnreg for circular statistics); ΔPLV = -0.03 +/- 0.03, p=0.32 ; see Figure 1).

Conclusion The slower reaction times observed in PWS confirm the idea that they encounter motor difficulties with the initiation of speech gestures. On the contrary, the similar performance in synchronization accuracy and consistency do not support the hypothesis that PWS face difficulties in predicting and anticipating recurrent events.



Speech Reaction Times (in the unpredictable task), Phase Angle and Phase locking Values (in the predictable task), for adults who stutter (PWS) and matched control subjects (PNS).

References

ALM, PA. (2004). Stuttering and the basal ganglia circuits: a critical review of possible relations. *J Commun Disord., 37*(4), 325-369.

CIVIER, O., BULLOCK, D., MAX, L. & GUENTHER, F. (2013). Computational modeling of stuttering caused by impairments in a basal ganglia thalamo-cortical circuit involved in syllable selection and initiation. *Brain and Language, 26*, 263-278.

ETCHELL, A., JOHNSON, B., et SOWMAN, P. (2014). Behavioral and multimodal neuroimaging evidence for a deficit in brain timing networks in stuttering: a hypothesis and theory. *Frontiers in human neuroscience, 8*, 467

FALK, S., MULLER, T., DALLA BELLA, S. (2015). Non-verbal sensorimotor timing deficits in children and adolescents who stutter. *Frontiers in Psychology, 6*, 101-114.

SLIS, A., SAVARIAUX, C., PERRIER, P. & GARNIER, M. (2020). Complexity of Rhythmic Tapping Task and Stuttering. *Proc. International Seminar on Speech Production, Rhodes Island.*.

OLANDER, L., SMITH, A., & ZELAZNIK, H. (2010). Evidence that a motor timing deficit is a factor in the development of stuttering. *J. Speech Lang. Hear. Res., 53*, 876–886.

SARES, A., DEROCHE, M., SHILLER, D., et al. (2019). Adults who stutter and metronome synchronization: evidence for a nonspeech timing deficit. *Annals of the New York Academy of Sciences, 1449*(1), 56-69.

STARKWEATHER, C. W., FRANKLIN, S. & SMIGO, T. M. (1984) Vocal and finger reaction times in stutterers and nonstutterers: Differences and correlations. *Journal of Speech, Language, and Hearing Research, 27*,(2), 193-196.

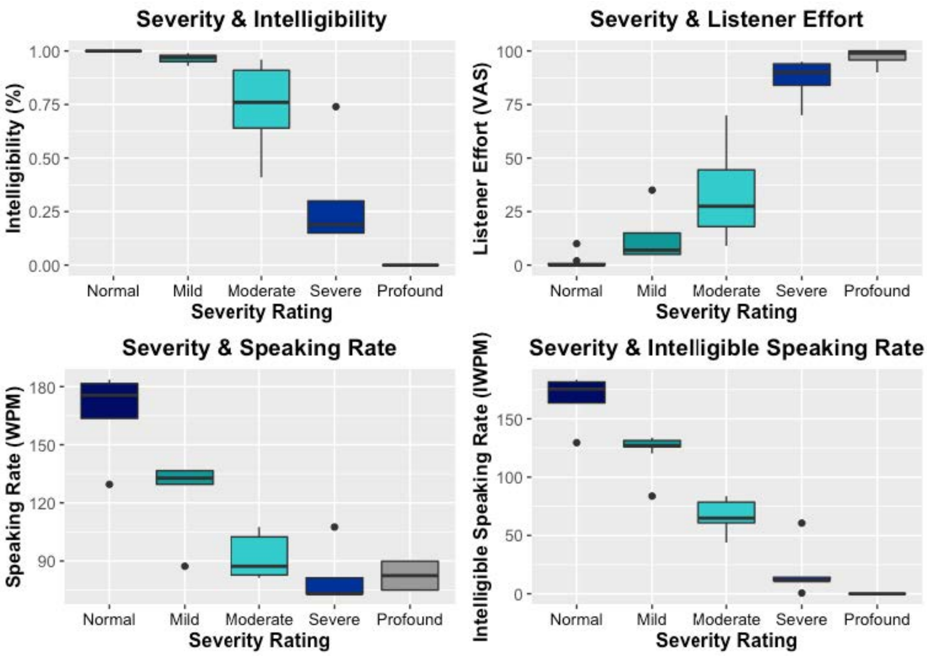
ASSESSING DYSARTHRIA SEVERITY OF REMOTE SPEECH RECORDINGS

AE Haenssler[1], JR Green[1], M Maffei[1], Z Scheier[3], A Clark[3], KM Burke[3], JP Onnela[2], JD Berry[3], KP Connaghan[1]
[1]*MGH Institute of Health Professions, Boston, Massachusetts, United States of America*
[2]*Harvard T.H. Chan School of Public Health, Boston, Massachusetts, United States of America*
[3]*Massachusetts General Hospital, Boston, Massachusetts, United States of America*

Introduction Efficacious approaches to rating dysarthria severity are important for both speech-language intervention and research. Though speech severity is generally quantified using adjectival labels (e.g., normal-profound) (King et al., 2012), the reliability and validity of these approaches is unclear due to the lack of standard definitions and methodology (Stipancic et al., 2021). A recent study sought to address this gap by examining the reliability and validity of clinician ratings of dysarthria severity, as well as the potential of other clinical measures (e.g., intelligibility, speaking rate, listener effort) to serve as severity-surrogate measures (Stipancic et al., 2021). While the findings indicated good reliability and validity across measures, the efficacy of these measures across speech recording methods and environments is unclear. The considerable growth in mHealth practices has increased access to clinical and research opportunities, while providing more ecologically valid settings than clinics and labs. However, little is known about the impact of field, as opposed to laboratory, recording conditions on the validity and reliability of clinical ratings. Therefore, the current study was designed to investigate the reliability and validity of clinician severity and severity-surrogate ratings from remotely acquired recordings using a smartphone app.

Methods Clinical severity ratings of speech recordings from 12 people with dysarthria due to amyotrophic lateral sclerosis have been collected to date. The speakers recorded themselves reading a short passage using the Beiwe smartphone research platform (Onnela et al., 2021). Offline, speaking rate was extracted and the recordings were combined with multispeaker babble at a -1-dB signal-to-noise-ratio. Thus far, three speech-language pathologists have independently transcribed and rated the samples. The rating tasks included 1) speech severity on a five-point scale, and 2) listener effort using a visual analogue scale. Analysis of the full data set will include evaluation of reliability of severity and severity-surrogate (intelligibility, intelligible speaking rate, and listener effort) measures using intraclass correlation coefficients. Validity of the measures will be evaluated by correlations across measures. In addition, the ability of severity-surrogate measures to differentiate between severity levels will be determined using one-way ANOVAs.

Results & Discussion Preliminary findings suggest high inter-rater reliability across the listener severity ratings. Ratings from at least two listeners were in agreement for all speakers, while the ratings of all listeners were in agreement for nine speakers. Significant correlation coefficients (r; $p<0.05$) indicated strong associations between severity ratings and the severity-surrogate measures. Visual inspection of the data (Figure 1) suggests that the severity-surrogate measures differentiated between some severity levels, with intelligible speaking rate particularly promising for differentiating across all levels. Overall, our preliminary findings suggest that the validity and reliability of severity and related measures documented in previous work (Stipancic et al., 2021) are replicated using recordings collected over a smartphone. Establishing the efficacy of these clinical approaches for remotely collected data promises to facilitate clinical and research participation by people with dysarthria.



Boxplots of severity-surrogate measures across listener-rated dysarthria severity levels

References

King, J. M., Watson, M., & Lof, G. L. (2012). Practice patterns of speech-language pathologists assessing intelligibility of dysarthric speech. *Journal of Medical Speech-Language Pathology, 20*(1), 1–17.

Onnela et al., (2021). Beiwe: A data collection platform for high-throughput digital phenotyping. *Journal of Open Source Software, 6*(68), 3417.

Stipancic, K. L., Palmer, K. M., Rowe, H. P., Yunusova, Y., Berry, J. D., & Green, J. R. (2021). "You say severe, I say mild": Toward an empirical classification of dysarthria severity. *Journal of Speech, Language and Hearing Research, 64*, 4719-4725.

MOTOR SKILL ACQUISITION: INSIGHTS FROM STUDIES OF IMPLICIT AND EXPLICIT LEARNING IN ADULTS WHO DO AND DO NOT STUTTER

Fiona Höbler[1], Tali Bitan[2], Luc Tremblay[1], Luc de Nil[1]
[1] *University of Toronto, Toronto, Canada*
[2] *University of Haifa, Haifa, Israel*

Complex motor skills are acquired under explicit and implicit conditions, depending on the stage of the learning process, the constraints on our cognitive capacities, as well as on the developmental stage during which they are acquired. When treating children and adults who experience speech motor difficulties, the majority of intervention approaches are delivered under explicit learning conditions, whereas successful maintenance of treatment outcomes often also relies on implicit learning and the automatization of speech skills. This research sought to investigate the nature of motor skill acquisition in adults who do (AWS) and do not stutter (ANS), and how implicit and explicit learning processes may differ between these individuals.

In our first study, 15 AWS and 15 ANS completed an implicit Alternating Serial Reaction Time (ASRT) task (Nissen & Bullemer, 1987) across two sessions, separated by 24 hours. The ASRT involved a 10-element probabilistic sequence, requiring participants to learn higher-order associations between alternating sequence elements. On the ASRT task, AWS were found to perform the implicit motor sequences more slowly, but with greater accuracy overall, than ANS. Male AWS and ANS groups differed significantly in terms of sequence-specific learning, while the female AWS group performed similarly to their ANS counterparts during the two sessions.

In the second study, 18 AWS and 18 ANS performed an explicit Finger-to-thumb Opposition Sequencing (FOS) task (Korman et al., 2007). This also involved two sessions separated by 24 hours, with participants receiving explicit instruction of the sequence. On the FOS task, no significant differences in explicit motor learning were found between AWS and ANS, either in performance across blocks or in consolidation between day one and day two of practice.

All participants completed basic measures of cognitive processing, including short-term (STM) and working memory (WM), as well as sustained attention. Cognitive measures revealed that the AWS and ANS groups in both studies exhibited lower WM scores, which were found to negatively correlate with ASRT performance variables among AWS participants.

These results demonstrated motor learning differences between AWS and ANS during implicit motor sequence learning, and most specifically among male AWS. No differences were found between the groups in explicit motor learning. Further research is warranted in individuals at earlier stages of motor development. If confirmed in children who stutter, these differences may point to the role of motor learning inefficiencies during early development, and may further our understanding of the incidence and persistence of speech motor disorders, in which sex differences can also play a role. Additionally, a better understanding of explicit and implicit learning processes, as well as increased awareness of the client's developmental and cognitive capacities, can help guide the appropriate focus and individualization of fluency therapies in the clinical setting.

References

Korman, M., Doyon, J., Doljansky, J., Carrier, J., Dagan, Y., & Karni, A. (2007). Daytime sleep condenses the time course of motor memory consolidation. *Nature Neuroscience, 10*(9), 1206–1213. https://doi.org/10.1038/nn1959

Nissen, M. J., & Bullemer, P. (1987). Attentional requirements of learning: Evidence from performance measures. *Cognitive Psychology, 19*(1), 1–32. https://doi.org/10.1016/0010-0285(87)90002-8

VALIDATION OF SELF-REPORTED PHENOTYPIC DATA FROM A GENOME-WIDE ASSOCIATION STUDY OF AUSTRALIANS WHO STUTTER

Sarah Horton[1,2], Ingrid Scheffer[3,4], Angela Morgan[1,2]
[1]*Murdoch Children's Research Institute, Parkville, Australia*
[2]*University of Melbourne, Melbourne, Australia*
[3]*University of Melbourne, Austin Health, Heidelberg, Australia*
[4]*Florey Institute of Neuroscience and Mental Health, Parkville, Australia*

Introduction Within large, population-level studies such as genome-wide association studies (GWAS), it is not always feasible to undergo 1:1 clinical assessment of all participants, resulting in reliance on self-report data. Boyce et al. (2022) reported stuttering frequency and severity in a cohort of 987 Australians enrolled in the Genetics of Stuttering GWAS. Following on from Boyce et al. (2022), the present study sought to validate participants' self-reported stuttering severity rating scores with 1:1 clinical speech pathology assessment. Here we also determined to examine (i) severity rating accuracy between participants with and without prior history of speech pathology intervention, and (ii) the association between objective stuttering severity and subjective impact of stuttering. It was hypothesised that there would be no significant difference in stuttering severity ratings between speech pathologists, participants who had accessed speech pathology intervention, and participants who had not accessed speech pathology intervention. It was also hypothesised that higher stuttering severity would correlate with higher subjective impact of stuttering.

Methods One hundred and thirty-seven Australians (42 female) aged six to 84 years (mean age 45.24 years) were recruited from a stuttering GWAS. Data were collected via online survey and face-to-face telehealth interview. Prior to interview, participants completed the OASES survey (Yaruss & Quesal, 2006) and indicated whether they had accessed previous speech pathology intervention. Following a conversational speech sample, speech pathologists and participants rated stuttering severity on a 10-point scale from no stuttering at all to extremely severe stuttering (Reilly et al., 2009). A t-test was used to investigate difference between speech pathologist and participant ratings, and between participant ratings with and without a history of speech pathology intervention. A Pearson correlation was used to investigate association between objective stuttering severity rating and subjective impact of stuttering, as measured by the OASES overall impact score.

Results & discussion There was no significant difference between speech pathologist and parent- or self-reported ratings of participant stuttering severity ($p = .296$). 81.16% of parent- and self-report ratings were within 1 point of the speech pathologist's rating on the 10-point scale, indicating clinical agreement (Onslow, Packman & Harrison, 2003). There was no significant difference in agreement of severity ratings between participants with or without a history of speech pathology intervention ($p = .113$ and $.784$, respectively), indicating that even participants who were not familiar with the rating scale were able to apply it accurately. Finally, there was a significant positive correlation between speech pathologist rated stuttering severity and overall impact score on the OASES survey ($r = .37$, $p = 0.001$), demonstrating an association between the two variables. The medium effect size indicates that subjective impact of stuttering may also be influenced by a range of other factors that were not assessed in this study. Overall, these findings indicate that self-report data appears to be a suitable method for collecting phenotype information for large cohorts of participants where individual assessment is not feasible.

References

Boyce, J.O., Jackson, V.E., Van Reyk, O., Parker, R., Vogel, A.P., Eising, E., Horton, S.E., Gillespie, N.A., Scheffer, I.E., Amor, D.J., & Hildebrand, M.S. (2022). The nature and impact of developmental stuttering across the lifespan, as reported by study participants. *Developmental Medicine & Child Neurology.*

Onslow, M., Packman, A., & Harrison, R.E. (2003). *The Lidcombe Program of early stuttering intervention: A clinician's guide.*

Reilly, S., Onslow, M., Packman, A., Wake, M., Bavin, E.L., Prior, M., Eadie, P., Cini, E., Bolzonello, C. & Ukoumunne, O.C. (2009). Predicting stuttering onset by the age of 3 years: A prospective, community cohort study. *Pediatrics, 123*(1), pp.270-277.

Yaruss, J.S. & Quesal, R.W. (2006). Overall Assessment of the Speaker's Experience of Stuttering (OASES): Documenting multiple outcomes in stuttering treatment. *Journal of fluency disorders, 31*(2), pp.90-115.

THE IMPACT OF COGNITIVE SYMPTOMS ON COPING WITH BULBAR AMYOTROPHIC LATERAL SCLEROSIS (ALS): PERCEIVED NEEDS FOR BULBAR SYMPTOM MANAGEMENT

Anna Huynh[1,2,3], Kerry Adams[4], Carolina Barnett-Tapia[1], Sanjay Kalra[5], Lorne Zinman[1,2], Yana Yunusova[1,2,3]
[1]*University of Toronto, Toronto, Canada*
[2]*Sunnybrook Health Sciences Centre, Toronto Canada*
[3]*KITE-Toronto Rehabilitation Institute, University Health Network, Toronto, Canada*
[4]*Alberta Health Services, Edmonton, Canada*
[5]*University of Alberta, Edmonton, Canada*

Background Bulbar symptoms are frequently associated with cognitive symptoms in ALS. The presence of bulbar and/or cognitive symptoms is further associated with a more debilitating disease presentation and poorer quality of life (Chio et al., 2009). As there is no cure for ALS, current practices are focused on symptom management. Existing guidelines for bulbar symptom management provide limited guidance to Speech-Language Pathologists (SLPs) for patients presenting with bulbar and cognitive symptoms (Pattee et al., 2019). Multiple calls to action have been made to ensure that clinical management is informed by patient and caregiver's priorities (Gwathmey & Berggren, 2019). However, there is an absence of research on experiences and needs of patients with bulbar and cognitive symptoms.

Objectives To describe the problems and coping approaches of patients with bulbar and cognitive symptoms; To identify the perceived needs of these patients for clinical management.

Methods Qualitative study design was informed by interpretive description. Seven interviews were conducted with patients and their caregivers. Purposive sampling was used to recruit patients with documented bulbar disease and a range of cognitive presentations. All patients had severe dysarthria and moderate-severe dysphagia. Four patients presented with mild to moderate cognitive symptoms based on cognitive test scores; three patients met the criteria for frontotemporal dementia (FTD). Thematic analysis was used to analyze interview data.

Results Participants identified problems related to bulbar and cognitive symptoms and healthcare services. Those with FTD experienced more problems participating in their daily lives and interacting with healthcare services. While patients without FTD engaged in coping strategies more independently, patients with FTD relied heavily on their caregivers to cope with changes in bulbar and cognitive function. Participants used three types of coping approaches: (i) problem-focused strategies such as seeking information; (ii) emotion-focused strategies such as avoiding challenging situations; and (iii) meaning-focused strategies such as attributing problems to other reasons. Participants also identified four needs for clinical support: (i) informational needs such as personalized information; (ii) instrumental needs such as navigational support; (iii) appraisal needs such as regular feedback; and (iv) emotional needs such as counselling. Those in the FTD group have a greater need for cognitive-specific information and intervention.

Discussion Patients with bulbar ALS have complex needs for bulbar symptom management, which becomes further complicated by the presence of significant cognitive disease (i.e., FTD). While patients and caregivers are making use of their existing resources to cope with bulbar symptoms, they are having difficulty adjusting to the rapid disease course, especially when cognitive symptoms are present. Patients and their families are expressing a need for more direct support with bulbar symptom management with greater consideration of cognitive symptoms from clinicians. There is a continuing need for education on the scope of SLP practice in the context of ALS multidisciplinary clinic.

References

Chiò, A., Logroscino, G., Hardiman, O., Swingler, R., Mitchell, D., Beghi, E., et al. (2009). Prognostic factors in ALS: A critical review. *Amyotrophic Lateral Sclerosis, 10*(6-5), pp. 310-323. https://www.tandfonline.com/doi/full/10.3109/17482960802566824

Pattee, G. L., Plowman, E. K., Garand, K. L., Costello, J., Brooks, B. R., Berry, J. D., et al. Provisional best practices guidelines for the evaluation of bulbar dysfunction in amyotrophic lateral sclerosis. *Muscle Nerve, 59*(5), pp. 531-536. https://onlinelibrary.wiley.com/doi/abs/10.1002/mus.26408

Gwathmey, K. G., & Berggren, K.N. (2019). Reflections on the implementation of the provisional best practices guidelines for the evaluation of bulbar dysfunction in ALS. *Muscle Nerve, 59*(5), pp. 523–4. https://onlinelibrary.wiley.com/doi/abs/10.1002/mus.26453

SPEECH DIADOCHOKINESIS IN STROKE SURVIVORS: EVALUATION OF AUTOMATED ANALYSIS PROCEDURES

Adam Jacks, Soomin Kim, Katarina Haley
University of North Carolina at Chapel Hill, Chapel Hill, NC, United States of America

Background Speech diadochokinesis (DDK) tasks are used to evaluate the speed and regularity of speechlike movements in both adults and children. Tanchip and colleagues recently evaluated the validity of automated algorithms for quantifying syllables in DDK tasks in patients with ALS (Tanchip et al., 2021). Little evidence is available on DDK performance in stroke survivors, although it is commonly suggested that some tasks help differentiate acquired apraxia of speech (AOS) from dysarthria (Duffy, 2020). In addition, Tanchip's validation study included only AMR and not SMR tasks, which are considered particularly relevant for AOS. The purpose of this study is to evaluate the accuracy and characterize performance on AMRs and SMRs in chronic stroke survivors using algorithms validated in people with ALS.

Method Ninety-one stroke survivors produced three AMR and one SMR tasks as part of a motor speech evaluation, for a total of 364 samples. The samples were analyzed with the energy algorithm detailed by Tanchip et al. (2021), and a custom script was used to display syllable boundaries in Praat TextGrids and calculate syllable rate, coefficient of variation of syllable durations (COVsyll), and scanning index (Ackermann & Hertrich, 1993). Ten percent of the sample were analyzed manually to determine accuracy of syllable counting. Additional work to be completed by the time of the conference includes: 1) compare manual vs. automated accuracy for the full sample, 2) complete perceptual ratings of rate consistency for comparison against scanning index and COVsyll, 3) compare results for participants with different speech profiles (e.g AOS, aphasia with phonemic paraphasia).

Results Syllable rates for all tasks were lower than reported in normative samples. Median AMRs were as follows for /p/: 4.19 syll/sec (IQR: 3.41-4.83), /t/: 3.92 (IQR: 3.27-4.74), /k/: 3.75 (IQR: 2.95-4.72); median rate for SMR was 2.69 syll/sec (IQR: 2.04-3.53). Of 48 samples (12 participants) re-analyzed manually for comparison with automated syllable counting, we found 21 samples (44%) where there was no difference between manual and automated counts. Across all rescored participants, the median deviation was 0 syllables and average deviation was 3.5 syllables (manual identified on average 3.5 more syllables than automated). In a few outlier cases the automated process missed substantially more syllables—this was particularly notable for weak syllables produced by some faster speakers (e.g. > 6 syll/sec) and those producing variant prosodic patterns (e.g. one participant used a strong weak pattern, PUH-puh-PUH-puh).

Discussion Automated algorithms that have been validated for identifying and quantifying syllables in DDK tasks in patients with ALS may be appropriate for similar purposes in stroke survivors with motor speech impairment. However, we found discrepancies between manual and automated counting, particularly for fast speakers, those using a variant prosodic pattern (e.g. strong-weak), and in the SMR (puh-tuh-kuh) task which was not used in Tanchip's validation study. The findings have implications for administration instructions, for example re-instructing participants who produce non-target prosodic patterns. Additional analyses in progress may shed light on the potential of acoustic metrics to quantify rate regularity (e.g. scanning index, COVsyll).

References

Ackermann, H., & Hertrich, I. (1993). Dysarthria in Friedreich's ataxia: Timing of speech segments. *Clinical linguistics & phonetics, 7*(1), 75-91.

Duffy, J. R. (2020). *Motor speech disorders: Substrates, differential diagnosis, and management* (4th Edition). Elsevier Health Sciences.

Tanchip, C., Guarin, D. L., McKinlay, S., Barnett, C., Kalra, S., Genge, A., ... & Yunusova, Y. (2022). Validating Automatic Diadochokinesis Analysis Methods Across Dysarthria Severity and Syllable Task in Amyotrophic Lateral Sclerosis. *Journal of Speech, Language, and Hearing Research, 65*(3), 940-953.

ELEVATED GLOBAL RESPONSE INHIBITION UNDERLIES STUTTERED SPEECH

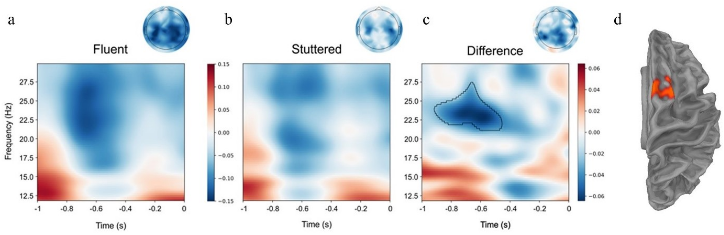
Eric S. Jackson[1], Joan Orpella[1], Graham Flick[1], M. Florencia Assaneo[2]
[1]*New York University, New York, United States of America*
[2]*National Autonomous University of Mexico, Ciudad de México, Mexico*

Introduction Stuttering is a neurodevelopmental communication disorder that manifests itself, most saliently, as intermittent interruptions in speech. Despite progress towards discovering structural and functional abnormalities in the brains of stutterers, little is known about the neural bases of stuttered speech. This is because neuroimaging studies have focused on activation associated with fluent speech in stutterers, primarily due to the difficulty of reliably eliciting stuttered speech during laboratory testing. We recently introduced a method for eliciting a comparable amount of stuttered and fluent speech during neuroimaging1. Using functional near-infrared spectroscopy, we provided evidence that a preparatory mechanism in the right hemisphere frontoparietal control network, which is related to inhibitory control, differentiates stuttered from fluent speech2. However, fNIRS relies on the "slow" hemodynamic response, leaving the timing/dynamics of speech preparation in stuttering underspecified. Here, we report results from a magnetoencephalography (MEG) study, allowing for greater specification of the temporal dynamics associated with stuttered speech. We tested the hypothesis that stuttered speech results from an overly active global inhibition mechanism3–5.

Experimental Design Twenty-nine adult stutterers participated. There were two experimental sessions: (1) Stuttering assessment and clinical interview following Jackson et al.1 to obtain participant-specific anticipated word lists likely to elicit stuttered speech. (2) MEG procedure during which participants produced 300 words (50 words repeated six times) from the list constructed during the clinical interview, to best ensure a balanced amount of stuttered and fluent trials. Words were presented on a screen one at a time followed by a pre-cue (one second before go), and then a cue to speak.

Data Analysis Trials were categorized into stuttered or fluent. (1) To determine differences in Beta power between stuttered and fluent trials, we first conducted a time-frequency decomposition separately for each type trial (stuttered and fluent, equalized in counts). The decomposition was performed for frequencies between 12 and 30 Hz, therefore spanning the entire Beta frequency band, and for times between the pre-cue presentation and the presentation of the cue to speak (-1 to 0, respectively; Fig 1a-c). A Stockwell transform (width = 0.5) was selected for the decomposition as this approach offers a good balance between temporal and spectral resolution. (2) To determine the cortical origin of the results from the time frequency analysis, we projected each participant's epochs for each condition to a source space template (fsaverage). In particular, for each participant we computed a forward model based on a 1-layer boundary element model and a minimum-norm inverse model (signal-to-noise ratio = 2; loose dipole fitting = 0.2, normal orientation) using a noise covariance matrix computed from all sensors averaged over a baseline period of 300ms across trials. This inverse model was applied to the participant's epochs (stuttered and fluent separately, equalized in counts) using dynamic statistical parameter mapping (dSPM). Once in source space (5124 sources), power spectral density was estimated for each trial of each condition using a multi-taper method. We then computed the averages for each condition from which the normalized difference in power (condition A – condition B / condition A + condition B) was estimated. Each participant's data was then morphed to a common template (MNI space) and entered into a cluster-based permutation test (one-tail) across participants ($N = 29$) with an initial t-threshold of $p < 0.01$ and a subsequent cluster threshold of $p < 0.05$.

Results Both fluent and stuttered trials showed the expected Beta suppression pattern after the presentation of the pre-cue (i.e., when the participant becomes aware of the need to produce the anticipated word) (Fig 1a-b, respectively). However, stuttered trials were characterized by greater Beta power (reduced Beta suppression) compared to fluent trials (Figure 1c). This power differential in the Beta band originated in a single cluster corresponding to the right pre-SMA (R-preSMA) (Figure 1d), a known node in the right subthalamic reactive action-stopping network. Conclusions. Results provide evidence that stuttered vs. fluent speech is associated with greater Beta power emanating from the R-preSMA, a known node in the global inhibition network. This finding is in line with proposals that stuttered speech results from overly active global response inhibition5, which interferes with the progression of successive motor programs.



Time-frequency results. The analysis window spans 1 second between the pre-cue (time -1) and the go cue (time 0). Time-frequency representation of the MEG signal during the pre-cue period in trials when subsequent speech was fluent (a), stuttered (b), and the difference (fluent minus stuttered) time-frequency plot with the significant cluster in the beta band highlighted (black contour) (c). Average topographies for frequencies 21-25 Hz and times -0.8 to -0.5 sec for each condition are plotted above the corresponding time-frequency matrices. (d): Power spectral density results in source space in axial view.

References

Arenas, R. (2017). Conceptualizing and investigating the contextual variability of stuttering: The speech and monitoring interaction (SAMI) framework. *Speech Lang Hear., 20*(1), pp. 15-28.

Hannah, R., & Aron, A. R. (2021). Towards real-world generalizability of a circuit for action-stopping. *Nat Rev Neurosci., 22*(9), pp. 538-552.

Jackson, E.S., Dravida, S., Zhang, X., Noah, J.A., Gracco, V.L., & Hirsch, J. (submitted). Activation in Right Dorsolateral Prefrontal Cortex Underlies Stuttering Anticipation.

Jackson, E.S., Gracco, V., & Zebrowski, P.M. (2020). Eliciting stuttering in laboratory contexts. *J Speech Lang Hear Res., 63*(1), pp. 143-150.

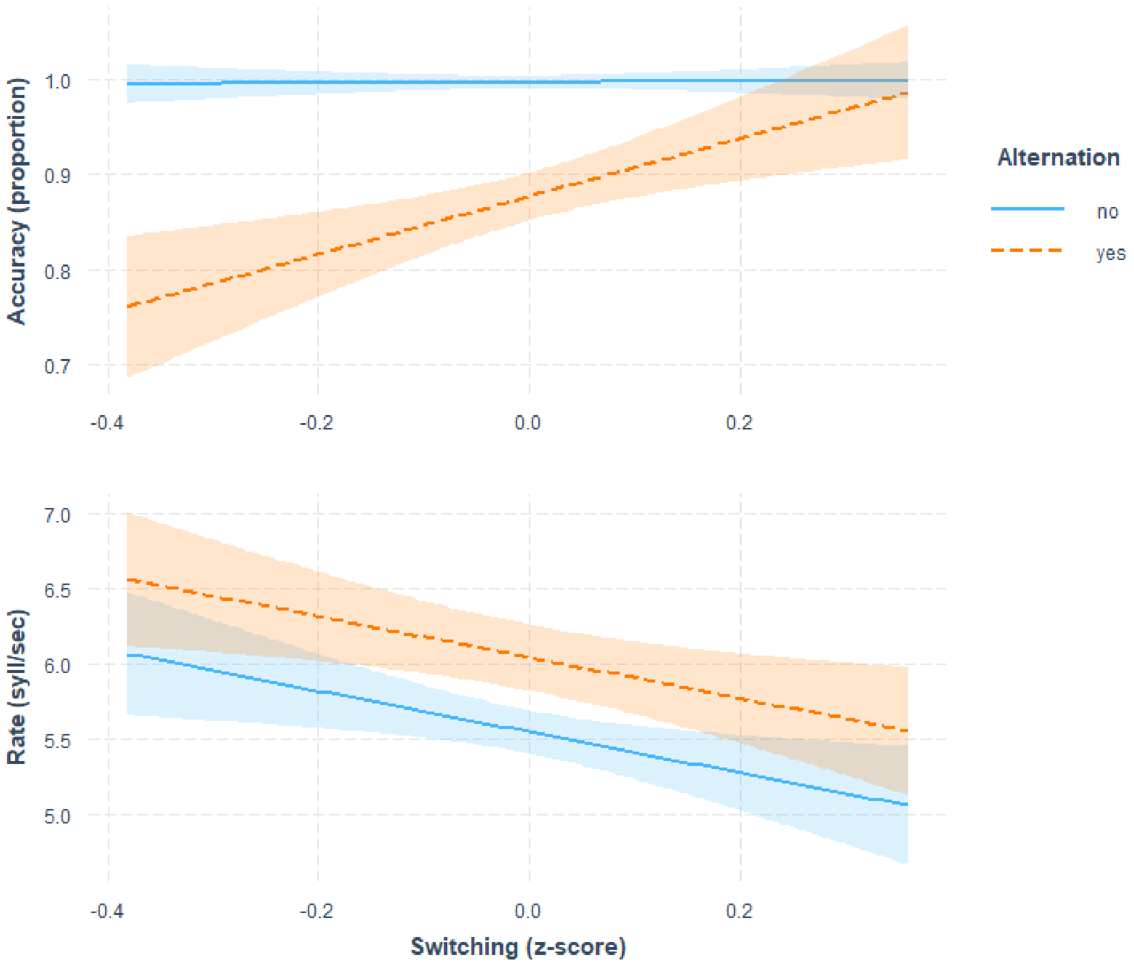
Neef, N.E., Anwander, A., Bütfering, C., et al. (2018). Structural connectivity of right frontal hyperactive areas scales with stuttering severity. *Brain, 141*(1), pp. 191-204.

DIADOCHOKINESIS PERFORMANCE AND ITS LINK TO COGNITIVE CONTROL: ALTERNATING VS. NON-ALTERNATING DDK

Esther Janse[1], Chen Shen[1,2], Esther de Kerf[1]
[1]*Radboud University, Nijmegen, the Netherlands*
[2]*Avans University of Applied Sciences, Den Bosch, the Netherlands*

The ability to produce speech at maximum performance levels, as indexed by oral diadochokinesis (DDK), is commonly used in clinical settings to evaluate speech motor control. In an earlier paper (Shen & Janse, 2020), we discussed the association between DDK performance and executive control ability, as observed in a sample of young healthy adult speakers producing *alternating* sequences (e.g., rapidly repeating nonsensical 'pataka', or 'katapa', or their real-word counterparts 'pakketten' or 'kapotte' in Dutch). More specifically, we previously tested which aspect of executive control (inhibition, updating of working memory, or switching, cf. (Miyake & Friedman, 2012) was associated with DDK performance. Results showed that particularly cognitive switching was associated with the accuracy with which alternating sequences were produced: young adults with better switching abilities showed higher DDK accuracy than those with poorer switching. To follow up on those results, we ask whether this observed association between DDK and cognitive switching is actually specific for rapid production of *alternating* sequences, or also holds for *non-alternating* sequences.

In the present study, we focused on *nonword* DDK sequences to compare performance of this same sample of 78 young healthy adult speakers of Dutch (Shen & Janse, 2020) on *alternating* ('pataka', and 'katapa') and *non-alternating* sequences ('papapa', 'tatata', and 'kakaka'). To investigate whether executive control abilities (i.e., indices of individual speakers' updating, inhibition, and switching abilities) were more strongly associated with production of alternating, as compared to non-alternating sequences, we set up separate mixed-effect regression models for DDK accuracy and rate. For the accuracy and rate models, the full models tested for interactions between the variable Alternation (yes/no) and each of the three cognitive measures (cf. Shen & Janse, 2020). We also investigated the possible link between DDK speed and accuracy, by including speed in the accuracy model (and vice versa), and by allowing this speed-accuracy relation to differ between alternating and non-alternating conditions. Participant was included as a random effect in the rate and accuracy models, as well as a random by-participant slope for the Alternation effect. These full models were then stripped in a step-wise manner, to arrive at the most parsimonious models. Of the three executive control abilities, switching was the only predictor for DDK accuracy and rate. Figure 1 shows that the association between cognitive switching and DDK *accuracy* was only observed for alternating sequences (for which accuracy was more variable than for non-alternating sequences). Rate particularly predicted accuracy of the alternating sequences, such that higher rates were associated with higher accuracy levels. The most parsimonious DDK *rate* model also included the positive association between accuracy and rate (across sequence conditions), as well as simple effects of Alternation and cognitive switching, such that those with better switching ability showed slower DDK rates. Thus, those with better cognitive switching ability showed more accurate (alternating) DDK production, and slower maximum rates for both alternating and non-alternating sequences. These combined results suggest that those with better switching have better control over their maximum speech performance, and show that the link between cognitive control and DDK performance also holds for non-alternating sequences.



Model plots depicting DDK accuracy (upper panel) and speech rate (bottom panel) as a function of speakers' switching ability (higher scores indicating better switching) for production of alternating and non-alternating DDK sequences.

References

Miyake, A., & Friedman, N. P. (2012). The nature and organization of individual differences in executive functions: Four general conclusions. *Current Directions in Psychological Science*, *21*(1), 8–14.

Shen, C. & Janse, E. (2020). Maximum speech performance and executive control in young adult speakers. *Journal of Speech, Language, and Hearing Research*, *63*(11), 3611-3627.

A SYSTEMATIC REVIEW OF THE POTENTIAL IMPACTS OF 16P11.2 DELETION SYNDROME ON MOTOR SPEECH DISORDERS

Suleyman Kahraman[1], Anja Lowit[1], Mario Parra Rodriguez[1], Suzanne Churcher[2]
[1] *University of Strathclyde, Glasgow, United Kingdom*
[2] *University of Sheffield, Sheffield, United Kingdom*

Introduction & Aims Speech disorders are a common feature in individuals with deletion carriers, including 16p11.2 deletion syndrome. However, despite this association, the relationship of this genetic disorder with speech has not been sufficiently clarified. This gap in the literature causes a challenge in both diagnosis and treatment planning as there is a lack of clinical information. This systematic review investigated the potential association between 16p11.2 deletion syndrome and motor speech disorders.

Methods Searches were performed in Web of Science, SCOPUS, PubMed, ProQuest, and MEDLINE databases after identifying keywords and inclusion-exclusion criteria under the aim of the study. Keywords consisted of speech or motor speech disorders and 16p11.2 deletion syndrome and their variances. Inclusion criteria were: the sample should consist of 16p11.2 deletion carriers, the study language must be English, published after January 1, 2010, participants with speech disorders. Exclusion criteria included studies examining the physiological characteristics of 16p11.2 deletion syndrome rather than speech dimension, participants with other syndromes besides 16p11.2 deletion syndrome, conference papers, systematic review and meta-analysis papers. The study was conducted based on the PRISMA 2020 statement. As a result of database searches, 790 studies were found. After applying the inclusion and exclusion criteria, six studies remained. The remaining six studies were evaluated for bias risk and quality assessment. After the quality assessment, it was decided that all six studies were sufficiently strong to be included in the review. A total of 227 16p11.2 deletion carriers and a control group of 235 individuals were included in the studies.

Results The results revealed the presence of a speech disorder accompanied by disturbances in the auditory feedback mechanism and language domains. With respect to the latter, 16p11.2 deletion carriers had deficits in intellectual function, language sub-domains of the Madison Speech Assessment Protocol (MSAP) and scored below their age level on the Peabody Picture Vocabulary Test-4 (PPVT-4). They also showed lower scores for syllable repetition and non-word repetition tests, percentage of consonants correct and the speech and motor sub-domains of the MSAP. In relation to the auditory feedback mechanism, reduced sensorimotor adaptation to sustained vowel identity changes was reported, and pitch compensation responses were excessive to unpredictable mid vocalization pitch perturbations. Speech problems presented as abnormalities in neuromuscular tone and oral motor movements such as tongue elevation and lateralization, reduced soft palate contraction and decreased range of motion in speech and non-speech tasks were reported. The majority of the 16p11.2 deletion carriers met the childhood apraxia of speech (CAS) criteria. However, the phenotype was different from CAS because of the presence of symptoms such as slow speech rate, equal stress and hypernasality associated with dysarthria for some participants.

Discussion & Conclusions Based on investigations to date, 16p11.2 deletion syndrome appears to frequently affect both speech and language domains. Speech disorders are primarily caused by motor coordination disorder and auditory feedback disturbances. Characterization of CAS and identifying subtypes of speech disorders in 16p11.2 deletion syndrome will play a crucial role in diagnosis and treatment planning.

References

Demopoulos, C., Kothare, H., Mizuiri, D., Henderson-Sabes, J., Fregeau, B., Tjernagel, J., Houde, J. F., Sherr, E. H., & Nagarajan, S. S. (2018). Abnormal speech motor control in individuals with 16p11. 2 deletions. *Scientific Reports, 8*(1), 1–10.

Fedorenko, E., Morgan, A., Murray, E., Cardinaux, A., Mei, C., Tager-Flusberg, 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), 310. https://doi.org/10.1038/ejhg.2015.230

Hanson, Bernier, R., Porche, K., Jackson, F. I., Goin-Kochel, R. P., Snyder, L. G., Snow, A. V., Wallace, A. S., Campe, K. L., Zhang, Y., Chen, Q., D'Angelo, D., Moreno-De-Luca, A., Orr, P. T., Boomer, K. B., Evans, D. W., Kanne, S., Berry, L., Miller, F. K., … Chung, W. K. (2015). The cognitive and behavioral phenotype of the 16p11.2 deletion in a clinically ascertained population. *Biological Psychiatry, 77*(9), 785–793. https://doi.org/10.1016/j.biopsych.2014.04.021

Hippolyte, L., Maillard, A. M., Rodriguez-Herreros, B., Pain, A., Martin-Brevet, S., Ferrari, C., Conus, P., Macé, A., Hadjikhani, N., Metspalu, A., Reigo, A., Kolk, A., Männik, K., Barker, M., Isidor, B., Le Caignec, C., Mignot, C., Schneider, L., Mottron, L., … Jacquemont, S. (2016). The Number of Genomic Copies at the 16p11.2 Locus Modulates Language, Verbal Memory, and Inhibition. *Biological Psychiatry, 80*(2), 129–139. https://doi.org/10.1016/j.biopsych.2015.10.021

Mei, C., Fedorenko, E., Amor, D. J., Boys, A., Hoeflin, C., Carew, P., Burgess, T., Fisher, S. E., "& Morgan, A. T. (2018). Deep phenotyping of speech and language skills in individuals with 16p11. 2 deletion. *European Journal of Human Genetics, 26*(5), 676.

Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., Shamseer, L., Tetzlaff, J. M., Akl, E. A., Brennan, S. E., Chou, R., Glanville, J., Grimshaw, J. M., Hróbjartsson, A., Lalu, M. M., Li, T., Loder, E. W., Mayo-Wilson, E., McDonald, S., … Moher, D. (2021). The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. P*LOS Medicine, 18*(3), e1003583. https://doi.org/10.1371/journal.pmed.1003583

Raca, G., Baas, B. S., Kirmani, S., Laffin, J. J., Jackson, C. A., Strand, E. A., Jakielski, K. J., "& Shriberg, L. D. (2013). Childhood Apraxia of Speech (CAS) in two patients with 16p11.2 microdeletion syndrome. *European Journal of Human Genetics, 21*(4), 455–459. https://doi.org/10.1038/ejhg.2012.165

PREDICTION ERRORS DRIVE AUDITORY-MOTOR ADAPTATION IN A HIERARCHICAL FACTS MODEL

Kwang Kim[1], Jessica Gaines[2], Ben Parrell[3], Vikram Ramanarayanan[4], Srikantan Nagarajan[1], John Houde[1]
[1]*University of California, San Francisco, United States of America*
[2]*University of California, Berkeley, Berkeley, California, United States of America*
[3]*University of Wisconsin–Madison, Madison, Wisconsin, United States of America*
[4]*Modality.AI, San Francisco, California, United States of America*

Numerous studies have reported that speech auditory-motor adaptation may be mostly, if not entirely, an implicit process (e.g., Keough et al., 2013) that is presumably driven by auditory prediction errors—discrepancies between predicted auditory feedback and actual auditory feedback (e.g., Kim & Max, 2021; Lametti et al., 2020). Here, the prediction error-based mechanisms involved in auditory-motor adaptation were examined via computational modeling. Specifically, we modified the recently proposed FACTS model in which state estimation mechanisms compute sensory prediction errors (Parrell et al., 2019) in order to simulate different possible adaptation mechanisms. As part of this process, we reimplemented the FACTS model in Python, replacing the CASY model of the vocal tract with the Maeda model (Maeda, 1990), that uses a modified set of constriction-based task variables as described in Gaines et al. (2021).

First, we tested whether adaptation could be driven through updates to the internal forward models that predict either a) the articulatory state, given the previous state and an efference copy of motor commands sent to the vocal tract, or b) auditory and/or somatosensory signals, given an articulatory state estimate. In the original FACTS design, neither type of model update yielded adaptive behavior in the subsequent movements. These findings suggested that adaptation may require learning that influences control policies rather than forward model updates (Hadjiosif et al., 2021).

Subsequently, we redesigned FACTS with a hierarchical architecture (e.g., Haar & Donchin, 2020) in which the task state is predicted from the articulatory state and corrected via auditory prediction errors (Fig. 1A). This task estimate forms part of the task-space control policy. Importantly, the task state prediction is generated by a Locally Weighted Projection Regression (LWPR, Klanke et al., 2008) model which can be updated for future movements based on state corrections (i.e., auditory prediction errors scaled by unscented Kalman gains). Additionally, since the Jacobian relationship between the articulatory and task states used in the transformation of desired task-space changes to articulatory-space motor commands can be derived from this LWPR mapping, this method allows us to examine how changing task-space alone or both task-space and articulatory-space control policies together influence adaptation.

To model adaptation, the LWPR articulatory-to-task mapping was updated after each trial whenever an auditory error was detected (i.e., when auditory feedback differed from the prediction). During simulations of consistent auditory perturbation (e.g., increased F1) across multiple trials, such changes in the task estimate in the control laws led to generating articulatory motor commands that anticipated the perturbation and decreased F1 across subsequent movements (Fig. 1B). The simulation also qualitatively matched previous reports including the incomplete extent of adaptation (e.g., Houde & Jordan, 1998; Kim et al., 2020; Parrell et al., 2017). In addition to the adaptive behaviors under auditory perturbation, the model also demonstrated "unlearning" behavior after the removal of the perturbation. Together, the simulations from the hierarchical architecture design demonstrate that 1) auditory prediction errors can drive speech auditory-motor adaptation and 2) adaptation is likely driven by changes to control policies rather than (only) to forward predictive models.

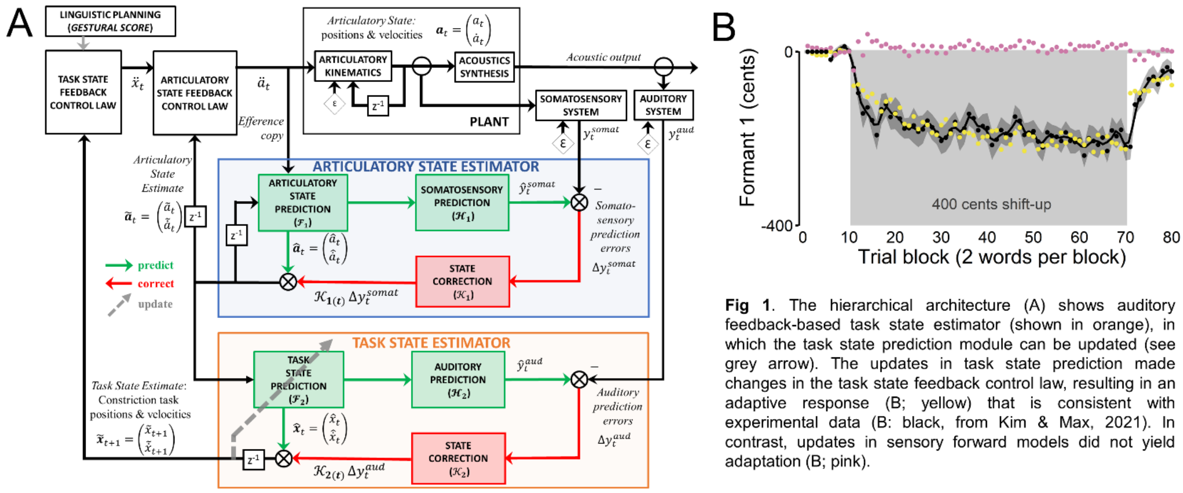


Fig 1. The hierarchical architecture (A) shows auditory feedback-based task state estimator (shown in orange), in which the task state prediction module can be updated (see grey arrow). The updates in task state prediction made changes in the task state feedback control law, resulting in an adaptive response (B; yellow) that is consistent with experimental data (B: black, from Kim & Max, 2021). In contrast, updates in sensory forward models did not yield adaptation (B; pink).

References

Gaines, J. L., Kim, K. S., Parrell, B., Ramanarayanan, V., Nagarajan, S. S., & Houde, J. F. (2021). Discrete constriction locations describe a comprehensive range of vocal tract shapes in the Maeda model. *JASA Express Letters, 1*(12), 124402. https://doi.org/10.1121/10.0009058

Haar, S., & Donchin, O. (2020). A Revised Computational Neuroanatomy for Motor Control. *Journal of Cognitive Neuroscience, 32*(10), 1823–1836. https://doi.org/10.1162/jocn_a_01602

Hadjiosif, A. M., Krakauer, J. W., & Haith, A. M. (2021). Did We Get Sensorimotor Adaptation Wrong? Implicit Adaptation as Direct Policy Updating Rather than Forward-Model-Based Learning. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience, 41*(12), 2747–2761. https://doi.org/10.1523/JNEUROSCI.2125-20.2021

Houde, J. F., & Jordan, M. I. (1998). Sensorimotor adaptation in speech production. *Science, 279*(5354), 1213–1216. https://doi.org/10.1126/science.279.5354.1213

Keough, D., Hawco, C., & Jones, J. A. (2013). Auditory-motor adaptation to frequency-altered auditory feedback occurs when participants ignore feedback. *BMC Neuroscience, 14*, 25. https://doi.org/10.1186/1471-2202-14-25

Kim, K. S., Daliri, A., Flanagan, J. R., & Max, L. (2020). Dissociated Development of Speech and Limb Sensorimotor Learning in Stuttering: Speech Auditory-motor Learning is Impaired in Both Children and Adults Who Stutter. *Neuroscience, 451*, 1–21. https://doi.org/10.1016/j.neuroscience.2020.10.014

Kim, K. S., & Max, L. (2021). Speech auditory-motor adaptation to formant-shifted feedback lacks an explicit component: Reduced adaptation in adults who stutter reflects limitations in implicit sensorimotor learning. *The European Journal of Neuroscience, 53*(9), 3093–3108. https://doi.org/10.1111/ejn.15175

Klanke, S., Vijayakumar, S., & Schaal, S. (2008). A Library for Locally Weighted Projection Regression. *Journal of Machine Learning Research, 9*(21), 623–626.

Lametti, D. R., Quek, M. Y. M., Prescott, C. B., Brittain, J.-S., & Watkins, K. E. (2020). The perils of learning to move while speaking: One-sided interference between speech and visuomotor adaptation. *Psychonomic Bulletin & Review, 27*(3), 544–552. https://doi.org/10.3758/s13423-020-01725-8

Maeda, S. (1990). Compensatory Articulation During Speech: Evidence from the Analysis and Synthesis of Vocal-Tract Shapes Using an Articulatory Model. In W. J. Hardcastle & A. Marchal (Eds.), *Speech Production and Speech Modelling* (pp. 131–149). Springer Netherlands. https://doi.org/10.1007/978-94-009-2037-8_6

Parrell, B., Agnew, Z., Nagarajan, S., Houde, J., & Ivry, R. B. (2017). Impaired Feedforward Control and Enhanced Feedback Control of Speech in Patients with Cerebellar Degeneration. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience, 37*(38), 9249–9258. https://doi.org/10.1523/JNEUROSCI.3363-16.2017

Parrell, B., Ramanarayanan, V., Nagarajan, S., & Houde, J. (2019). The FACTS model of speech motor control: Fusing state estimation and task-based control. *PLoS Computational Biology, 15*(9), e1007321. https://doi.org/10.1371/journal.pcbi.1007321

THE USE OF THE VOICE TRAINER (APP) IN ATAXIC DYSARTHRIA

Simone Knuijt, Bert de Swart, Bart van de Warrenburg, Saskia Scholten, Jorik Nonnekes, Hanneke Kalf
Radboudumc, Nijmegen, the Netherlands

Background Dysarthria is a common feature in neurodegenerative diseases. The Dutch app Voice trainer was developed based on the principles of the Pitch Limiting Voice Treatment (PLVT)(De Swart et al., 2003). This app gives people with Parkinson's disease real time feedback on loudness and pitch during exercises and can be helpful to generalize the speaking technique in daily live. The aim of this pilot study was to assess whether the Voice trainer can also be useful in the treatment of patients with ataxic dysarthria to better control their loudness and pitch.

Method Patients with a pure neurodegenerative ataxia were recruited at the department of Rehabilitation. The treatment consisted of five sessions (face or online) of 30 minutes within three weeks, to control the speaking technique in automatic sequences and spontaneous speech. Patients were encouraged to practice at home or to use the Voice trainer in situations discussed with the speech therapist. Patients were assessed before treatment (T0), directly after treatment (T1) and three months after treatment (T2). The assessment battery consisted of:

- The Dutch sentence intelligibility test (Martens et al., 2010): reading aloud 18 semantically unpredictable sentences, randomly generated from a database of 1200 sentences.
- The Radboud Dysarthria Assessment (Knuijt et al., 2017), including a perceptual rating of dysarthria severity and a short self-evaluation questionnaire.
- The Communication Participation Item Bank (CPIB-10): a short questionnaire on communicative participation.
- A short questionnaire about the usability of the Voice trainer.

Results In total, 25 patients with ataxic dysarthria were included and 21 patients completed the treatment and follow-up assessments between September 2020 and October 2021. No statistically significant change in mean intelligibility was found directly after treatment ($p=.57$) or after three months ($p=.31$), neither did the severity of dysarthria change. There were small positive changes on both the self-evaluation questionnaire and the CPIB-10, but these differences were not statistically significant. On the question how important the Voice trainer was to learn a better speaking technique, patients gave a mean VAS-score of 6.7 (range 2-10). On T2, 85% of the patients still used the app during conversations or to practice with.

Discussion & conclusion We did not find statistically significant speech changes upon treatment with the Voice trainer app. However, patients generally judged that the Voice trainer was useful in learning another speaking technique and most of them still used the app after three months. A listening experiment to evaluate the effect on naturalness of speech and an additional study with objective acoustic measurements are in progress and will be presented at the meeting.

References

Knuijt, S., Kalf, J. G., Van Engelen, B. G. M., De Swart, B. J. M., & Geurts, A. C. H. (2017). The Dutch Dysarthria Assessment: development and clinimetric evaluation. *Folia phoniatrica et logopaedica : official organ of the International Association of Logopedics and Phoniatrics, 69.*

Martens, H., Van Nuffelen, G., De Bodt, M. (2010). *Nederlands spraakverstaanbaarheidsonderzoek - zinsniveau.* Belsele: VVL.

de Swart, B. J., Willemse, S. C., Maassen, B. A., & Horstink, M. W. (2003). Improvement of voicing in patients with Parkinson's disease by speech therapy. *Neurology,60*(3), pp. 498-500.

EFFECTS OF SYLLABLE FREQUENCY ON ACCURACY AND FLUENCY IN ADULTS WITH DEVELOPMENTAL STUTTERING

Alexandra Korzeczek[1], Jana Wiechmann[2],Katharina Hammers[2], Nicole Neef[1], Arno Olthoff[1], Martin Sommer[1], Joana Cholin[2]
[1]*University Medical Centre Göttingen, Göttingen, Germany*
[2]*Bielefeld University, Bielefeld, Germany*

Developmental stuttering involves involuntary losses of speech motor control. The resulting dysfluencies might be attributed to discoordinated linguistic and motor planning levels. At the interface of these two levels, abstract phonological syllables are transformed into motor programs, as postulated by psycholinguistic speech production models (e.g. Levelt et al., 1999). Prior experimental evidence suggests that fluent speakers are faster in retrieving stored motor programs for high-frequency syllables compared to the online assembly of motor programs of low-frequency syllables (Cholin et al., 2006). Similarly, speakers with Apraxia of Speech produce less errors in high-frequency compared to low-frequency syllables, supporting the assumption of two qualitatively different routes of motor programming (Aichert & Ziegler, 2004).

In the current study, we investigated the syllable frequency effect in developmental stuttering. 19 adults with developmental stuttering (AWS) and 19 fluently speaking controls read aloud 300 German pseudowords (two-, three-, four-syllabic) whose initial syllables were either of high- or low-frequency (e.g. Fliegeuslürei /fliː-gɔɪs-lyː-raɪ/ vs. Flahgeuslürei /flaː-gɔɪs-lyː-raɪ/). As dependent variables, we measured fluency (the percentage of stuttering events such as speech blocks, repetitions, and prolongations) and accuracy (the percentage of phonological/phonetic errors such as additions, elisions, substitutions, and metathesis of consonants and vowels). Results showed no effect of syllable frequency on the percentage of stuttering events in AWS. By contrast, AWS produced more errors on pseudowords with high-frequency first syllables, while controls produced more errors on pseudowords with low-frequency first syllables. Our finding suggests that the fast access or retrieval of stored motor programs of high-frequency syllables is unreliable in AWS. One possible explanation might be an affected speech motor automaticity (e.g. Smits-Bandstra et al., 2006). Within the psycholinguistic model of Levelt et al. (1999), an imbalance between the dual-route functioning could be the consequence and might contribute to speech-motor breakdowns.

References

Aichert, I., & Ziegler, W. (2004). Syllable frequency and syllable structure in apraxia of speech. *Brain and Language, 88*(1), 148–159. `https://doi.org/10.1016/S0093-934X(03)00296-7`

Cholin, J., Levelt, W. J. M., & Schiller, N. O. (2006). Effects of syllable frequency in speech production. *Cognition, 99*(2), 205–235. `https://doi.org/10.1016/j.cognition.2005.01.009`

Levelt, W. J. M., Roelofs, A., & Meyer, A. S. (1999). A theory of lexical access in speech production. *Behavioral and Brain Sciences, 22*(01). `https://doi.org/10.1017/S0140525X99001776`

Smits-Bandstra, S., De Nil, L., & Rochon, E. (2006). The transition to increased automaticity during finger sequence learning in adult males who stutter. *Journal of Fluency Disorders, 31*(1), 22-42; quiz 39-40. `https://doi.org/10.1016/j.jfludis.2005.11.004`

ASSESSMENT OF ATYPICAL SPEECH IN MULTIPLE SCLEROSIS VIA A MULTIMODAL DIALOGUE PLATFORM: AN EXPLORATORY STUDY

Hardik Kothare[1], Michael Neumann[1], Jackson Liscombe[1], Oliver Roesler[1], Doug Habberstad[1], William Burke[1], Andrew Cornish[1], Lakshmi Arbatti[2], Abhishek Hosamath[2], David Fox[1], David Pautler[1], David Suendermann-Oeft[1], Ira Shoulson[2,3], Vikram Ramanarayanan[1,4]

[1]*Modality.AI, Inc.*

[2]*Grey Matter Technologies, Inc.*

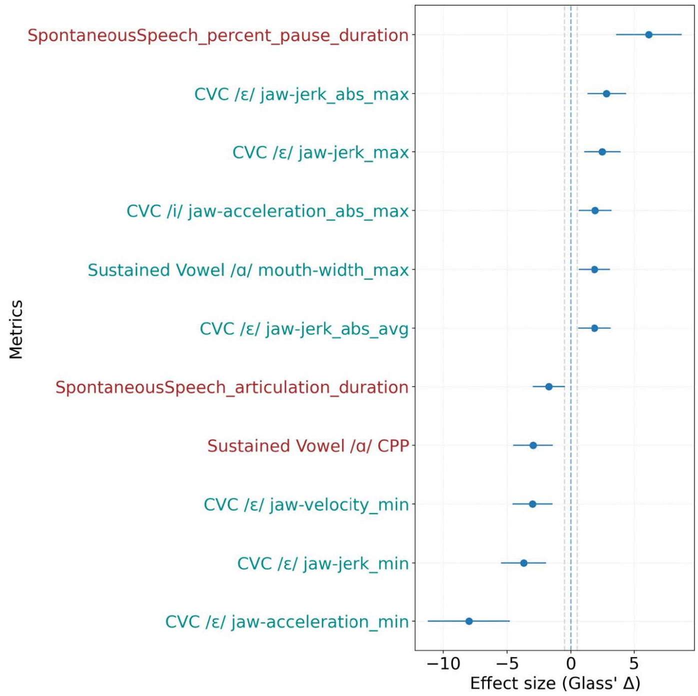
[3]*University of Rochester, Rochester, New York, United States of America*

[4]*University of California, San Francisco, California, United States of America*

The prevalence of dysarthria in Multiple Sclerosis (MS) is reported to be around 45% (Noffs et al., 2018) with most people manifesting mild severity (Gerald et al., 1987). However, impaired speech production in people with MS (PwMS) is known to impact quality of life in this population (Piacentini et al., 2014), highlighting the need to define speech-related biomarkers for remote patient monitoring and tracking disease progression and the outcomes of therapeutic interventions in MS (Noffs et al., 2018). We present an exploratory study investigating the feasibility of a novel multimodal dialogue platform (Ramanarayanan et al., 2020) with real-time extraction of speech acoustic and facial kinematic metrics in assessing impaired speech production in people with MS.

We present findings from 9 PwMS (self-reported diagnosis) and 9 age-matched healthy controls (all female, mean age = 40 ± 8 years) who took part in this ongoing study. These participants were guided by a virtual conversational agent, Tina, through a battery of tasks that elicited speech and facial behaviours like sustained vowel phonation, counting up of numbers in a single breath, repeating consonant-vowel-consonant (CVC) words, alternating-motion rate diadochokinesis, reading sentences and passages, picture description and production of spontaneous speech on a topic of their choice. A multimodal analytics module automatically extracted features that captured the acoustic and timing-related properties of speech (e.g., F0, articulation duration, articulation rate) and facial and articulatory kinematics (e.g., jaw acceleration, lip velocity) during the tasks. To calibrate for subject-specific camera setups, facial metrics in pixels were normalised for each participant by the inter-caruncular distance between the eyes in pixels. At the end of their interactive session, participants filled out three survey instruments: the short form of the Communicative Participation Item Bank (CPIB-S), the Schwab and England Activities of Daily Living scale and the Patient Report of Problems (PROP™).

We performed non-parametric Kruskal-Wallis tests at an alpha threshold of 0.01 to look at differences in metrics between PwMS and controls (see figure). We found that PwMS showed greater values of higher-order derivatives of the vertical movement of the jaw (i.e., higher acceleration and jerk, which indicate lack of smoothness in movement) during the production of /ɛ/ and /i/ CVC words. PwMS also exhibited a pattern of shorter articulatory duration during spontaneous speech production accompanied by larger percentage of pause durations. During sustained phonation of the vowel /ɑ/, patients had a wider mouth opening than controls. The cepstral peak prominence (CPP) value during this sustained vowel production was also lower in patients indicating a relative degradation in voice quality. These findings support the feasibility of assessing and monitoring objective measures of atypical speech production in MS through the use of a novel multimodal conversational technology.



Effect sizes of acoustic (brown) and facial metrics (teal) that show statistically significant differences between PwMS and controls at an alpha threshold of 0.01.

References

Gerald, Fiona J. Fitz, Bruce E. Murdoch, & Helen J. Chenery. (1987). Multiple sclerosis: Associated speech and language disorders. *Australian Journal of Human Communication Disorders 15*(2), 15-35.

Noffs, Gustavo, Thushara Perera, Scott C. Kolbe, Camille J. Shanahan, Frederique MC Boonstra, Andrew Evans, Helmut Butzkueven, Anneke van der Walt, & Adam P. Vogel. (2018). What speech can tell us: A systematic review of dysarthria characteristics in Multiple Sclerosis. *Autoimmunity reviews 17*(12): 1202-1209.

Piacentini, Valentina, Ilaria Mauri, Davide Cattaneo, Marco Gilardone, Angelo Montesano, & Antonio Schindler. (2014). Relationship between quality of life and dysarthria in patients with multiple sclerosis. *Archives of physical medicine and rehabilitation 95*(11): 2047-2054.

Ramanarayanan, V., Roesler, O., Neumann, M., Pautler, D., Habberstad, D., Cornish, A., Kothare, H., Murali, V., Liscombe, J., Schnelle-Walka, D. & Lange, P.L., 2020. Toward Remote Patient Monitoring of Speech, Video, Cognitive and Respiratory Biomarkers Using Multimodal Dialog Technology. In *INTERSPEECH* (pp. 492-493).

THE ASSOCIATION BETWEEN LONGITUDINAL DECLINES IN PHONETIC ACCURACY AND SPEECH INTELLIGIBILITY IN AMYOTROPHIC LATERAL SCLEROSIS

Hannah Rowe[1], Kaila L Stipancic[2], Thomas Campbell[3], Jordan Green[1]
[1]*MGH Institute of Health Professions, Boston, Massachusetts, United States of America*
[2]*University at Buffalo, Buffalo, United States of America*
[3]*University of Texas at Dallas, Richardson, Texas, United States of America*

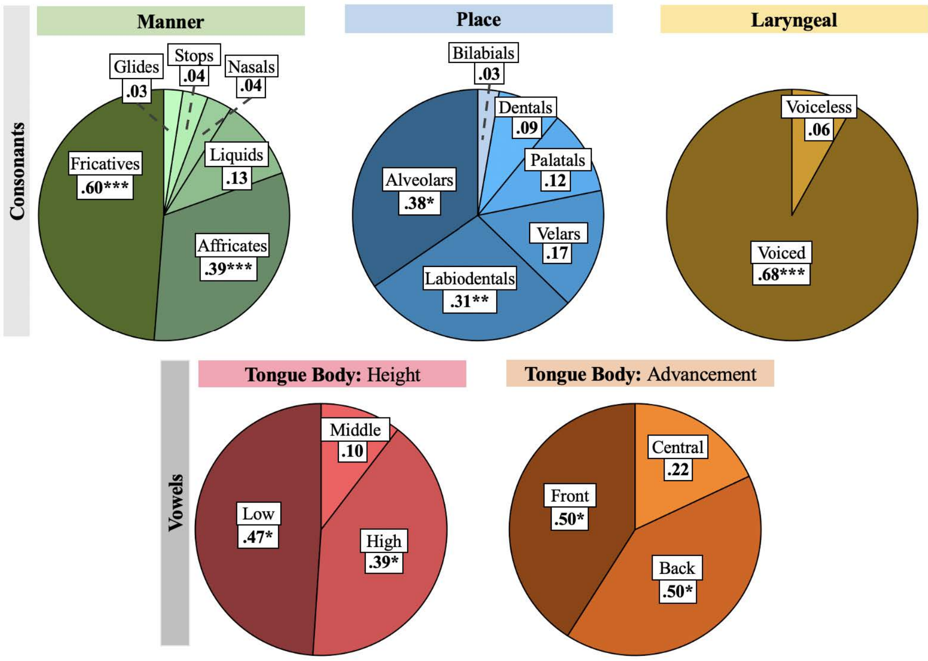
Introduction Dysarthria, defined as a neuromotor disorder of speech execution, often results in reductions in speech intelligibility. In progressive dysarthrias, such as those associated with amyotrophic lateral sclerosis (ALS), deficits in intelligibility have been linked to declines in articulatory performance [1]–[4], which has typically been characterized using speech acoustics or kinematics [5]–[7]. Despite the recent surge in research on speech motor impairments in ALS, surprisingly few studies have examined (1) how these articulatory impairments constrain speech sound production over time and (2) how the resulting speech errors degrade speech intelligibility. Information about the timing and types of phonetic errors that impact intelligibility is important for identifying efficacious treatment targets to optimize intelligibility. Prior work by Kent and colleagues examined phonetic contrasts in speakers with ALS [8], but their stimuli was limited to single words and their study was cross-sectional. In this study, we extended the work of Kent and colleagues by investigating the impacts of impaired phonetic features on intelligibility in a paragraph-reading task and examining how phonetic performance declines over time. Our research questions were:

1. Which phonetic features have the greatest impact on intelligibility in speakers with ALS?
2. What is the extent of decline for each feature over time in speakers with ALS?

Methods Twenty-one participants with ALS read the Bamboo Passage over multiple sessions, separated on average by 330.33 days. Phonemic and orthographic transcriptions were completed for all speech samples. Intelligibility was determined by calculating the percent words correctly transcribed orthographically by naïve listeners. The software PEPPER (Programs to Examine Phonetic and Phonologic Evaluation Records) was used to complete phonetic analyses. In this study, we use "phonetic features" to indicate distinctive features, which include the following feature groups: Manner (e.g., Fricatives); Place (e.g., Bilabials); Laryngeal (e.g., Voiceless consonants); and Tongue Body, which were separated into Height (e.g., Low vowels) and Advancement (e.g., Back vowels). For RQ1, we conducted linear mixed effects (LME) models for each group of features. For RQ2, we first stratified the data into the first session (i.e., Timepoint 1) and the final session (i.e., Timepoint 2). Then, to examine extent of decline, we conducted LME models for each group of features and used lsmeans() to assess the contrasts between features.

Results Intelligibility was most associated with inaccuracy in the following phonetic features (see Figure 1): Fricatives and Affricates; Alveolars and Labiodentals; Voiced consonants; Low and High vowels; and Back and Front vowels. The following features declined to the greatest extent: Affricates, Fricatives, and Stops; Labiodentals; and Voiceless consonants. There were no significant differences in the extent of decline in vowel articulation.

Discussion Taken together, our results (1) identify candidate treatment targets for speakers with reduced intelligibility due to ALS and (2) provide insights into how the underlying physiologic constraints secondary to neurodegeneration differentially impact phonemes over time. Additional work is needed to examine the decline of speech performance from disease onset and whether our results are maintained with a larger sample size.



Standardized beta coefficients illustrating the impact of each phonetic feature on intelligibility.

References

M. S. De Bodt, M. E. Hernandez-Diaz Huici, & P. H. Van De Heyning. (2002). Intelligibility as a linear combination of dimensions in dysarthric speech. *J. Commun. Disord., vol. 35*, no. 3, pp. 283–292. https://doi.org/10.1016/S0021-9924(02)00065-5

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

J. J. Sidtis, J. S. Ahn, C. Gomez, & D. Sidtis. (2011). Speech characteristics associated with three genotypes of ataxia. J. Commun. Disord., vol. 44, no. 4, pp. 478–492, Jul. 2011. https://doi.org/10.1016/j.jcomdis.2011.03.002

J. Lee, K. C. Hustad, and G. Weismer, “Predicting Speech Intelligibility With a Multiple Speech Subsystems Approach in Children With Cerebral Palsy,” J. Speech Lang. Hear. Res., vol. 57, no. 5, pp. 1666–1678, Oct. 2014, https://doi.org/10.1044/2014_JSLHR-S-13-0292

P. Rong and J. R. Green, “Predicting Speech Intelligibility Based on Spatial Tongue–Jaw Coupling in Persons With Amyotrophic Lateral Sclerosis: The Impact of Tongue Weakness and Jaw Adaptation,” J. Speech Lang. Hear. Res., vol. 62, no. 8S, pp. 3085–3103, Aug. 2019, https://doi.org/10.1044/2018_JSLHR-S-CSMC7-18-0116

Y. Yunusova, J. R. Green, L. Greenwood, J. Wang, G. L. Pattee, and L. Zinman, “Tongue Movements and Their Acoustic Consequences in Amyotrophic Lateral Sclerosis,” Folia Phoniatr. Logop., vol. 64, no. 2, pp. 94–102, 2012. https://doi.org/10.1159/000336890

J. Wang et al., "Automatic Prediction of Intelligible Speaking Rate for Individuals with ALS from Speech Acoustic and Articulatory Samples," Int. J. Speech Lang. Pathol., vol. 20, no. 6, pp. 669–679, Nov. 2018. https://doi.org/10.1080/17549507.2018.1508499

R. D. Kent, G. Weismer, J. F. Kent, and J. C. Rosenbek, "Toward Phonetic Intelligibility Testing in Dysarthria," J. Speech Hear. Disord., vol. 54, no. 4, pp. 482–499, Nov. 1989. https://doi.org/10.1044/jshd.5404.482

THE ROLE OF THE SUPPLEMENTARY MOTOR AREA IN SPEECH PRODUCTION: EVIDENCE FROM PARTICIPANTS WHO DO, AND DO NOT STUTTER

Charlotte E. E. Wiltshire, Nicole Benker, Rosa Hufschmidt, Philip Hoole
Ludwig-Maximilians-University, Munich, Germany

Introduction The basal-ganglia-thalamo-cortical network is thought to underly the coordination of speech and non-speech movements. Neuroimaging studies have revealed anatomical and functional differences in this network in people who stutter (Frankford et al., 2021; Chang & Guenther, 2020). Stuttering is markedly reduced when speaking with an external cue (such as a metronome) compared with internally cued speech (e.g. conversational speech). Two neural loops that pass through the basal ganglia may explain this "rhythm effect": An "internal timing network" comprising a basal-ganglia-SMA loop and an "external timing network" comprising a pre-motor-basal-ganglia-cerebellum loop (Alm, 2004). In this study, we will examine the hypotheses that for both people who do, and do not, stutter, the SMA is 1) involved in the coordination of speech movements and 2) is particularly sensitive to internally cued speech. We further hypothesise that these effects will be strongest in people who stutter, representing differences in the underlying function of the basal-ganglia-SMA motor loop. The rationale and methodology have been pre-registered as part of a Stage 1 registered report (`https://osf.io/hpve5/`).

Method Twenty-six participants who do stutter and 26 who do not stutter (Total 52 participants) will complete two sessions. For each session, repetitive Transcranial Magnetic Stimulation (rTMS; 0.6 Hz, 15 minutes) will be used to temporarily disrupt the function of the SMA or the hand representation of the primary motor cortex (Hand-M1; control site). Before and after rTMS, Electromagnetic Articulography (EMA) will be used to record speech movements whilst participants repeat simple speech sequences (e.g. "bi da gu") with and without a metronome, i.e. externally or internally cued speech production. The variability of speech movements over repeated utterances (coefficient of variation) will be measured. Motor Evoked Potentials (MEPs) will be elicited using TMS over Hand-M1 before and immediately after rTMS to measure changes in cortical excitability.

Results Here, we present data from the first five participants in each group. The complete data set will be presented at the conference. rTMS over Hand-M1 successfully reduced MEP amplitude (p=.004), as expected. Overall, speech variability increased after receiving rTMS to the SMA during the internal condition (p=.018) but not the external condition (p=.065). There was no change in variability for both internal and external speech conditions following rTMS to Hand-M1 (p=.55). When broken down by group, mean variability during internally and externally cued speech increased for both the stuttering and control groups following rTMS to the SMA, but neither comparison reached significance. This is perhaps not surprising given the small number of participants per group (n=5). There was no change in mean variability for either group following Hand-M1 stimulation.

Conclusion rTMS-induced disruption to the SMA impaired speech motor control during internally cued speech, but not externally cued speech. This suggests a specific role for the SMA for co-ordinating internally cued speech. Results from the full data set will further inform on whether this effect is different for people who stutter and controls.

References

Alm, P. A. (2004). Stuttering and the basal ganglia circuits: A critical review of possible relations. *Journal of Communication Disorders, 37*(4), 325–369.

Chang, S.-E., & Guenther, F. H. (2020). Involvement of the Cortico-Basal Ganglia-Thalamocortical Loop in Developmental Stuttering. *Frontiers in Psychology, 10*.

Frankford, S. A., Heller Murray, E. S., Masapollo, M., Cai, S., Tourville, J. A., Nieto-Castañón, A., & Guenther, F. H. (2021). The neural circuitry underlying the "rhythm effect" in stuttering. *Journal of Speech, Language, and Hearing Research, 64*(6S), 2325-2346.

Poster Session II

SPEECH PERCEPTION TRAINING WITH OROFACIAL SOMATOSENSORY STIMULATION AFFECTS SPEECH PRODUCTION

Monica Ashokumar[1,2,3], Jean-Luc Schwartz[1,2,3], Takayuki Ito[1,2,3,4]
[1]*Grenoble Alpes University, Grenoble, France*
[2]*GIPSA-Lab, Grenoble, France*
[3]*CNRS, Grenoble, France*
[4]*Haskins Laboratories, New Haven, United States of America*

Sensorimotor adaptation to speech motor training results in changes in speech perception (Nasir & Ostry, 2009; Shiller et al., 2009). Given the interactive nature of speech perception and production mechanisms, the reverse can also be true, that is, speech perceptual training may change speech production. Importantly, somatosensory inputs during speech motor learning appear to intervene in recalibration effects in speech perception (Ohashi & Ito, 2019). Hence, receiving a specific pair of auditory-somatosensory inputs may be a key to formulate or calibrate representations in both directions: production-to-perception and perception-to-production. In this study, we examined whether speech perception training with somatosensory stimulation can modify speech production.

We carried out a perception training paradigm based on a vowel identification task and compared the production performance of corresponding speech sounds before and after the perception training phase. Fourteen native speakers of French participated in this study. In the training, the participants were asked to identify if the French vowel they heard was /e/ (lip-spreading) or /ø/ (lip-rounding). The auditory stimuli consisted of a synthesised eight-member /e/-/ø/ continuum. Each stimulus was presented one at a time in a pseudo-random order. Along with the presentation of the vowel sounds, somatosensory stimulation associated with facial skin deformation was applied in the backward direction, which corresponds to the articulatory movement for the production of the vowel /e/. In the pre- and post-test, we recorded the corresponding vowels /e/ and /ø/ in the form of the French words 'dé' [/de/, "dice" in English] and 'deux' [/dø/, "two" in English]. The participants were asked to produce these vowels using a picture-naming task. The first, second and third formants in the recorded vowel sounds were extracted and were compared to assess a possible difference between pre- and post-tests.

We found that the third formant of the post-test was significantly increased in the vowel /e/, but not in the vowel /ø/. The first and second formants were not changed between the pre- and post-test for both vowels. Considering that lip spreading results in an increase of the third formant due to an increase of the front cavity resonance (Schwartz et al., 1993), we postulate that receiving somatosensory stimulation related to lip-spreading action for the production of /e/ along the perception training phase may induce a change in production after the training phase, resulting in an increase in the third formant value for /e/. Since orofacial somatosensory inputs associated with facial skin deformation can provide information of articulatory movement, the repetitive exposure to a pair of auditory-somatosensory stimulation related to a specific speech sound can modify the production of the corresponding speech sound. This could result from speech perceptual recalibration produced by the association of auditory and somatosensory inputs. These findings shed insights on the role of the somatosensory system in speech learning.

References

Nasir, S. M., & Ostry, D. J. (2009). Auditory plasticity and speech motor learning. *Proceedings of the National Academy of Sciences, 106*(48), 20470–20475. https://doi.org/10.1073/pnas.0907032106

Ohashi, H., & Ito, T. (2019). Recalibration of auditory perception of speech due to orofacial somatosensory inputs during speech motor adaptation. *Journal of Neurophysiology, 122*(5), 2076–2084. https://doi.org/10.1152/jn.00028.2019

Schwartz, J.-L., Beautemps, D., Abry, C., & Escudier, P. (1993). Inter-individual and cross-linguistic strategies for the production of the [i] vs. [Y] contrast. *Journal of Phonetics, 21*(4), 411–425.

Shiller, D. M., Sato, M., Gracco, V. L., & Baum, S. R. (2009). Perceptual recalibration of speech sounds following speech motor learning. *The Journal of the Acoustical Society of America, 125*(2), 1103–1113. https://doi.org/10.1121/1.3058638

ARTICULATORY CHARACTERISTICS OF PHARYNGEALIZATION IN TASHLHIYT

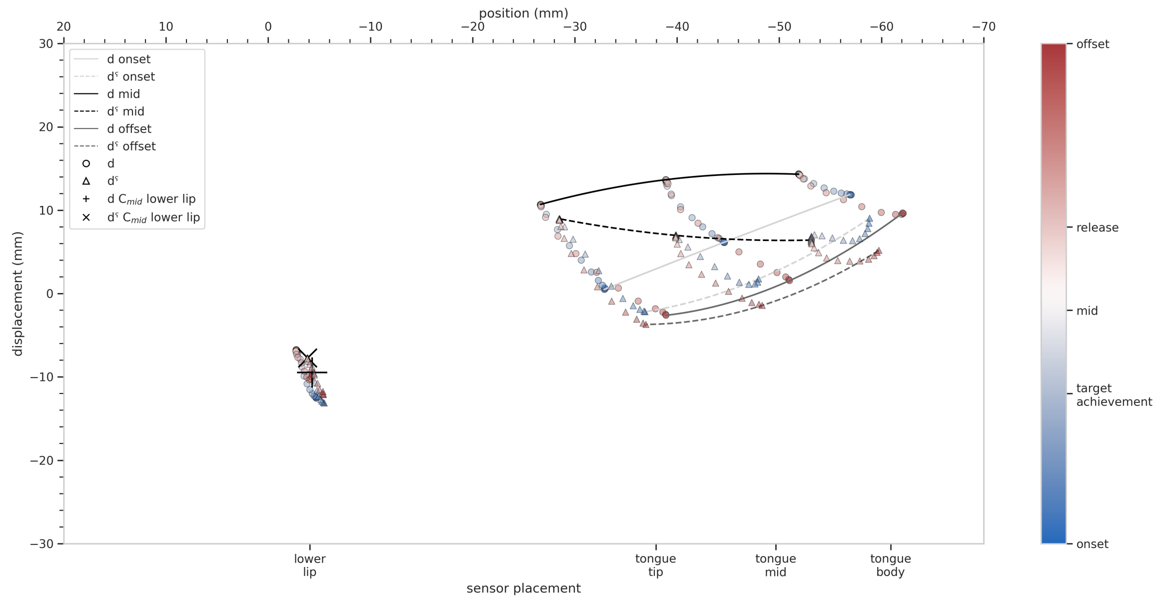
Philipp Buech, Anne Hermes, Rachid Ridouane
CNRS, Grenoble, France
University of Sorbonne Nouvelle, Paris, France

Pharyngealization is a type of secondary articulation and has been reported to mainly affect surrounding vowels (Maddieson, 1984; Ladefoged & Maddieson, 1996). The acoustic correlates of pharyngealization are well-described (mostly on Arabic varieties; cf. Al-Tamimi, 2017; Yeou, 2001)). However, articulatory studies are rarer and report on diverse mechanisms for pharyngealization, esp., regarding the tongue dorsum (e.g., being either raised, lowered, or retracted; cf. Embarki et al., 2011; Zeroual et al., 2011; Alwabari, 2020). This study provides a detailed articulatory analysis of plain and pharyngealized coronal consonants in Tashlhiyt, differing in manner (stops, fricatives, laterals, tap) and quantity (singletons vs. geminates). The results reveal that pharyngealization entails a global modification of the tongue shape/posture during the target consonants, but also the surrounding vowels.

Six male Tashlhiyt speakers were recorded using electromagnetic articulography (EMA, AG 501). Sensors were placed on the tongue tip, the tongue mid, and the tongue body. In this study, we analyze a subset of a larger corpus, consisting of [aCa] sequences containing plain [d, dː, t, tː, ɾ, ɾː, l, lː z, zː, s, sː] and pharyngealized [dˤ, dˤː, tˤ, tˤː, ɾˤ, ɾˤː, lˤ, lˤː, zˤ, zˤː, sˤ, sˤː] consonants. Articulatory annotation was done semi-automatically to identify movement onset/offset, target achievement, and release. We quantified the data by applying measurements to analyze positional tongue shape differences over the time course of the VCV sequence. The positions of the sensors were measured at each 10% from the movement onset over the target achievement and release to the movement offset. Tongue curvature was calculated at each time point based on the three tongue sensors and its degree as the height of an idealized triangle represented by the three tongue sensors.

We find important differences when comparing plain and pharyngealized consonants (Fig. 1). Most prominent is a lowering of the tongue for pharyngealized consonants across speakers and consonant manners and quantities, reflected by the sensors on the tongue mid and tongue body. For the tongue shape, we see general patterns for pharyngealization in that the curvature values are lower, but also imply speaker-specific variation. Globally, all six speakers used tongue shape modifications, but differed in their exact implementation: e.g., S6 showed a clear shift from a convex tongue shape for plain stops and laterals (singletons and geminates) to a concave shape for their pharyngealized counterparts, whereas S5 just showed a less convex shape for pharyngealized consonants (instead of a shift). For all speakers, these modifications are highest during the consonant, but are already present in the preceding vowel (i.e., at movement onset towards target achievement) and fade out into the following vowel (i.e., at release to movement offset).

Our results provide further evidence that the scope of pharyngealization is larger than the consonant itself. In addition, tongue lowering was identified as an important articulatory characteristic for pharyngealization.



Representative example of [ada] (circles) vs [adˤa] (triangles) productions (averaged across repetitions) from one speaker (S1), displaying the positions of the sensors on the lower lip (left), tongue tip, tongue mid and tongue body. Colors (from blue to red) represent the movement of the sensors from the onset (blue) to the offset of the consonant (red). The lines represent the interpolated tongue shape over tongue tip, tongue mid, and tongue body for the plain (solid lines) and pharyngealized (dashed lines) consonant at the movement onset (light grey), mid of consonant (black), and at movement offset (dark grey).

References

Al-Tamimi, J. (2017). Revisiting acoustic correlates of pharyngealization in Jordanian and Moroccan Arabic: Implications for formal representations. *Laboratory Phonology, 8*(1), 1-40. `https://doi.org/10.5334/labphon.19`

Alwabari, S. (2020). Phonological and Physiological Constraints on Assimilatory Pharyngealization in Arabic: Ultrasound Study (PhD dissertation). University of Ottawa.

Embarki, M., Ouni, S., Yeou, M., Guilleminot, C., & Maqtari, S. A. (2011). Acoustic and electromagnetic articulographic study of pharyngealisation. Coarticulatory effects as an index of stylistic and regional variation in Arabic. In Z. M. Hassan & B. Heselwood (Eds.), *Instrumental Studies in Arabic Phonetics* (pp. 193-216). John Benjamins Publishing Company.

Ladefoged, P., & Maddieson, I. (1996). *Sounds of the World's Languages.* Wiley-Blackwell.

Maddieson, I. (1984). *Patterns of Sounds.* Cambridge University Press.

Yeou, M. (2001). Pharyngealization in Arabic: Modelling, acoustic analysis, airflow and perception. *Revue de La Faculté des Lettres El Jadida, 6*, 51-70.

Zeroual, C., Esling, J. H., & Hoole, P. (2011). EMA, endoscopic, ultrasound and acoustic study of two secondary articulations in Moroccan Arabic. In Z. M. Hassan & B. Heselwood (Eds.), *Instrumental Studies in Arabic Phonetics* (pp. 277-298). John Benjamins Publishing Company.

VARIABILITY IN ANTICIPATORY V-TO-V COARTICULATION IN FRENCH

Daria D'Alessandro, Cécile Fougeron
CNRS, Grenoble, France
University of Sorbonne Nouvelle, Paris, France

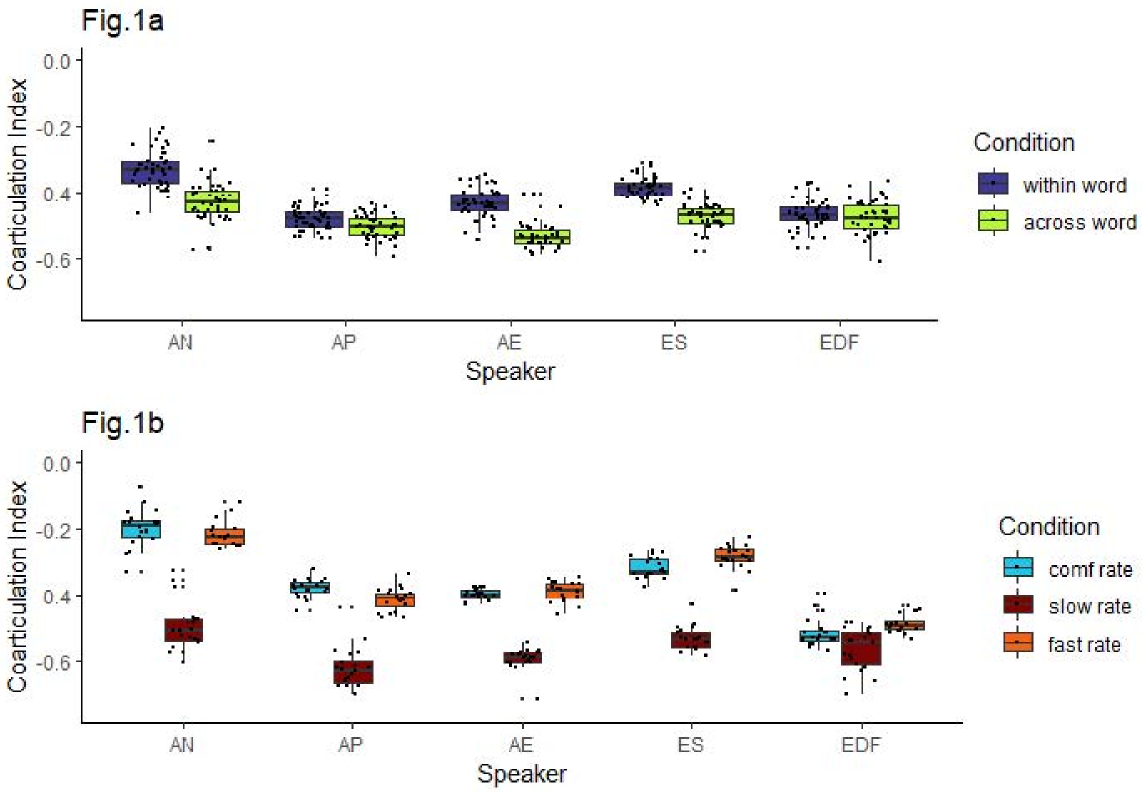
Premising that it is planned, anticipatory Vowel-to-Vowel coarticulation provides a cue on the size of the encoded speech units, since it reflects the coordination of elements that are planned together in the same unit (e.g. Whalen & Chen, 2019). Several studies have shown that coarticulation varies, e.g., according to speakers (Zellou, 2017), age (D'Alessandro & Fougeron, 2021), rate (Matthies et al., 2001), style (Scarborough & Zellou, 2013). This variability needs to be accounted for, either by explaining in which way the coordination between elements planned together can change, or by modelling the relationship between elements pertaining to different speech plans. In this study, we aim to further document the variability in anticipatory V-to-V coarticulation in French, in terms of inter- and intra- speaker variability, as well as variability according to prosody and rate.

In two production experiments, we examined anticipatory coarticulation of /i/ on /a/ in 'papi' sequences produced by five female speakers. In Experiment 1, speakers read, over a large amount of repetitions, sentences manipulating the boundary between /pa/ and /pi/, being either a syllable boundary (within the word ***papi***, 'grandpa') or a word boundary (across the words *pa**pa pi**lote*, 'dad pilots'). Each sequence occurred 45 times. In experiment 2, speakers repeated sentences including the word *papi* 20 times in a row at a comfortable rate and at either a fast (6syll/s.) or a slow (1.4syll/s.) rate following a timer on the screen (Didirkova et al., 2020). Coarticulation Index is measured as the acoustic impact of /i/ on the F2-F1 compacity of /a/ as follows: ((F2-F1)a-(F2-F1)i)/(F2-F1)i

Results for experiment 1 (Fig.1a) show that within-word coarticulation is speaker-dependent, both in terms of the amount of coarticulation (e.g., higher index for speakers AN and ES) and in terms of stability across tokens: while speakers AP and ES are stable in their coarticulation patterns, the others (especially AN) are more variable across repetitions. Comparison between within-word /papi/ and across-words coarticulation /pa#pi/ aimed at testing whether coordination (or absence of) across words translates into different coarticulation patterns. Results are also speaker-dependent: while three speakers reduce coarticulation across words, speakers EDF and AP coarticulate to the same extent across words than within a word. The within-word and across-word conditions do not differ in terms of intra-speaker variability.

In experiment 2, speech rate was manipulated to challenge the flexibility of the co-ordination between syllables within a word. The results (Fig.1b) show that none of the speakers increased coarticulation with an increase of rate (note that AP did not speak faster). However, at slow rate, coarticulation is drastically reduced for four out of five speakers, suggesting either that between-syllable coordination at slow rate is modified to reduce overlap, or that the two syllables are not encoded in the same speech plan.

This variability in V-to-V coarticulation patterns raises questions regarding (a) the flexibility in the specification or implementation of the coordination between elements in the same speech plan, (b) the flexibility in the construction of the speech plans according to uttering conditions.



Coarticulation Index per speaker and condition for Experiment 1 on the top (Fig.1a) and Experiment 2 at the bottom (Fig.1b). The higher the index, the greater the coarticulation.

References

D'Alessandro, D., & Fougeron, C. (2021). Changes in Anticipatory VtoV Coarticulation in French during Adulthood. *Languages, 6*(4), 181. `https://doi.org/10.3390/languages6040181`

Didirková I., Lancia L. & Fougeron C. Adaptations sur le F1 et le débit en réponse à diverses perturbations (2020) *Journée d'Etude sur la Parole 2020.* `https://aclanthology.org/2020.jeptalnrecital-jep.19`

Matthies M, Perrier P, Perkell JS, Zandipour M. Variation in anticipatory coarticulation with changes in clarity and rate. *Journal of Speech, Language, and Hearing Research, 44*(2):340-53 (2001). `https://doi.org/10.1044/1092-4388(2001/028)`

Scarborough, R., Zellou, G. Clarity in communication:"Clear" speech authenticity and lexical neighborhood density effects in speech production and perception. *The Journal of the Acoustical Society of America, 134*(5), 3793-3807 (2013). `https://doi.org/10.1121/1.4824120`

Zellou, G. Individual differences in the production of nasal coarticulation and perceptual compensation. *Journal of Phonetics, 61*, 13-29 (2017). doi: 10.1121/10.0000951

Whalen, D. H., & Chen, W. R. Variability and central tendencies in speech production. *Frontiers in Communication, 49* (2019). `https://doi.org/10.3389/fcomm.2019.00049`

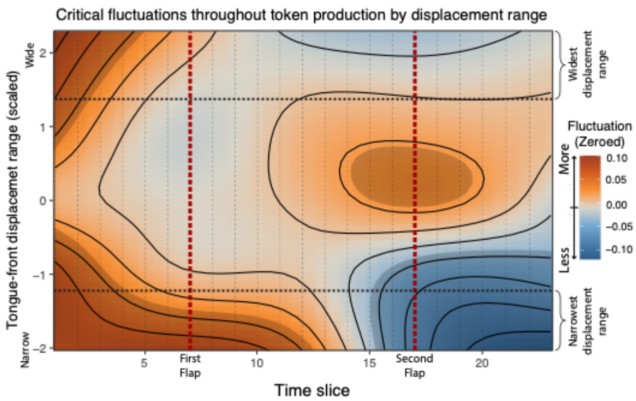
GAIT CHANGE IN TONGUE MOVEMENT IN AMERICAN AND NEW ZEALAND ENGLISH

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

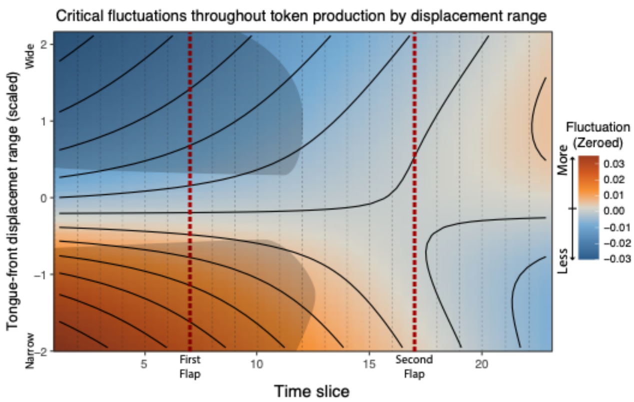
In a recent paper (Derrick & Gick, 2021), we demonstrated that gait change leading to wider movement rate ranges, analogous to switching from walking to running, occurs in the tongue during speech. North American English (NAE) speakers were recorded producing phrases with /VɾVɾV/ sequences with rhotic and non-rhotic vowels; results showed speakers who have a wide speech rate also categorically switch tongue motion patterns between slow and fast speech. We hypothesized that motor systems make use of existing learned structures to solve new problems: NAE speakers build on existing alternations between flaps/taps, vowels and rhotic tongue shapes (Derrick & Gick, 2011; Delattre & Freeman, 1968) to solve speech-rate problems. The present study hypothesizes that a language lacking this varied movement inventory will not exhibit this gait-change, instead following its standard gait, achieving as much speech rate range as it can using its more limited inventory. New Zealand English, a non-rhotic dialect of English, is one such language.

We recorded 10 New Zealand English (NZE) speakers using Electromagnetic articulography (EMA). We had them read phrases with /VɾVɾV/ sequences. Each speaker used reiterant speech ("mama**ma**mamama"), played at 3, 4, 5, 6, and 7 syllables/second, to guide their speech rate. We used a critical fluctuation analysis of tongue motion derived from Schiepek & Strunk (2010) to show effort during the time-course of speech.

The wider the speech rate range, the greater their tongue-front displacement range. For NAE (Figure 1), speakers with narrow displacement ranges had one stable speech-gait pattern, while speakers with wide tongue-front displacement ranges showed two stable patterns, as shown by end-state comfort (Schiepek % Strunk, 2010). Speakers in the middle stretched one less-stable gait pattern to achieve different speech rates, instead showing beginning-state comfort—evidence of improvisation. The NAE results indicate categorical motion solutions may emerge in any motor system, providing that system with access to wider movement-rate ranges. For NZE (Figure 2), however, speakers never produced two separate gaits: Even speakers with the widest speech-rate ranges stretched one gait to produce all their speech rates. As their tongue-front displacement ranges increased, they demonstrated less end-state comfort. These findings suggest that the speech motor system bootstraps off existing degrees of freedom in the movement inventory to allow speakers to access wider movement-rate ranges.



NAE: end-state comfort at narrowest and widest tongue-front displacement ranges (orange first), beginning-state comfort in the middle (orange last).



NZE: end-state comfort at narrowest tongue-front displacement ranges (orange first), beginning-state comfort as tongue-front displacement increases (blue first).

References

Delattre, P., & Freeman, D. (1968) A dialect study of American Rs by X-ray motion picture. *Linguistics, 44*,29-68.

Derrick, D. and Gick, B. (2011). Individual variation in English flaps and taps: A case of categorical phonetics. *Canadian Journal of Linguistics, 56*(3),307–319.

Derrick, D., & Gick, B. (2021) Gait change in tongue movement, *Scientific Reports 11*(16565), 1-14.

Rosenbaum, D. A., Vaughan, J., Barnes, H. J., & Jorgensen, M. J. (1992). Time course of movement planning: Selection of handgrips for object manipulation. *Journal of Experimental Psychology, Learning, Memory, and Cognition. 18*, 1058–1073.

Schiepek, G., & Strunk, G. (2010). The identification of critical fluctuations and phase transitions in short term and coarse-grained time series-a method for real-time monitoring of human change processes. *Biological Cybernetics, 102*, 197–207.

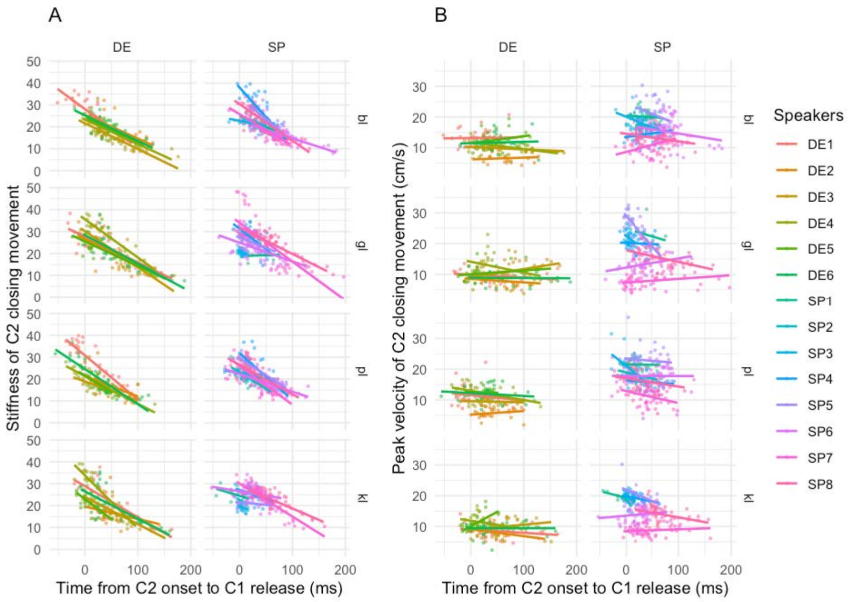
FLEXIBILITY AND ABSTRACTNESS OF ARTICULATORY COORDINATION IN GERMAN AND SPANISH WORD-INITIAL STOP-LATERAL CLUSTERS

Shihao Du[1], Adamantios I. Gafos[1,2]
[1]*University of Potsdam, Potsdam, Germany*
[2]*Haskins Laboratories, New Haven, Connecticut, United States of America*

Using data from Arabic C1C2 sequences, Gafos et al. (2020) show that the later C2 initiates its movement in reference to C1 constriction release, the higher the amplitude-normalized peak velocity (also known as stiffness; Cooke 1980, Kelso 1986) of C2's closing movement. Similar relations between stiffness and movement onset are found in other areas of motor control (for table tennis, Bootsma and van Wieringen 1990; for juggling, Beek 1989).

If the relation uncovered in Gafos et al. (2020) instantiates a general principle of coordination, then it ought to be found elsewhere. We pursue this prediction in German (DE) and Spanish (SP) word-initial /bl, pl, gl, kl/ sequences. In Electromagnetic Articulography data from 6 DE and 8 SP speakers, we assessed the relation between C2 closing movement stiffness and C2 onset-to-C1 release lag, which quantifies how early C2 initiates its closing movement in relation to C1 release. Linear-mixed-effects models with this lag as predictor and C2 closing stiffness as dependent variable confirm the presence of the predicted relation (negative correlation; -0.133, $t = -32.612$, $p < 0.0001$) across languages, clusters and speakers (Figure 1A): C2's closing stiffness is adjusted on-line in response to a flexible timing of C2 movement onset; the earlier C2's closing movement is initiated, the lower its stiffness (and vice versa). Prior work on timing in C1C2 suggests that peak velocity of C2 is a predictor of overlap between C1 and C2 (Jun 1995; 2004). Our data do not support this. As seen in Figure 1B, the lag measure, which indexes overlap in C1C2, does not covary (0.0052 cm/s, $t = -1.71$, $p > 0.05$) with C2 peak velocity (cf. Roon et al. 2021 on Moroccan Arabic).

In sum, our results provide further evidence for the relation in Gafos et al. (2020). We discuss potential interpretations for speech and other skilled action. For speech, in a C1C2, if C2 starts early within C1, then the stiffness of C2 is lowered in order to ensure that some lag in time is achieved between the C1, C2 constrictions. If C2's stiffness were not lowered, with C2 rising swiftly towards its constriction, this would result in no gap between the plateaus of C1 and C2, compromising recovery of C1's identity. Similarly, in the table tennis context (Bootsma and van Wieringen 1990), initiating an attacking stroke with less force (lower stiffness) when the ball is close(r) to the player (later movement onset) would result in failure to achieve ball-racket contact. Two properties of the uncovered relation are underscored: lawful flexibility and abstractness. The first is revealed in the degrees of freedom with which movements corresponding to any given effector begin; the second in the invariance of the observed relation with respect to the effectors involved (/bd, db, dg, gd, br, rb, kr, rk, kl, lk, lb, nk/ in Gafos et al. 2020 and stop-lateral clusters here) as well as in the presence of this relation with the more abstract control parameter of stiffness rather than peak velocity.



Scatterplots with linear regression lines showing a robust negative correlation between C2 onset-to-C1 release lag and the stiffness of C2 closing movement across languages (DE: German, SP: Spanish) and clusters (/bl, gl, pl, kl/) for each speaker (left panel A). In contrast, no robust relation is observed between C2 onset-to-C1 release lag and the peak velocity of C2 closing movement across languages, clusters, or speakers (right panel B).

References

Beek, P. J. (1989). Timing and pase locking in cascade juggling. *Ecological Psychology, 1*(1), 55–96. https://doi.org/10.1207/s15326969eco0101_4

Bootsma, R. J., & van Wieringen, P. C. W. (1990). Timing an attacking forehand drive in table tennis. *Journal of Experimental Psychology: Human Perception and Performance, 16*(1), 21–29. https://doi.org/10.1037/0096-1523.16.1.21

Cooke, J. D. (1980). The organization of simple, skilled movements. In G. Stelmach & J. Requin (Eds.), *Tutorials in motor behavior* (pp. 199-212). Amsterdam: North-Holland.

Gafos, A. I., Roeser, J., Sotiropoulou, S., Hoole, P., & Zeroual, C. (2020). Structure in mind, structure in vocal tract. *Natural Language & Linguistic Theory, 38*(1), 43–75. https://doi.org/10.1007/s11049-019-09445-y

Jun, J. (1995). Perceptual and articulatory factors in place assimilation: An optimality theoretic approach [Ph.D.]. University of California, Los Angeles.

Jun, J. (2004). Place assimilation. In B. Hayes, R. Kirchner, & D. Steriade (Eds.), *Phonetically Based Phonology* (pp. 58–86). Cambridge University Press.

Kelso, J. A. S. (1986). Pattern formation in speech and limb movements involving many degrees of freedom. *Experimental Brain Research, 15*, 105–128.

Roon, K. D., Hoole, P., Zeroual, C., Du, S., & Gafos, A. I. (2021). Stiffness and articulatory overlap in Moroccan Arabic consonant clusters. *Laboratory Phonology: Journal of the Association for Laboratory Phonology, 12*(1), 8. https://doi.org/10.5334/labphon.272

SOMATOSENSORY AND MOTOR CORTEX BOTH CAUSALLY CONTRIBUTE TO SPEECH MOTOR LEARNING

Matthias K. Franken, Timothy F. Manning, Alexandra Williams, David J. Ostry
McGill University, Quebec, Canada

Speakers readily adapt their articulation to various sensory perturbations, such as formant alterations, pitch shifts, or loudness modulations. It is thought that this sensorimotor adaptation, much like in non-speech motor learning, relies on comparing predicted sensory feedback with observed sensory feedback. However, the neural mechanisms that support this learning are poorly understood. Interestingly, there is accumulating evidence that the somatosensory system plays an active role in motor learning (Kumar et al., 2019). For example, somatosensory stimulation such as facial skin deformation may alter subsequent speech motor learning (Ito & Ostry, 2010). In the present study, we investigated whether the primary somatosensory cortex and the primary motor cortex are causally involved in speech motor learning. The hypothesis is that if a brain area participates in speech motor learning, then disruption of its activity using magnetic brain stimulation prior to learning will either reduce learning or eliminate it altogether.

Fifty participants performed a speech motor learning task, in which they produced the words ‘shame’ (/ʃeɪm/) and ‘shake’ (/ʃeɪk/) while receiving auditory feedback through headphones. After thirty baseline productions, the frequency spectrum of the initial fricative sound was shifted in real-time by 3 semitones, resulting in a shift from /ʃ/ towards /s/, and held at this value for the remaining 150 speech productions. Such altered auditory feedback typically leads to adaptation, in which speakers compensate for the altered auditory feedback by shifting their vocal output in the opposite direction. Participants performed the speech adaptation task after continuous theta-burst stimulation, which was used to disrupt activity during learning in the lip area of either primary motor cortex or primary somatosensory cortex, in a between-participants design.

The findings to date show that whereas participants show speech motor adaptation in a control condition, adaptation is blocked by disrupting either primary motor cortex or primary somatosensory cortex. Motor-evoked potentials measured at the lips before and after theta-burst stimulation in the somatosensory condition confirm that blocking adaptation in this condition was not caused by indirect effects of somatosensory stimulation on primary motor cortex.

These results show that the primary somatosensory cortex is causally involved in speech motor learning, in line with recent studies suggesting that motor learning is driven, at least in part, by the somatosensory system. In addition, this is the first study showing that both primary motor and primary somatosensory cortex causally contribute to motor learning. The current results show that even in altered auditory feedback, the somatosensory system may be crucial for updating the memory of the adaptive motor commands and their associated somatosensory representations.

References

Ito, T., & Ostry, D. J. (2010). Somatosensory Contribution to Motor Learning Due to Facial Skin Deformation. *Journal of Neurophysiology, 104*(3), 1230–1238. https://doi.org/10.1152/jn.00199.2010

Kumar, N., Manning, T. F., & Ostry, D. J. (2019). Somatosensory cortex participates in the consolidation of human motor memory. *PLOS Biology, 17*(10), e3000469. https://doi.org/10.1371/journal.pbio.3000469

BILINGUALS READILY ACQUIRE LANGUAGE SPECIFIC SENSORIMOTOR MAPS FOR SPEECH

Daniel R Lametti[1], Emma Wheeler[1], Imane Hocine[2], Douglas M Shiller[2]
[1]*Acadia University, Wolfville, Canada*
[2]*University of Montréal, Canada*

Introduction The extent to which the bilingual brain maintains shared versus separate language representations has been the focus of significant research. However, the degree to which bilinguals maintain separate articulatory phonetic representations (i.e., speech motor plans) for their first (L1) and second (L2) language remains unclear (Chang, 2019; Shiller et al. 2022). The observation that many bilinguals who learn a second language after childhood speak with an accent in their L2 seems to argue for shared speech motor plans. In this case, the accent may reflect the misapplication of L1 speech motor plans to the production of L2 speech sounds. This suggests that such bilinguals may be limited in their ability to acquire language-specific speech motor plans. Here, using real-time alterations in auditory feedback, we examine the extent to which proficient L2 French speakers can acquire language-specific speech motor plans in their L1 and L2.

Methods Thirty L1-English/L2-French bilinguals were recruited in Wolfville, Nova Scotia—an English-speaking region of Canada. Participants produced thirty different English sentences (drawn from the phonetically-balanced Harvard corpus) and their French translations, with the English and French sentences randomly interleaved. Following Lametti et al. 2018, participants spoke into a microphone and heard the sound of their own speech in real-time through headphones. After a baseline period of English and French sentence production, the first (F1) and second (F2) formants of all vowels were altered in a manner that depended on language. Half of the participants (n=15) experienced a 57 mel *decrease* in F1 and a 57 mel *increase* in F2 during English sentence production, and the opposite alteration (a 57 mel increase in F1 and a 57 mel decrease in F2) during French sentence production. The remaining participants experienced the manipulation with the opposite language correspondence—that is, a -57 mel change in F1, +57 mel change in F2 for French and a +57 mel change in F1, -57 mel change in F2 for English. The aim of the manipulation was to induce language-specific changes in speech production such that unique motor plans would be required to produce vowels in each language, including vowels that are shared between languages.

Results Adaptation to altered auditory feedback was measured as the degree of change in formant production that counteracted the F1/F2 manipulation. Robust changes in produced vowel formants were observed during English and French sentence production that depended on language. The changes offset approximately 30% of the formant alteration in each language such that vowels shared between French and English were produced with different first and second formants. No difference was observed in the amount of compensation for altered feedback in English and French, even though the participants were first language English speakers. The results suggest that proficient bilinguals can acquire language-specific speech motor plans, even for vowel sounds that are similar or identical in each language. The work supports the idea that bilinguals can acquire separate articulatory-phonetic representations for the languages they speak.

References

Chang CB (2019). The phonetics of second language learning and bilingualism. In *The Routledge Handbook of Phonetics* (pp. 427–447). Routledge.

Lametti DR, Smith HS, Watkins KE, Shiller DM (2018). Robust Sensorimotor Learning during Variable Sentence-Level Speech. Current *Biology, 28*(19):3106-3113.

Shiller DM, Bobbitt SG, Lametti DR (2022). Immediate Cross Language Transfer of Novel Articulatory Plans for Speech in Bilinguals. *PsyArXiv.*

EFFECTS OF DELAYED AUDITORY FEEDBACK DEPEND ON THE PRESENCE OF PHONETIC INFORMATION IN THE FEEDBACK SIGNAL

Monique C. Tardif, Matthias Heyne, Ashley Petitjean, Caroline Fox, Jason W. Bohland
University of Pittsburgh, Pittsburgh, United States of America

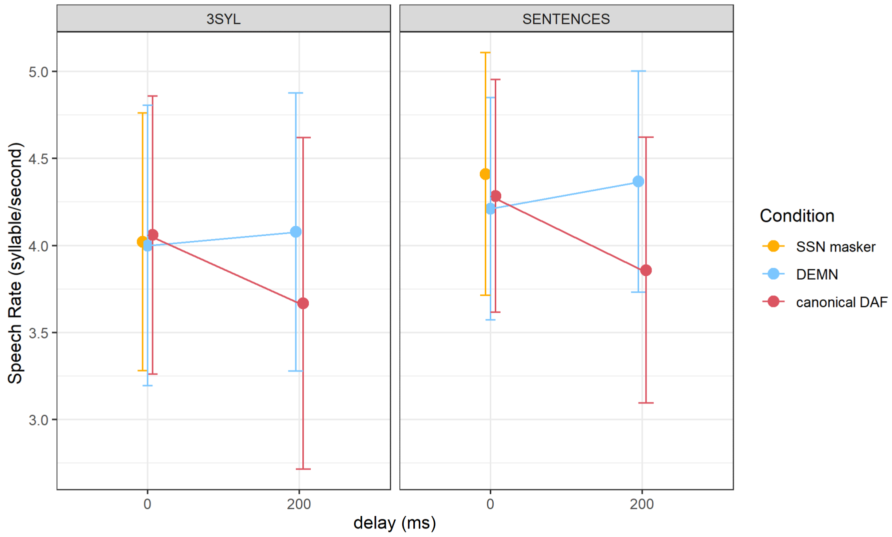
Delayed auditory feedback (DAF) profoundly affects speech, leading to reduced rate, increased intensity and pitch (Fairbanks, 1995), and speech errors (Yates, 1963). The Displaced Rhythm Hypothesis (DRH; Howell et al., 1983) proposed that only "the rhythmic properties of the syllable units, not their identity" drive these disruptions, which should be equivalent "whether the delayed signal contains any information about the phonetic identity of the sounds or not."

Some studies have provided support for the DRH using experiments that altered DAF signal content (e.g., envelope-modulated square wave; Howell & Archer, 1984) or manipulated rhythm and phonemic content in spoken stimuli (Müller et al., 2000; Kaspar & Rübeling, 2011). However, these studies did not widely vary speech materials or utilize well-matched non-speech feedback signals as a contrast with canonical DAF. Here we report speech rate changes induced by canonical DAF and delayed envelope-modulated noise (DEMN) during sentence and nonsense syllable production. We hypothesized that DEMN would be less disruptive than DAF, eliciting smaller rate reductions.

17 native English speakers (ages: 20-52, 12 women) without history of speech, language, or hearing disorders passed hearing screenings and provided the current dataset. In the syllable task, participants produced repeated 3-syllable sequences synchronized to a visual metronome. The sentence task used TIMIT sentences (Garofolo et al., 1993) presented visually and aurally, which participants repeated while approximating the rate and rhythm of the exemplar. Speech was recorded using a head-worn microphone, and auditory feedback was delivered over headphones, with latencies (0 vs 200ms) controlled using PsychPortAudio and an RME audio device. Canonical DAF signals were amplified +5 dB from input level. For DEMN, speech-shaped noise (SSN) was generated from long-term spectra of TIMIT sentences [7] and multiplied by RMS speech envelopes estimated over brief sliding windows and delivered at 90dB SPL. A constant SSN masker provided a control without phonetic or rhythmic information.

Participants completed two blocks of each task (50 sentences / 62 syllable sequences per block) with canonical DAF and one block each with a mixture of DEMN and SSN masker. Speech rate was estimated for each trial from recordings using the VocalToolkit plugin (Corretge, 2021) for Praat (Boersma & Weenink, 2018). We conservatively removed trials with rate estimates outside the median 95% per participant. A linear mixed effects model (Bates et al., 2015) with random intercept for participant was used to estimate effects of delay, task, and feedback condition on speech rate.

Figure 1 shows mean speech rates across delays, tasks and conditions. The mixed-effects model demonstrated significant main effects of delay (200 vs. 0ms) and task (sentences vs. syllables). Critically, a significant interaction between feedback condition and delay revealed that canonical DAF, but not DEMN, elicited rate reductions. Thus, our current results do not support the displaced rhythm hypothesis, though only one dependent variable (rate) has been considered. The reduced effect of DEMN on rate cannot be explained by feedback intensity differences since SSN feedback conditions had increased amplification. DAF effects on speech rate thus do not represent generic auditory interference effects but rather depend at least partly on phonetic information present in the feedback signal. These results present challenges for current models of speech production.



Mean speech rates for the syllable (left) and sentences (right) tasks as a function of delay. Error bars represent standard deviations across participants. Colored lines and markers indicate different feedback conditions. SSN masker = constant speech-shaped noise masker, DEMN = (delayed) enveloped-modulated noise, DAF = canonical (delayed) auditory feedback. The SSN masker condition was constant throughout a trial so has no matched 200 ms delay condition.

References

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4.

Boersma, P., & Weenink, D. (2018). Praat: Doing phonetics by computer [Computer program]. *Journal of Statistical Software, 67*(1), 1–48.

Corretge, R. (2021). Praat Vocal Toolkit [Computer program].

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

Yates, A. J. (1963). Delayed auditory feedback. *Psychological Bulletin, 60*(3), 213.

Howell, P., Powell, D. J., & Khan, I. (1983). Amplitude contour of the delayed signal and interference in delayed auditory feedback tasks. *Journal of Experimental Psychology: Human Perception and Performance, 9*(5), 772.

Howell, P., & Archer, A. (1984). Susceptibility to the effects of delayed auditory feedback. *Perception & Psychophysics, 36*(3), 296–302.

Müller, K., Aschersleben, G., Esser, K., & Müsseler, J. (2000). Effects of delayed auditory feedback on speech: Just a problem of displaced rhythm? In *Rhythm perception and production* (pp. 249–261). Swets & Zeitlinger.

Kaspar K. & Rübeling, H. (2011). Rhythmic Versus Phonemic Interference in Delayed Auditory Feedback. *Journal of Speech, Language, and Hearing Research, 54*(3), 932–943. https://doi.org/10.1044/1092-4388(2010/10-0109)

Garofolo, J., Lamel, L., Fisher, W., Fiscus, J., Pallett, D., & Dahlgren, N. (1993). DARPA TIMIT acoustic phonetic speech corpus. *Linguistic Data Consortium*, Philadelphia.

SPEECH ELICITATION METHODS FOR MEASURING ARTICULATORY CONTROL VARY: DOES IT MATTER?

Claudia I. Abbiati, Kim R. Bauerly, Shelley L. Velleman
University of Vermont, Burlington, United States of America

Articulatory variability is a widely used measure for examining speech-motor control in speakers with motor speech disorders, including stuttering (Jackson et al., 2016), dysarthria (Kuruvilla-Dugdale & Mefferd, 2017), and childhood apraxia of speech (Case & Grigos, 2021). The most common measure of articulatory variability is the spatiotemporal index (STI), which quantifies the variability of articulatory movement patterns across several repetitions of the same sentence (Smith et al., 1995).

Inconsistent movement patterns result in higher STI values whereas stable motoric patterns across repetitions yield lower STI values. While findings are relatively consistent (Jackson et al., 2016; Kuruvilla-Dugdale & Mefferd, 2017; Case & Grigos, 2021), there are differences with respect to how repetitions of the same sentence are elicited. Some studies elicit repetitions consecutively and may include brief pauses in between productions (Howell et al., 2009), while other studies elicit repetitions one at a time in a pseudorandomized order (Usler & Walsh, 2018). Furthermore, some studies elicit repetitions by presenting stimuli auditorily (Kuruvilla-Dugdale & Mefferd, 2017), while other studies use visual and auditory cues (Jackson et al., 2016). It is unknown whether these different elicitation methods affect STI values. Therefore, the aims of this study were to determine whether articulatory variability, as indexed using lip aperture (LA) STI, differed in neurotypical adults when repeating sentences varying in length and complexity, when (1) produced as consecutive (CR) versus isolated repetitions (IR) and (2) presented using auditory (A) versus auditory and visual (AV) elicitation methods.

Nineteen monolingual English adults (age range: 18-35 years; ten females and nine males) produced ten sentences varying in length and complexity (i.e., short-simple, short-complex, long-simple, long-complex). Sentences were purposefully loaded with sounds requiring lip movement and have been used in prior studies to examine articulatory variability (MacPherson & Smith, 2013). Productions were elicited under four different conditions: IR+A, IR+AV, CR+A, CR+AV. The condition orders were counterbalanced. Instructions and stimuli were prerecorded and presented on a laptop. Northern Digital's electromagnetic articulography Wave system (Northern Digital Inc.) was used to measure articulatory movements. Participants were fitted with a 3X3mm sensor on the midline border of their upper lip and lower lip using Glustitch PeriAcryl90 (Gluestitch, Inc.) and a head-mounted 5DOF sensor to track head motion. A labile microphone was positioned three feet in front of participants to simultaneously capture the acoustic signal. Data were imported into SMASH (Green et al., 2013) and analyzed using a custom MATLAB program.

Three-dimensional lip movements were corrected for head movements and smoothed using a 15Hz low-pass filter. All sentences were segmented starting with the release of [b] in "birds" and ending at the release of [p] in "played" for short sentences and "pond" for long sentences (e.g., short-simple: The birds and the butterflies played; long-complex: The birds that saw butterflies played by the pond). Articulatory variability was assessed for each sentence type.

Data analysis is currently underway. Preliminary results for 11 participants show a mean increase in LA-STI under IR+A versus CR+A for long-complex (M = 3.75), long-simple (M = 2.50), short-complex (M = 2.44), and short-simple (M = 2.04) sentences.

There is a trend for larger differences between elicitation conditions for longer and more complex sentences. Findings will provide important insight into the speech elicitation methods used for measuring articulatory variability.

References

Case J, Grigos MI. (2021). The effect of practice on variability in childhood apraxia of speech: A multidimensional analysis. *American Journal of Speech-Language Pathology, 30*:1477-1495.

Green JR, Wang J, Wilson DL. (2013). SMASH: A tool for articulatory data processing and analysis. In: Bimbot F, Cerisara C, Fougeron C, Gravier G, Lamel L, Pellegrino F, Perrier P eds. *Interspeech. International Speech Communication Association*:1331–1335.

Howell P, Anderson AJ, Bartrip J, Bailey E. (2009). Comparison of acoustic and kinematic approaches to measuring utterance-level speech variability. *Journal of Speech, Language, and Hearing Research, 52*(4):1088–1096.

Jackson E, Tiede M, Beal D, Whalen DH. (2016). The impact of social-cognitive stress on speech variability, determinism, and stability in adults who do and do not stutter. *Journal of Speech, Language, and Hearing Research, 59*(6):1295-1314.

Kuruvilla-Dugdale M, Mefferd A. (2017). Spatiotemporal movement variability in ALS: Speaking rate effects on tongue, lower lip, and jaw motor control. *Journal of Communication Disorders, 67*:22-34.

MacPherson MK, Smith A. (2013). Influences of sentence length and syntactic complexity on the speech motor control of children who stutter. *Journal of Speech, Language, and Hearing Research, 56*(1):89-102.

Smith A, Goffman L, Zelaznik HN, Ying G, McGillem C. (1995). Spatiotemporal stability and patterning of speech movement sequences. *Experimental Brain Research, 104*(3):493-501.

Usler ER, Walsh B. (2018). The effects of syntactic complexity and sentence length on the speech motor control of school-age children who stutter. *Journal of Speech, Language, and Hearing Research, 61*(9):2157-2167.

MOTOR LEARNING AND DEVELOPMENTAL STUTTERING: A SYSTEMATIC REVIEW

Fiona Höbler[1], Silvia Isabella[2], Ala Refai[1], Tali Bitan[1,3], Luc Tremblay[1], Luc de Nil[1]
[1] *University of Toronto, Toronto Canada*
[2] *Hospital for Sick Children Research Institute, Toronto, Canada*
[3] *University of Haifa, Haifa, Israel*

Fluent speech production relies on the smooth and efficient coordination, integration and automaticity of several neural and sensorimotor processing. These processes have been found to be disrupted in developmental stuttering. Studies have pointed to impairments in areas of motor planning and coordination not only on speech tasks, but also on non-speech motor planning and coordination. Motor learning paradigms of sequential and adaptive movement have demonstrated differences between individuals who do and do not stutter in skill acquisition, retention and transfer. However, the accumulative findings of the motor learning literature in developmental stuttering have not yet been evaluated. This systematic review aimed at identifying and evaluating current evidence of motor skill learning in individuals who stutter at distinct stages of the lifespan, and to interpret the research findings as they relate to theories of motor learning, speech motor development, and the treatment of stuttering.

A systematic search of the literature was carried out on the 8th December 2020 across six electronic databases, including: Embase, MEDLINE, PsycINFO, CINAHL, ASSIA, and AMED. 4,362 unique records were screened by independent reviewers based on a-priori defined inclusion criteria. Following consensus between two reviewers, 60 eligible articles were included in the final review. Reported research data were categorized as relevant to motor performance, where measurements were obtained across trials within one practice session, or motor learning outcomes, where performance measures were based on delayed tests of retention, consolidation and/or follow-up. Parameters of motor learning were extracted from the study data and summarised in the results of this review.

The majority of research reports involved adult participants who do and do not stutter (n=50). Studies also frequently involved male participants only. Ten studies reported on motor skill learning in children who do and do not stutter, with two studies including both adults and children. Among these, a smaller number of studies investigated motor learning outcomes among adults (n=12) and children who stutter (n=2), where performance gains were measured beyond one practice session. Significant differences have been found between adults who do and do not stutter on both speech and non-speech tasks. The evidence among children who stutter is still limited and less clear.

Here we report on the parameters of motor learning, including the structure of practice trials and type of feedback provided, that have been found to differentiate individuals who do and do not stutter at distinct stages of development. We also highlight the need for extended follow-up to investigate the long-term learning effects and maintenance abilities of persons who stutter, along with further research into the influences of extended practice and transfer. Further research that includes larger and more sex-balanced participant samples is warranted. Motor learning research can shed light on the important considerations for the structure and delivery of fluency intervention methods.

THE EFFECT OF ERROR-CLAMPING AUDITORY FEEDBACK ON CENTERING

Kwang S. Kim, Srikantan Nagarajan, John Houde
University of California, San Francisco, California, United States of America

In the last decade, several studies have reported on a speech phenomenon called centering, wherein the formants (F1 and F2) of a speaker's vowel production, which deviate from the median at utterance onset, move closer the median as production of the vowel progresses (Niziolek et al., 2013; Tomassi et al., 2022). This within-utterance formant movement of centering was thought to be at least partly driven by auditory feedback, since centering was correlated with speaking-induced suppression, a measure of auditory feedback processing during speaking (Niziolek et al., 2013). Subsequent studies that investigated the role of auditory feedback in centering by adding masking noise have however reported conflicting results. In one study, centering was reduced under masking noise conditions (Niziolek et al., 2015), but the same effect was not found in two other studies (Niziolek & Bakst, 2019; Parrell et al., 2021).

Conflicting results on centering when masking noise was added in auditory feedback during speaking could have occurred because masking noise not only decreased the quality of auditory feedback, but also affected the speech motor system's general reliance on feedback-based mechanisms. When the auditory feedback becomes noisy and unreliable, the speech motor control system may change its belief (how much it weighs sensory prediction errors in making state corrections, see Parrell et al., 2019 for more details) in multiple sensory domains. In turn, the system may decide to rely more on somatosensory feedback, for example, which could explain the enhancement of centering observed under some masking noise conditions (Niziolek & Bakst, 2019).

In the current study, we attempted to reduce the use of auditory information in centering without the potential risks of affecting the general reliability of sensory feedback-based mechanisms. Specifically, we error-clamped auditory feedback to the "ideal" median production so that the error-clamped auditory feedback informed the system that the ongoing production was ideal, and no correction was necessary. Hence, we hypothesized that centering would be reduced in error-clamped trials if the mechanism relied on auditory feedback. First, we first asked participants to speak under un-perturbed auditory feedback through over-ear headphones ("baseline"). With this data, we calculated the median timecourse of F1 and F2 across their productions and saved the resulting median production for the subsequent error-clamp session. In the next session, participants produced the words with their median production replacing their normal auditory feedback ("error-clamped").

On average, there was less centering in the error-clamped trials compared to the baseline trials. In addition, although the initial vowel variability was not different between the two conditions, the mid vowel variability was significantly larger in the error-clamped trials. Importantly, the centering behavior, while reduced, was still present in the error-clamped trials, suggesting that there may be other mechanisms involved in centering (e.g., somatosensory feedback). Combined together, our findings demonstrate that centering relies on auditory feedback, potentially in addition to other mechanisms.

References

Niziolek, C. A., & Bakst, S. (2019). Self-correction in L1 and L2 vowel production. In S. Calhoun, P. Escudero, M. Tabain, & P. Warren (Eds.), *Proceedings of the 19th International Congress of Phonetic Sciences* (pp. 3185–3189). *Australasian Speech Science and Technology Association Inc.*

Niziolek, C. A., Nagarajan, S. S., & Houde, J. F. (2013). What does motor efference copy represent? Evidence from speech production. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience, 33*(41), 16110–16116. https://doi.org/10.1523/JNEUROSCI.2137-13.2013

Niziolek, C. A., Nagarajan, S. S., & Houde, J. F. (2015). The contribution of auditory feedback to corrective movements in vowel formant trajectories. In The Scottish Consortium for ICPhS 2015 (Ed.). *Proceedings of the 18th International Congress of Phonetic Sciences*, 4.

Parrell, B., Ivry, R. B., Nagarajan, S. S., & Houde, J. F. (2021). Intact Correction for Self-Produced Vowel Formant Variability in Individuals With Cerebellar Ataxia Regardless of Auditory Feedback Availability. *Journal of Speech, Language, and Hearing Research: JSLHR, 64*(6S), 2234–2247. https://doi.org/10.1044/2021_JSLHR-20-00270

Parrell, B., Ramanarayanan, V., Nagarajan, S., & Houde, J. (2019). The FACTS model of speech motor control: Fusing state estimation and task-based control. *PLoS Computational Biology, 15*(9), e1007321. https://doi.org/10.1371/journal.pcbi.1007321

Tomassi, N. E., Weerathunge, H. R., Cushman, M. R., Bohland, J. W., & Stepp, C. E. (2022). Assessing Ecologically Valid Methods of Auditory Feedback Measurement in Individuals With Typical Speech. *Journal of Speech, Language, and Hearing Research: JSLHR, 65*(1), 121–135. https://doi.org/10.1044/2021_JSLHR-21-00377

THE EFFECT OF DRAMA CLASSES ON SPEECH PRODUCTION IN CHILDREN WITH DYSARTHRIA: A SURVEY OF PARENTAL PERCEPTIONS

Anja Kuschmann, Inoka Mirihagalla Kankanamalage
University of Strathclyde, Glasgow, United Kingdom

Background Participating in performing arts activities that involve vocal activities such as group singing and drama classes can have psychosocial benefits including increased self-confidence (Barnish & Barran, 2020). In recent years, there has been increasing interest in the therapeutic benefits of such activities for adults with motor speech disorders. Group singing, for instance, improved health-related quality of life, mood, breathing and voice quality in adults with dysarthria due to Parkinson's disease (e.g. Fogg-Rogers et al., 2016). Psychosocial benefits were also reported in children and young people with learning disabilities (LD), who participated in performing arts activities (e.g. Wu et al., 2020; Zyga et al., 2018). Losardo et al. (2019) offered drama classes to young people with LD, and reported measurable improvements in intelligibility, fluency and the ability to initiate social conversations. These changes were confirmed by parental reports.

Aim The current study aimed to determine whether children with dysarthria due to physical disabilities such as cerebral palsy might also benefit from drama classes, and if so whether the benefits are psychosocial in nature or extend to improvements in speech production.

Methods Parents were asked to complete an e-survey to report any changes they may have noticed in their children's speech, communicative participation and social skills as a result of participating in drama classes offered by a small charity for children with dysarthria. The e-survey was designed and distributed via Qualtrics. Binary choice questions, Likert scale questions and open questions were used to collect information. Results were analysed using descriptive statistics.

Results Six parents responded to the survey (50% response rate) and all responses were included in the analysis. Responses were provided for four girls and two boys, aged between 6-18 years, who attended sessions for 1-24 months. Children's motor speech difficulties were reported as moderate (four children) or severe (two children).

Parents reported improvements in all areas investigated i.e. speech, communicative participation and social skills. In terms of speech, 5/6 parents agreed that their children produced more words in one breath; 4/6 parents also reported that their children's voice was louder as well as more stable. In addition, 4/6 parents reported that family, friends and unfamiliar people found it easier to understand their child. Fewer parents noticed changes to nasality (2/6). Regarding communicative participation, parents reported remarkable improvements for seeking clarifications (4/6), listening to others (3/6) and willingness to share ideas (3/6). Moderate improvements were noticed in the children's ability to maintain a conversation (4/6) and to express themselves (3/6). Five parents further reported that their children developed more self-confidence.

Conclusion Improvements in all areas surveyed suggest that the children with dysarthria benefitted from participating in drama classes not just in terms of psychosocial development but also in terms of positive changes to speech production. Further studies are

warranted to quantify the observed changes to speech using objective measures.

References

Barnish, M.S. & Barran, S.M. (2020). A systematic review of active group-based dance, singing, music therapy and theatrical interventions for quality of life, functional communication, speech, motor function and cognitive status in people with Parkinson's disease. *BMC Neurology, 20*, 1-15.

Fogg-Rogers, L., Buetow, S., Talmage, A., McCann, C. M., Leão, S. H. S., Tippett, L., Leung, J.... Purdy, S. C. (2016). Choral singing therapy following stroke or Parkinson's disease: an exploration of participants' experiences, *Disability and Rehabilitation, 38*(10), 952-962.

Losardo, A., Davidson, D. & McCullough, K. (2019). Stages of Success: The Theatre and Therapy Project: Speech-language pathology and theatre education students work together in a program for adolescents and young adults with developmental disabilities. *The ASHA Leader*, March 2019.

Wu, J., Chen, K., Ma, Y. & Vomočilová, J. (2020). Early intervention for children with intellectual and developmental disability using drama therapy techniques. *Children and Youth Services Review, 109*, 1-7.

Zyga, O., Russ, S. W., Meeker, H. & Kirk J. (2018). A preliminary investigation of a school-based musical theater intervention program for children with intellectual disabilities. *Journal of Intellectual Disability, 22*(3), 262-278.

FLUENCY FRIENDS: A SPEECH RECOGNITION-BASED VIDEO GAME FOR PEOPLE WHO STUTTER

Lukas Latacz[1], Susan Fosnot[2], Erich Reiter[1]
[1]*SAY IT Labs BV, Leuven, Belgium*
[2]*Chapman University, Orange, California, United States*

We are investigating how a speech-recognition-based video game can be used to motivate fluency practice in conjunction with and beyond the clinic. This hybrid approach of combining modern technology with traditional speech therapy for people who stutter (PWS) is a novel approach of practicing speech independently while receiving immediate feedback.

The video game – Fluency Friends (SAY IT Labs, 2022) – has been in development for several years and is based on the Fluency Development System for Young Children (Meyers and Woodford, 1992) which teaches motoric and cognitively-based fluency exercises specifically for young children. This therapy program uses slow speech to talk efficiently without compromising the prosodic elements of natural fluent speech (rhythm, stress patterns, loudness, pitch, and duration of speech). It also contains elements of social discourse such as turn-taking and talking under pressure. The video game additionally incorporates mindfulness and breathing techniques to offer players learning opportunities on how to control anxiety and stress (Hardy, 2015), (Kuypers, 2011). Early prototypes of the game were tested at a 1-week specialized camp for children who stutter (CWS) and again in the home and clinic settings. Each prototype contained several speech-controlled mini-games that focused on 1 particular aspect of speech, such as rate of speech, pitch, or linguistic stress.

The current version of Fluency Friends has evolved into a narrative video-game that provides intrinsic motivation through gaming elements such as a relatable story for PWS – a quest to get their voice back –which includes many challenges and battles, collecting rewards and getting upgrades for voice use. There are many elements that are relevant to speech and language pathology including in-game scaffolding, immediate feedback to players as they play the game, and automatic data-collection to track the progress of the player.

At the conference, we will report on our latest research using the video game. We are currently investigating how the latest version of the Fluency Friends can motivate independent practice and promote speech motor skills. Subjects are PWS between 6 to 17 years of age who qualify as having moderate to severe stuttering in accordance with the SSI-4 (Riley, 2009), without any other speech or language co-morbidities. We are measuring effects of playing the game regularly for several months, using user-reported outcome measures and acoustic and linguistic measurements that are relevant to the exercises in the video game.

References

Hardy, S. T. (2015). *Asanas for autism and special needs: Yoga to help children with their emotions, self-regulation, and body awareness.* Jessica Kingsley Publishers.

Kuypers, L. (2011). The Zones of Regulation. The Zones of Regulation. https://www.zonesofregulation.com/

Meyers, S. & Woodford, L. (1992). The fluency development system for young children. United Education Services.

Riley, G. (2009). *The Stuttering Severity Instrument for Adults and Children* (SSI-4) (4th ed.). Austin, TX: PRO-ED.

SAY IT Labs (2022), Fluency Friends, Available at: https://www.sayitlabs.com/fluency-friends

CLEARSPEECHTOGETHER – AN SLT/PEER SUPPORTED SPEECH INTERVENTION MODEL FOR PEOPLE WITH PROGRESSIVE ATAXIAS

Anja Lowit[1], Jessica Cox[1], Aisling Egan[1], Marios Hadjivassiliou[2]
[1]*Strathclyde University, Glasgow, United Kingdom*
[2]*Sheffield Teaching Hospital, Sheffield, United Kingdom*

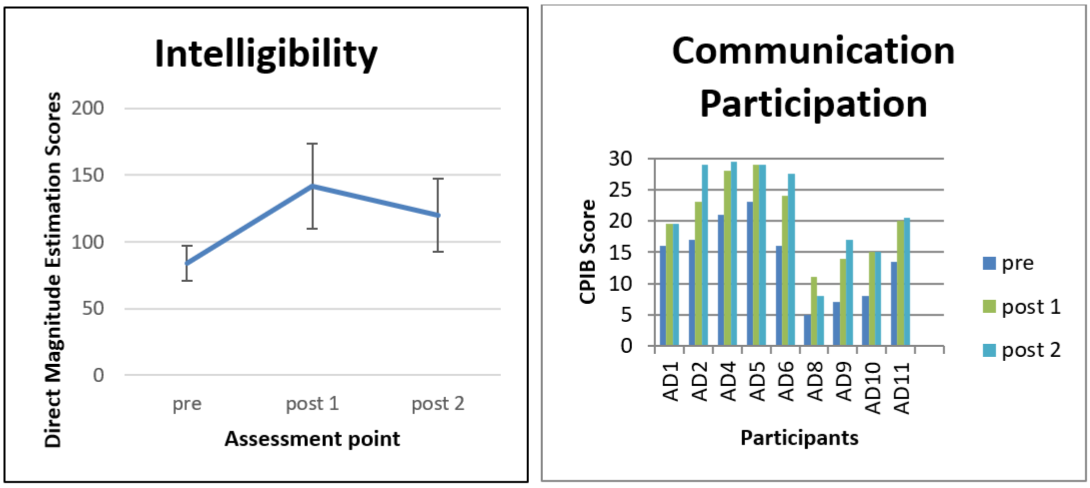
Background Progressive ataxias such as Friedreich's Ataxia or Spino-Cerebellar Ataxia frequently result in speech difficulties. Ataxic dysarthria affects all articulatory systems, i.e. breath support, voice quality and articulation. It is characterised by uncontrolled movements, resulting in excess pitch and loudness excursions and rhythm disturbances (Duffy, 2019). As with all dysarthrias, these problems can impact on communication participation and psycho-social wellbeing. Whilst there is a lack of a reliable evidence base for effective speech intervention (Vogel et al. 2014), a number of small scale studies have recently demonstrated the potential for improvements in physiological function, intelligibility as well as dysarthria impact across a wider range of ataxia types. However, with one exception, these treatments all required intensive input from clinicians over an extended period of time, which can reduce patient access to such treatment. One study employed a home practice app, which reduces clinician time, but at the same time provides less opportunities to practise speaking in real life contexts and psycho-social support (Vogel et al., 2019). We propose a new model of care – ClearSpeechTogether – which combines individual SLT and peer led group therapy to maximise patient practice whilst minimising pressure on clinicians' workloads.

This study piloted ClearSpeechTogether to establish feasibility and accessibility of the approach as well as potential benefits and adverse effects on patients with progressive ataxia.

Method We aimed to recruit 10 participants to run two groups. All had progressive ataxia of various causes, and mild-moderate speech and gross motor impairment. Intervention consisted of 4 individual sessions over 2 weeks, followed by 20 patient led, clinician supported group sessions over 4 weeks. Treatment focused on "Loud and Clear". Communication outcome measures included maximum phonation time (MPT), voice quality, reading and monologue intelligibility, and communication participation and confidence.

Results Recruitment was effective with a waiting list to join the study. Attrition was at 20% with one person dropping out after assessment and another after completing the individual therapy phase. The former could be replaced and 11 participants were thus recruited and 9 completed treatment. Acceptability was high and no adverse effects were reported. Statistical tests indicated no change in MPT, but significantly reduced strain in perceptual voice quality, improved intelligibility in reading (but not the monologue) and increased participation and confidence (Figure 1). Participant interviews highlighted the added value of the group sessions, both from a psycho-social perspective as well as to support internalisation and carry over of speech strategies.

Discussion ClearSpeechTogether presented an effective intervention approach for a small sample of people with progressive ataxia. It matched, if not exceeded the communication outcomes previously reported for intensive, individual therapy and did so at a considerably lower cost in terms of clinician time. In addition, it provided added value in tackling psycho-social issues that can be difficult to address in one-to-one therapy. ClearSpeechTogether focuses on generic speech strategies involving effective voice production and clear articulation, and should thus be applicable to a wide range of patients with acquired motor speech disorders. We hope to validate this in a future randomised controlled trial.



Reading Intelligibility and Communication Participation Item Bank (CPIB) results for pre versus immediate and long-term post-treatment assessments

References

Duffy, J. R. (2019). *Motor Speech Disorders - Substrates, Differential Diagnosis, and Management* (4th ed.). Elsevier, Mosby.

Vogel, A. P., Folker, J., & Poole, M. L. (2014). Treatment for speech disorder in Friedreich ataxia and other hereditary ataxia syndromes. *Cochrane Database of Systematic Reviews(10)*. `https://doi.org/10.1002/14651858.CD008953.pub2`

Vogel, A. P., Stoll, L. H., Oettinger, A., Rommel, N., Kraus, E. M., Timmann, D., Scott, D., Atay, C., Storey, E., Schöls, L., & Synofzik, M. (2019). Speech treatment improves dysarthria in multisystemic ataxia: a rater-blinded, controlled pilot-study in ARSACS. *J Neurol, 266*(5), 1260-1266. `https://doi.org/10.1007/s00415-019-09258-4`

A RANDOMIZED CONTROLLED TRIAL OF ASSIST FOR CHILDHOOD APRAXIA OF SPEECH: INITIAL FINDINGS

Edwin Maas[1], Susan Caspari[1], Molly Beiting[1], Christina Gildersleeve-Neumann[2], Ruth Stoeckel[3], Jingwei Wu[1]
[1] *Temple University, Philadelphia, United States of America*
[2] *Portland State University, Portland, Oregon, United States of America*
[3] *Mayo Clinic, Rochester, Minnesota, United States of America (retired)*

Introduction Integral stimulation-based treatments for CAS have replicated support from small-sample studies using single-case experimental designs (Murray et al., 2014). The primary purpose of this study was to examine the initial efficacy of ASSIST (Apraxia of Speech Systematic Integral Stimulation Treatment; Maas et al., 2019) for CAS in a randomized controlled trial design. In addition, a question that has arisen clinically is whether targeting simple or complex items leads to greater gains. Therefore, a secondary purpose was to examine the effect of utterance complexity. The study involved two recruitment cycles (summer 2019, summer 2022); here we report data from summer 2019.

Methods This study used a parallel-groups randomized controlled group design with a delayed treatment control group. The study took place in the context of an intensive summer camp (4 days/week, 4 weeks). Seventeen children with CAS (ages 4-9) were randomly assigned to either immediate (weeks 2-3; n=9) or delayed treatment (weeks 5-6; n=8), and to simple (n=8) or complex conditions (n=9). Outcome data were collected in weeks 1, 4, and 7 (T1, T2, and T3). Each child received 16 hours of individual ASSIST (2 hours/day). ASSIST includes pre-practice and systematic criteria for adaptive practice according to principles of motor learning (Maas et al., 2008).

For each child, we developed 30 personally meaningful words/phrase targets; only 4 were practiced in any session; new items were substituted in as targets were mastered. Targets were embedded in carrier frames, which differed in phonetic complexity (Jakielski, 2017) to create simple and complex utterances. When not in treatment, children engaged in camp activities not focused on speech/language. Data were collected by blinded data collectors and analyzed by blinded analysts. Change scores based on target accuracy were submitted to independent samples t-tests (α=0.05). To address ASSIST efficacy, we examine changes from T1 to T2. To address complexity effects, we examine changes from before to after treatment (T1-T2 for immediate group; T2-T3 for delayed group).

Results Initial data based on all 30 items from 15 children revealed numerically greater change in the ASSIST group (16%) than in the Control group (5%), but this difference was not significant (t_{13}=1.10, p=0.293). When examining only items that were treated, the difference (41% vs. -4%) showed a non-significant trend (t_{13}=1.78, p=0.099). For the complexity analysis involving all 30 items, numerically greater gains were observed for the Complex (11%) than for the Simple condition (-7%), but this difference was not significant (t_{13}=1.29, p=0.219).

The same non-significant pattern was observed when examining only treated items (19% vs. -7%; t_{13}=0.74, p=0.475).

Discussion Although no significant differences were observed in this initial, incomplete sample, numerical patterns are consistent with an advantage for ASSIST over no treatment, and for targeting complex over simple utterances. If these patterns hold with the full sample, this would support the initial efficacy of ASSIST and suggest that practicing more complex utterances may enhance learning.

References

Maas, E., Gildersleeve-Neumann, C., Jakielski, K., Kovacs, N., Stoeckel, R., Vradelis, H., & Welsh, M. (2019). Bang for your buck: A single-case experimental design study of practice amount and distribution in treatment for childhood apraxia of speech. *Journal of Speech, Language, and Hearing Research, 62*, 3160-3182.

Murray, E., McCabe, P., & Ballard, K. J. (2014). A systematic review of treatment outcomes for children with Childhood Apraxia of Speech. *American Journal of Speech-Language Pathology, 23*, 486-504.

THE RELATIONSHIP BETWEEN NEUROGENIC STUTTERING AND HYPOKINETIC DYSARTHRIA IN PARKINSON'S DISEASE

Fatemeh Mollaei, Sophie Atkins, Megan Lewis
University of Reading, Early Gate, Whiteknights, Reading, United Kingdom

Introduction Parkinson's disease (PD) is a multifaceted disorder with motor and non-motor symptoms. Speech impairments are one of the common symptoms with 90% of individuals with PD showing speech and auditory processing deficits including prosody, phonation, and articulation (Duffy, 2019). These deficits can be broken down into two main categories: hypokinetic dysarthria and neurogenic stuttering (Goberman et al., 2010). While there has been much focus on the hypokinetic dysarthria characteristics of PD, there is little investigation into the neurogenic stuttering and its relation to hypokinetic dysarthria in PD. Here we investigated the relationship between these two main speech impairments in PD to better understand and be able to design a holistic assessment and management plans for these deficits in PD.

Methods We included 32 individuals with PD in the analysis. All participants were tested off-medication for 12 hours. Disfluencies were differentiated into within-word disfluencies (WW: sound and syllable repetitions, blocks, and sound prolongations) and between-whole-word disfluencies (BW: phrase repetitions, revisions, and interjections) according to the guidelines provided by Yaruss (1998). We used a cut off of higher than 2% for WW disfluencies as indicative of stuttering, which is somewhat lower than 3% threshold suggested by Conture (2001). However, their threshold was based on children population, and they also mentioned that there can be variations. For the characterisation of the subtype of dysarthria and its severity, specifically hypokinetic dysarthria, we rated participants speech samples (rainbow passage; Fairbanks, 1960) on 43 perceptual characteristics related to the phonatory (e.g., loudness and pitch), articulatory, resonatory, prosodic, and respiratory dimensions of speech.

Results Overall disfluencies (combined BW and WW) were significantly correlated with dysarthria severity, articulation severity, and prosody severity. Significant correlation between WW disfluencies and articulation severity was found. PD participants with neurogenic stuttering were found to have significantly more severe dysarthria than PD participants without neurogenic stuttering.

Conclusion Our findings associate neurogenic stuttering and hypokinetic dysarthria in PD. While the present study is one of the early studies to investigate the relationship of dysarthria and neurogenic stuttering and more work is needed to replicate and confirm this relationship, it has implications for planning holistic assessment and treatment approaches for PD with regards to dysarthria and especially the articulation subtype.

References

Conture, E. G. (2001). Stuttering: Its nature, diagnosis, and treatment. Pearson College Division.

Duffy, J. R. (2019). *Motor speech disorders e-book: Substrates, differential diagnosis, and management.* Elsevier Health Sciences.

Fairbanks, G. (1960). *The rainbow passage. Voice and articulation drillbook*, 2, 127-127.

Goberman, A. M., Blomgren, M., & Metzger, E. (2010). Characteristics of speech disfluency in Parkinson disease. *Journal of Neurolinguistics, 23*(5), 470-478.

Yaruss, J. S. (1998). Describing the consequences of disorders: Stuttering and the international classification of impairments, disabilities, and handicaps. *Journal of Speech, Language, and Hearing Research, 41*(2), 249-257.

AN INVESTIGATION INTO ACOUSTIC SPEECH MARKERS IN HYPOKINETIC DYSARTHRIA ASSOCIATED WITH PARKINSON'S DISEASE

Mridhula Murali, Joan Ma, Robin Lickley
Queen Margaret University, Edinburgh, United Kingdom

Previous research has identified certain overarching features of hypokinetic dysarthria (HD) associated with Parkinson's Disease (PD) and found that HD often manifests differently between individuals. This study attempted to find trackable acoustic speech markers that might be more sensitive to speech changes over time and could highlight HD changes as PD progresses. These markers could also act as supplements to current perceptual methods of evaluation as acoustic analysis is quantifiable and relies of standard methods of analysis. This study focused on acoustic analysis of speech associated with hypokinetic dysarthria to identify acoustic speech markers linked to changes in speech functions over time in people with Parkinson's Disease (PwPD). The acoustic markers would have to be consistently present and detectable, six months apart.

Speech data was collected from 63 PwPD and 47 control speakers using an online podcast software at two time points six months apart). Recordings of a standard reading passage, minimal pairs, sustained phonation, and spontaneous speech were collected. Acoustic parameters pertaining to voice, articulation and prosody were investigated and perceptual severity ratings given by two speech and language therapists for both data collection points. Speech recordings of PwPD were then grouped into mild, moderate, and severe based on the severity ratings and sub-grouped based on whether their overall speech intelligibility has changed or not six months apart (for example, mild, change group & mild, no change group). This was done to check whether the acoustic parameters were able to a) confirm the results of the perceptual ratings and b) if acoustic parameters could pick up on a change where it was not auditorily perceptible. Speech was also compared to a control group at each data collection point to evaluate if the acoustic features were distinct enough from control speech over time.

An attempt was made to try to identify these markers both within group and within speaker to establish whether overall group markers exist. If individual differences were found to be more prominent than group differences, it could necessitate the creation of speaker-specific 'voice profiles' in the future that could be used to track speech changes more efficiently. Results from this study could help identify acoustic features that are more susceptible to detecting changes in the speech of PwPD, which could be used against voice samples of an individual and tracked over time.

The potential use of online data collection tool in this population in the current study identified a possible avenue for remote data collection and speech monitoring applications in the future. However, the possible pitfalls of this method of data collection with PwPD, including data collection design and loss of data backups will be discussed. Trackable acoustic speech markers may benefit in creating a clearer link between changes in speech function by better isolating how acoustic correlates can inform speech motor symptomatology. In addition, the acoustic parameters identified in this study could be used to investigate whether they are equally sensitive to changes in speech that may occur because of dopaminergic medication, or speech and language therapy.

References

Harel, B. T., Cannizzaro, M. S., Cohen, H., Reilly, N., & Snyder, P. J. (2004). Acoustic characteristics of Parkinsonian speech: A potential biomarker of early disease progression and treatment. *Journal of Neurolinguistics, 17*(6), 439–453.

Murali, M. (2022). Using a podcast application to collect high-quality speech data online for acoustic analysis in people with Parkinson's disease. In *SAGE Research Methods Cases.* `https://dx.doi.org/10.4135/9781529600575`

Skodda, S., Rinsche, H., & Schlegel, U. (2009). Progression of dysprosody in Parkinson's disease over time-A longitudinal study. *Movement Disorders, 24*(5), 716.

PERCEPTUAL CLASSIFICATION OF SPEAKERS WITH DYSARTHRIA OR APRAXIA OF SPEECH: EFFECTS OF SPEAKER, SPEECH TASK, AND LISTENER'S EXPERTISE

Michaela Pernon[1,2,3], Frédéric Assal[1], Ina Kodrasi[4], Marina Laganaro[1]
[1] *University of Geneva, Geneva, Switzerland*
[2] *CNRS-Université Sorbonne Nouvelle, Paris, France*
[3] *Hôpital Fondation A. de Rothschild, Paris, France*
[4] *Idiap Research Institute, Martigny, Switzerland*

Purpose In clinical practice, the diagnosis of motor speech disorders (MSDs), apraxia of speech (AoS) and dysarthria, is mainly based on perceptual approaches (Kent et al., 1996). However, the accuracy of perceptual classification may be quite low (Bunton et al., 2007; Mumby et al., 2007; Zyski & Weisiger, 1987) all the more that the different types of MSD share several perceptive signs. The aim of the present study was to examine (i) the accuracy of speech-language pathologists (SLPs) listeners in perceptually classifying apraxia of speech (AoS) and dysarthria, (ii) the impact of speech task, severity of MSD and listener's expertise on classification and (ii) the perceptual features used by SLPs.

Method Speech samples from 29 neurotypical speakers, 14 with hypokinetic dysarthria associated with Parkinson's disease (HD), 10 with post-stroke AoS and 6 with mixed dysarthria (flaccid and spastic) associated with amyotrophic lateral sclerosis (MD-FlSp), were classified by 40 listeners, 20 Expert SLPs and 20 Student SLPs. In an auditory-perceptual forced-choice classification task, listeners rated recorded speech samples elicited in spontaneous speech, text reading, oral diadochokinetic tasks (DDK) and a sample concatenating text reading and DDK, following a diagnostic approach. For each speech sample, SLPs answered three dichotomic questions: (1) neurotypical versus pathological speaker, (2) AoS versus dysarthria, (3) MD-FlSp versus HD, and a multiple-choice question indicating on which features their decision was based on.

Results Overall classification accuracy was 72%, with good inter-rater reliability. Correct classification of speech samples was higher for dysarthria than for AoS, and listeners classified HD more accurately than MD-FlSp. HD speakers were confused with neurotypical speakers and AoS speakers with MD-FlSp speakers. Effects of listener's expertise, MSD severity and speech task on classification accuracy were also found. Samples elicited with spontaneous speech and text reached the best classification rates. An average number of three perceptual features was used for correct classifications, and their type and combination differed between the three MSDs.

Conclusions The perceptual classification of MSDs in a diagnostic approach reaches substantial performance only in Expert SLPs and with continuous speech samples, albeit with lower accuracy for AoS. Overall although far from beeing perfect, classification accuracy was however higher than in previous studies (Bunton et al., 2007; Fonville et al., 2008; Van der Graaff et al., 2009; Zyski & Weisiger, 1987) probably due to the 3-step procedure that was close to the clinical decision approach.

References

Bunton, K., Kent, R. D., Duffy, J. R., Rosenbek, J. C., & Kent, J. F. (2007). Listener agreement for auditory-perceptual ratings of dysarthria. *Journal of speech, language, and hearing research*. https://doi.org/10.1044/1092-4388(2007/102)

Fonville, S., Van Der Worp, H. B., Maat, P., Aldenhoven, M., Algra, A., & Van Gijn, J. (2008). Accuracy and inter-observer variation in the classification of dysarthria from speech recordings. *Journal of Neurology, 255(10)*, 1545-1548. https://doi.org/10.1007/s00415-008-0978-4

Kent, R. D. (1996). Hearing and believing: Some limits to the auditory-perceptual assessment of speech and voice disorders. *American Journal of Speech-Language Pathology*, 5(3), 7-23. https://doi.org/10.1044/1058-0360.0503.07

Mumby, K., Bowen, A., & Hesketh, A. (2007). Apraxia of speech: how reliable are speech and language therapists' diagnoses? *Clinical Rehabilitation, 21*(8), 760-767. https://doi.org/10.1177/0269215507077285

Van der Graaff, M., Kuiper, T., Zwinderman, A., Van de Warrenburg, B., Poels, P., Offeringa, A., Van der Kooi, A., Speelman, H., & De Visser, M. (2009). Clinical identification of dysarthria types among neurologists, residents in neurology and speech therapists. *European neurology, 61*(5), 295-300. https://doi.org/10.1159/000206855

Zyski, B. J., & Weisiger, B. E. (1987). Identification of dysarthria types based on perceptual analysis. *Journal of Communication Disorders, 20*(5), 367-378. https://doi.org/10.1016/0021-9924(87)90025-6

EFFECTS OF AGE, DISEASE, AND L-DOPA ON AIRFLOW IN PARKINSONIAN DYSARTHRIA

Clara Ponchard[1], Alain Ghio[2], François Viallet [2,3], Lise Crevier Buchman[1,4], Didier Demolin [1]

[1] *Sorbonne Nouvelle, Paris, France*

[2]*Aix-Marseille Université, CNRS LPL, Aix-en-Provence, France*

[3]*Centre Hospitalier du Pays d'Aix, Aix-en-Provence*

[4]*Hôpital Foch, Université Paris-Saclay, Paris, France*

Introduction Parkinson's disease (PD) represents a specific disorder of the central nervous system, characterized by chronic dysfunction of the basal ganglia, which plays an essential role in controlling the execution of learned motor plans. If the motor expression of the symptoms involves mainly the limbs, the muscles involved in speech production are also subject to specific dysfunctions grouped under the term dysarthria (Pinto & Ghio, 2008). The main objective of our study is to compare the aerodynamic performances between healthy and pathological subjects in order to identify if the aerodynamic parameters allow to characterize the parkinsonian speech and what are the parameters of variations (disease, age, sex, L-DOPA treatment).

Methods Our corpus consists of 40 speakers (20 parkinsonian patients + 20 control subjects) recorded in the Neurology Department of the Centre Hospitalier du Pays d'Aix in Aix-en-Provence (Ghio et al., 2012). The speakers were native French-speaking women aged 40 to 90 years. There was no significant difference in age between the two groups. We analyzed subglottic pressure measurements estimated from the peak of intraoral pressure during the production of the consonant [p] as well as the pressure decay during the sentence "papa ne m'a pas parlé de beau papa". For Parkinsonian subjects, data were analyzed in the two pharmacological states: with L-DOPA (ON-DOPA) and with withdrawal (OFF-DOPA).

Results & Discussion OFF-DOPA patients have a significantly lower subglottic pressure than controls showing that PD has an impact on the sound pressure level which is significantly lower. In ON-DOPA, the values are similar to those of the controls but only in subjects under 60 years old, after which the pressure is lower. Thus, L-DOPA treatment increases the sound pressure level in subjects under 60 years old but would not improve it consistently. The drop in pressure is greater in the OFF-DOPA condition, particularly at the end of production, indicating a lesser control of pneumo-phonic coordination.

Older healthy subjects produce higher pressure levels than younger healthy subjects with a mean difference of 2 hPa. This increase may be due to two factors: (1) a decrease in lung elasticity that may result in an increase in the value of the pressure measured in the lung volume (Bouhuys, 1977), (2) the need for higher subglottic pressures due to changes in laryngeal airway resistance and vocal fold closure. This increase is not present in ON-DOPA subjects. In contrast, in OFF-DOPA there is an average increase of 0.8 hPa which suggests that people with PD are still able to compensate in vocal efficiency but at much lower levels than control subjects.

References

Pinto, S. & Ghio, A. (2008) Troubles du contrôle moteur de la parole : contribution de l'étude des dysarthries et dysphonies à la compréhension de la parole normale. *Revue française de linguistique appliquée, 13*(2), pp. 45-57.

Bouhuys, A. (1977). *The Physiology of Breathing: A Textbook for Medical Students*. Grune & Stratton, New-York.

Ghio, A., Pouchoulin, G., Teston, B., Pinto, S., Fredouille, C. et al. (2012). How to manage sound, physiological and clinical data of 2500 dysphonic and dysarthric speakers? *Speech Communication, 54*(5), pp.664-679. https://doi.org/10.1016/j.specom.2011.04.002

Pinto, S. & Ghio, A. (2008) Troubles du contrôle moteur de la parole : contribution de l'étude des dysarthries et dysphonies à la compréhension de la parole normale. *Revue française de linguistique appliquée, 13*(2), pp. 45-57.

A COMPARISON OF SOUND PRODUCTION TREATMENT AND METRICAL PACING THERAPY FOR APRAXIA OF SPEECH

Charlotte Purcell[1], Julie Wambaugh[2], Edwin Maas[3]
[1]*Rocky Mountain University of Health Professions, Provo, Utah, United States of America*
[2]*Temple University, Philadelphia, United States of America*
[3]*University of Utah, Salt Lake City, Utah, United States of America*

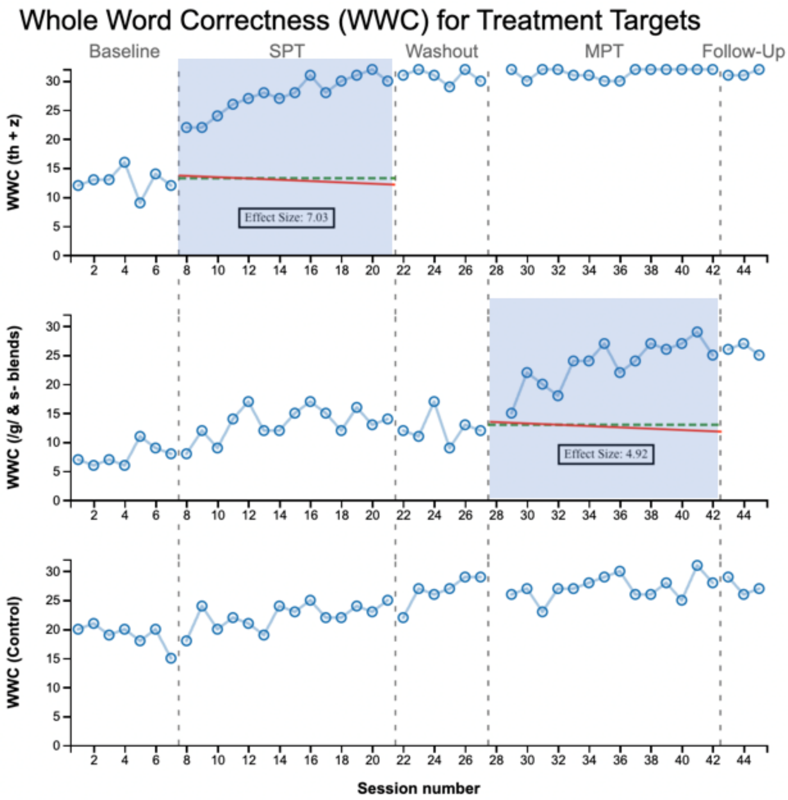
Introduction The purpose of this investigation is to compare two different specific treatment protocols for acquired apraxia of speech (AOS) that vary in theoretical construction: Sound Production Treatment (SPT; Bailey, Eatchel, & Wambaugh, 2015; Wambaugh et al., 1998) and Metrical Pacing Therapy (MPT; Brendel & Ziegler, 2008).

Aims This treatment comparison was designed to replicate each treatment's effects, determine the relative efficacy of SPT and MPT in a clinical environment, and provide insights into participant characteristics that may impact treatment outcomes with the ultimate goal of informing clinical practice relative to selection of AOS treatments.

Method The participants were four individuals with chronic, moderately severe AOS. All were diagnosed with concomitant aphasia using the Western Aphasia Battery-R (Kertesz, 2007). A nonconcurrent multiple baseline ABACA/ACABA design across participants and behaviors was employed to examine the effects of each treatment on whole word correctness (WWC) for acquisition and generalization targets. The treating clinician/primary investigator (PI) conducted probes consisting of treatment targets, response generalization targets, and untreated control targets across all treatment phases. These target items, which were real multisyllabic words, were probed at every visit. Items were presented in a random order, with the participant asked to repeat the PI's verbal production. Following baseline, there were two treatment periods (SPT and MPT), with a washout period between treatments. The treatment order was randomized for the first participant and then alternated for the remaining participants. Treatments were applied to two sets of target stimuli and conducted three times per week, for five weeks per intervention. Treatment followed the respective protocols as published. Follow-up probing was conducted two weeks post-treatment. The primary dependent variable was WWC, which encompassed articulatory accuracy and prosodic features of word production, after Strand et al. (2006) and Maas et al. (2012). A secondary dependent variable of communicative participation, the Communication Participation Item Bank (CPIB; Baylor et al., 2013), was administered at three time points in the study (in each A phase).

Results Two participants (P1 and P3) responded with a statistically significant change in the treated probes for both treatments and in the response generalization probe only during SPT in this investigation (see Figure 1 for an example of results for one participant). Another participant (P2) experienced statistically significant differences with the response generalization probe during SPT. The fourth participant (P4) experienced statistically significant changes with the target and response generalization in the treated set during MPT.

Discussion Final results compare the behavioral change and outcomes associated with the two interventions for all participants using several performance indicators (e.g., slope, effect sizes, maintenance, changes in accuracy levels, CPIB). The present findings suggest that SPT results in greater learning than MPT for most speakers with AOS. The presentation will address findings relative to the literature, discuss clinical implications, and provide directions for future research.



Participant 1 Treatment Targets (Note: The red and green lines within the graph demonstrate the use of the conservative dual-criterion method to assist with visual analysis.)

References

Bailey, D. J., Eatchel, K., & Wambaugh, J. (2015). Sound production treatment: Synthesis and quantification of outcomes. *American Journal of Speech-Language Pathology, 24*, S798-S814.

Baylor, C., Yorkston, K., Eadie, T., Kim, J., Chung, H., & Amtmann, D. (2013). The communication participation item bank (CPIB): Item bank calibration and development of a disorder-generic short-form. *Journal of Speech, Language, and Hearing Research, 56*(4), 1190-1208.

Brendel, B., & Ziegler, W. (2008). Effectiveness of metrical pacing in the treatment of apraxia of speech. *Aphasiology, 22*(1), 77-102.

Fisher, W. W., Kelley, M. E., & Lomas, J. E. (2003). Visual aids and structured criteria for improving visual inspection and interpretation of single-case designs. *Journal of Applied Behavior Analysis, 36*, 387–406.

Kertesz, A., Kertesz, A., Raven, J. C., & PsychCorp (Firm). (2007). WAB-R: Western Aphasia Battery-Revised. PsychCorp.

Maas, E., Butalla, C. E., & Farinella, K. A. (2012). Feedback frequency in treatment for childhood apraxia of speech. *American Journal of Speech-Language Pathology, 21*, 239-257.

Strand, E. A., Stoeckel, R., & Baas, B. (2006). Treatment of severe childhood apraxia of speech: A treatment efficacy study. *Journal of Medical Speech-Language Pathology, 14*(4), 297-307.

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.

THE EFFECT OF MASKING NOISE ON ORAL CANCER SPEECH ACOUSTICS AND KINEMATICS

Teja Rebernik[1,2], Bence M. Halpern[2,3,4], Thomas B. Tienkamp[1], Roel Jonkers[1], Aude Noiray[5,6], Rob J. J. H. van Son[2,3], Michiel W. M. van den Brekel[2,3], Max J. H. Witjes[7], Martijn Wieling[1,6]

[1] *University of Groningen, Groningen, The Netherlands*
[2] *Netherlands Cancer Institute, Amsterdam, The Netherlands*
[3] *University of Amsterdam, Amsterdam, The Netherlands*
[4] *Delft University of Technology, Delft, The Netherlands*
[5] *Laboratoire Dynamique du Langage (DDL), Lyon, France*
[6] *Haskins laboratories, New Haven, CT, United States of America*
[7] *University Medical Centre Groningen, Groningen, The Netherlands*

Introduction When speaking, adults predominantly rely on their feedforward system and can produce relatively unimpaired speech even in the absence of auditory and somatosensory feedback (Schliesser & Coleman, 1968). If auditory or somatosensory feedback are consistently altered, speakers adapt their feedforward maps to produce speech in line with their acoustic and articulatory targets. However, relatively little is known about long-term speech motor adaptation in adulthood if such a change in feedback becomes permanent. The goal of the study is to investigate the functioning of the feedforward mechanism in adults who had undergone a partial glossectomy and/or mandibulectomy. It is known that these procedures limit tongue flexibility and movement (e.g., Bressmann et al., 2004), which also has an impact on speech. We aim to decipher how these speakers develop a new speaking strategy, and how adaptable the speech motor control system is in adulthood.

Methods We collected data from 12 native Dutch speakers above 50 years of age who have undergone surgical treatment for oral cancer with stage 3 or 4 tumours (six female) as well as seven control speakers (three female). Participants produced both unique sentences and target words embedded in carrier phrases. Their speech productions were recorded acoustically, using a Sennheiser ME 66 shotgun microphone, and kinematically, using a VOX-EMA electromagnetic articulograph (see Rebernik et al., 2021, for accuracy assessment). For the current analysis, each speaker produced ten repetitions of five real words (/sok/, /ʃok/, /bi:t/, /bu:t/ and /ba:t/) embedded in the following carrier phrase 'Hij heeft tamme target word gezegd' ('He said target word'). Half of the repetitions were produced while the participant was wearing headphones through which pink noise was played at a comfortable volume, masking their auditory feedback.

For our acoustic analysis, we manually extracted the first and second formants of the corner vowels /a, i, u/ and calculated the triangular vowel space area (tVSA). For our kinematic analysis, we investigated the articulation of sibilants /s/ and /ʃ/. We used Mview (developed by Mark Tiede, Haskins Laboratories) to identify and extract kinematic landmarks in the tongue tip trajectory, including gestural onset, peak velocity, target and gestural offset.

Predictions Masking noise results in increased speech loudness (i.e., the Lombard effect). Based on prior studies investigating kinematic differences in loud vs. normal speech, we generally expect faster and bigger lingual and jaw movement for loud speech (e.g., Kearney et al., 2017). Additionally, for partial glossectomy/mandibulectomy speakers, we expect smaller gestures due to decreased range of movement in non-perturbed speech, and larger endpoint variability under masking noise.

References

Bressmann, T., Sader, R., Whitehill, T. L., & Samman, N. (2004). Consonant intelligibility and tongue motility in patients with partial glossectomy. *Journal of Oral and Maxillofacial Surgery, 62*(3), 298-303. https://doi.org/10.1016/j.joms.2003.04.017

Kearney, E., Giles, R., Haworth, B., Faloutsos, P., Baljko, M., & Yunusova, Y. (2017). Sentence-level movements in Parkinson's disease: loud, clear, and slow speech. *Journal of Speech, Language, and Hearing Research, 60*(12), 3426-3440. https://doi.org/10.1044/2017_JSLHR-S-17-0075

Rebernik, T., Jacobi, J., Tiede, M., & Wieling, M. (2021). Accuracy Assessment of Two Electromagnetic Articulographs: Northern Digital Inc. WAVE and Northern Digital Inc. VOX. *Journal of Speech, Language, and Hearing Research, 64*(7), 2637-2667. https://doi.org/10.1044/2021_JSLHR-20-00394

Schliesser, H. F., & Coleman, R. O. (1968). Effectiveness of certain procedures for alteration of auditory and oral tactile sensation for speech. *Perceptual and Motor Skills, 26*, 175-281. https://doi.org/10.2466/pms.1968.26.1.275

THE EFFICACY OF ACOUSTIC-BASED ARTICULATORY IMPAIRMENT PHENOTYPES FOR CHARACTERIZING AND CLASSIFYING DIVERGENT NEURODEGENERATIVE DISEASES

Hannah Rowe[1], Perman Gochyyev[1], Adam Lammert, Anja Lowit[2], Kristie Spencer[3], Jordan Green[1]
[1]*MGH Institute of Health Professions, Boston, Massachusetts, United States of America*
[2]*Strathclyde University, Glasgow, United Kingdom*
[3]*University of Washington, Seattle, Washington, United States of America*

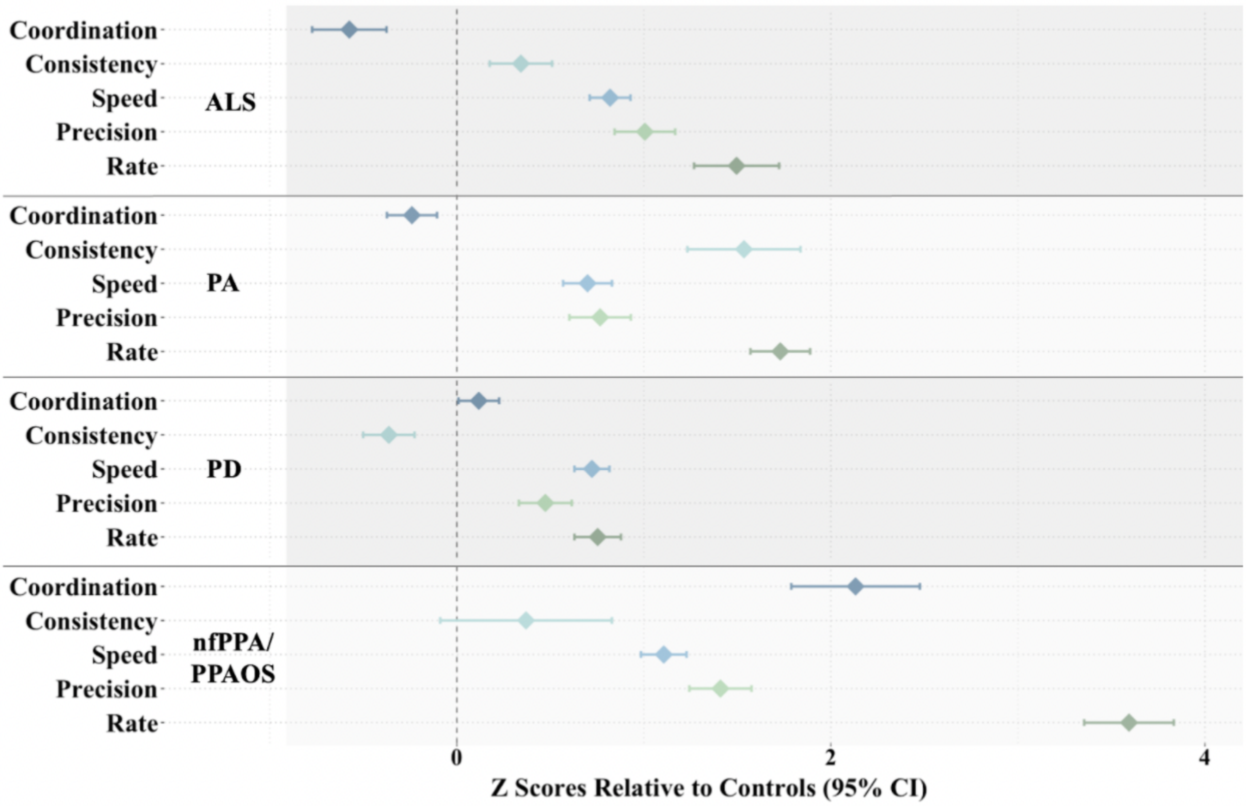
Introduction Neurodegenerative diseases often lead to communication deficits that can have devastating effects on patient quality of life (Batista & Pereira, 2016; Hartelius et al., 2008). Yet, there is no validated set of objective measures that can comprehensively characterize speech abnormalities (Green et al., 2013). Quantitative profiles of neurodegenerative diseases are important for (1) informing differential diagnosis and (2) identifying phenotypic variability that may motivate the development of more targeted treatments. Features from the articulatory subsystem are of particular interest because articulatory deficits are closely related to intelligibility loss (Rong et al., 2015; Lee et al., 2014; De Bodt et al., 2002; Weismer et al., 2001) and have demonstrated diagnostic efficacy for different movement disorders (Green et al., 2013; Cordella et al., 2017; Takakura et al., 2019; Shellikeri et al., 2016). However, the large number of descriptors used for articulation has created challenges with interpreting measures and deciphering the constructs that may be important for differential diagnosis and treatment (Berisha et al., 2021; Miller, 1992). Thus, our recent work sought to develop a unifying framework (i.e., Coordination, Consistency, Speed, Precision, and Rate) for characterizing the diversity of articulatory abnormalities. In the current study, we used this framework to comprehensively profile articulatory impairments in four divergent neurodegenerative diseases: amyotrophic lateral sclerosis (ALS), progressive ataxia (PA), Parkinson's disease (PD), and nonfluent primary progressive aphasia with progressive apraxia of speech (nfPPA+PAOS). We then assessed the efficacy of articulatory phenotyping for disease classification. Our research questions were:

1. (a) What are the articulatory phenotypes of neurodegenerative populations known to have divergent speech motor deficits?
   (b) Do the features contribute unique information to the characterization of articulatory function?
2. (a) What is the efficacy of articulatory phenotypes for classifying divergent neurodegenerative diseases?
   (b) Is a profile of articulatory features more diagnostically useful than individual features?

Methods Speech samples were obtained from MGH and from several outside collaborators. Acoustic measures were extracted from audio recordings of 217 participants (46 with ALS, 52 with PA, 60 with PD, 20 with nfPPA+PAOS, and 39 controls) during the sequential motion rate task. Acoustic profiles were comprised of the proportion of stop-gap to syllable length (*Coordination*), across-repetition variability in voice onset time (*Consistency*), second formant slope in /ka/ (*Speed*), across-consonant variability in second formant slopes of /pa/, /ta/, /ka/ (*Precision*), and syllables per second (*Rate*).

Results The primary features that distinguished the groups were reduced *Consistency* for PA, preserved *Rate* for PD, and reduced *Coordination* and *Rate* for nfPPA+PAOS (see Figure 1). Pairwise correlations revealed weak to moderate correlations for all pairs of acoustic features, suggesting that the features may contribute unique information to the characterization of articulatory function. Univariate and multivariate linear discriminant analyses were then used to determine the classification accuracy for each feature alone and all five features together. ROC curve analyses revealed that disease classification accuracy was strongest when all features were included in the model.

Discussion Overall, there was evidence of articulatory impairment phenotypes for the four populations, and the phenotypes demonstrated high classification accuracy for three of the four groups. Our findings highlight the phenotypic variability present across neurodegenerative motor diseases that may (1) provide complementary information for differential diagnosis and (2) inform the development of individualized speech treatments and sensitive outcome measures.



Forest plot demonstrating articulatory performance across diseases. Each dot represents the mean z score (diagnosis group compared to controls) for each articulatory component. Z scores to the right of the vertical dotted line indicate that the diagnosis group is more impaired than controls.

References

Batista, P. and A. Pereira, "Quality of Life in Patients with Neurodegenerative Diseases," *J. Neurol. Neurosci., vol. 7*, no. 1, 2016.

Berisha, V. et al., "Digital medicine and the curse of dimensionality," *Npj Digit. Med., vol. 4*, no. 1, p. 153, Dec. 2021.

Cordella, C., B. C. Dickerson, M. Quimby, Y. Yunusova, and J. R. Green, "Slowed articulation rate is a sensitive diagnostic marker for identifying non-fluent primary progressive aphasia," *Aphasiology, vol. 31*, no. 2, pp. 241–260, Feb. 2017

De Bodt, M. S., M. E. Hernández-Diaz Huici, and P. H. Van De Heyning, "Intelligibility as a linear combination of dimensions in dysarthric speech," *J. Commun. Disord., vol. 35*, no. 3, pp. 283–292, May 2002

Green, J. R. et al., "Bulbar and speech motor assessment in ALS: Challenges and future directions," Amyotroph. Lateral Scler. *Front. Degener., vol. 14*, no. 7–8, pp. 494–500, Dec. 2013.

Hartelius, L., M. Elmberg, R. Holm, A.-S. Lövberg, and S. Nikolaidis, "Living with Dysarthria: Evaluation of a Self-Report Questionnaire," *Folia Phoniatr. Logop., vol. 60*, no. 1, pp. 11–19, 2008

Lee, J., K. C. Hustad, and G. Weismer, "Predicting Speech Intelligibility With a Multiple Speech Subsystems Approach in Children With Cerebral Palsy," *J. Speech Lang. Hear. Res., vol. 57*, no. 5, pp. 1666–1678, Oct. 2014

Miller, N. "Variability in speech dyspraxia," *Clinical Linguistics & Phonetics, vol. 6*, no. 1+2, pp. 77–85, 1992.

Rong, P., Y. Yunusova, J. Wang, and J. R. Green, "Predicting Early Bulbar Decline in Amyotrophic Lateral Sclerosis: A Speech Subsystem Approach," *Behav. Neurol., vol. 2015*, pp. 1–11, 2015

Shellikeri, S. et al., "Speech Movement Measures as Markers of Bulbar Disease in Amyotrophic Lateral Sclerosis," *J. Speech Lang. Hear. Res., vol. 59*, no. 5, pp. 887–899, Oct. 2016

Takakura, Y. et al., "Sub-classification of apraxia of speech in patients with cerebrovascular and neurodegenerative diseases," *Brain Cogn., vol. 130*, pp. 1–10, Mar. 2019

Weismer, G., J.-Y. Jeng, J. S. Laures, R. D. Kent, and J. F. Kent, "Acoustic and Intelligibility Characteristics of Sentence Production in Neurogenic Speech Disorders," *Folia Phoniatr. Logop., vol. 53*, no. 1, pp. 1–18, 2001

EXAMINING THE IMPACT OF DYSARTHRIA ON COMMUNICATIVE PARTICIPATION IN PAKISTANI CULTURE

Nameeka Shahid, Anja Lowit, Anja Kuschmann
University of Strathclyde, Glasgow, United Kingdom

Background Acquired dysarthria can contribute significantly to changed communication patterns. This can result in increased barriers to communication leading to isolation, a decline in psychological well-being, and decreased quality of life (Palmer et al., 2019). Individual experiences of disability vary based on cultural differences as different cultures have different perceptions with regard to impairments and/or disabilities. (Rao et al., 2007). Whilst there is a growing body of research that has examined the impact of dysarthria on communicative participation in Western cultures, little research has been conducted on this subject in Asian cultures including Pakistan.

Aim The aim of the study was to investigate the impact of acquired dysarthria on the communicative participation of people living in Pakistan, to determine cultural differences and how these impact patients living with dysarthria.

Methodology Twenty-six adults diagnosed with acquired dysarthria following stroke, Parkinson's disease, multiple sclerosis, or traumatic brain injury took part in semi-structured interviews and completed four self-reporting questionnaires — Communication Participation Item Bank (CPIB), Dysarthria Impact Profile (DIP), Voice Handicap Index (VHI) and Short Form-36 (SF-36). The interview data were analyzed using thematic analysis.

Results The results from the questionnaires showed that all the participants indicated the impact of dysarthria on their communication. Several themes emerged from the interviews including Independence, Psychological issues, Awareness, Healthcare, Socioeconomic factors, Gender roles, Attitudes as well as religion and traditions. Findings from the interviews uncovered the lived experiences of people with dysarthria in Pakistan including barriers and challenges, and the impact of attitudes, and cultural, psychological, and social practices on their lives. Interviews revealed that in Pakistan disability is regarded negatively and is associated with stigma and public displays of sympathy such as sighs and prayers. Cultural perceptions of gender roles increase the impact of dysarthria in Pakistan further, in particular for the female population, most of whom cannot make decisions on their own. A change of role for men is also viewed negatively in Pakistani culture. The interviews further revealed a lack of awareness among participants regarding their conditions, and lack of awareness regarding communication disorders in the general population, exacerbating the impact of dysarthria in this culture further.

Conclusion The results highlight the importance of examining cultural issues when investigating the impact of acquired dysarthria on communicative participation. This study provides valuable information about cultural perceptions of motor speech disorders and related disabilities in Pakistan. The findings will be useful for clinicians in Pakistan to better understand the factors that might affect their patients' psychosocial wellbeing. Additionally, the findings will be valuable for clinicians working with Pakistani immigrant populations in contextualizing cultural influences on patient well-being and communicative participation.

References

Palmer, A., Carder, P., White, D., Saunders, G., Woo, H., Graville, D., & Newsom, J. (2019). The Impact of Communication Impairments on the Social Relationships of Older Adults: Pathways to Psychological Well-Being. *Journal Of Speech, Language, And Hearing Research, 62*(1), 1-21. https://doi.org/10.1044/2018_jslhr-s-17-0495

Rao, D., Feinglass, J., & Corrigan, P. (2007). Racial and Ethnic Disparities in Mental Illness Stigma. *Journal Of Nervous & Mental Disease, 195*(12), 1020-1023. https://doi.org/10.1097/nmd.0b013e31815c046e

Dickson, S., Barbour, R. S., Brady, M., Clark, A. M., & Paton, G. (2008). Patient's experiences of disruptions associated with post-stroke dysarthria. *International Journal of Language & Communication Disorders, 43*(2), 135-153. https://doi.org/10.1080/13682820701862228

Mackenzie, C., & Lowit, A. (2007). Behavioural intervention effects in dysarthria following stroke: communication effectiveness, intelligibility and dysarthria impact. *International Journal of Language & Communication Disorders, 42*(2), 131-153. https://doi.org/10.1080/13682820600861776

Walshe, M., Peach, R. K., & Miller, N. (2009a). Dysarthria Impact Profile: development of a scale to measure psychosocial effects. *International Journal of Language & Communication Disorders, 44*(5), 693-715. https://doi.org/10.1080/13682820802317536

Duffy, Joseph R. *Motor Speech Disorders: Substrates, Differential Diagnosis, and Management,* Elsevier, 2013

King, R. B & Bernardo, A.I.B (Eds); 2016. The Psychology of Asian Learners)

Masasa, T.L. Cultural beliefs towards disability: their influence on rehabilitation.

Pakistan Bureau of Statistics. Population Census. Available from http://www.pbs.gov.pk/content/population-census [Accessed on 7th October 2019]

Clarke, V. & Braun, V. (2018). Using thematic analysis in counselling and psychotherapy research: A critical reflection. *Counselling and Psychotherapy Research Journal, 18*(2), 107-110.

EFFECTS ON VOICE SOUND LEVEL AND DYSPHONIA AFTER HICOMMUNICATION: A NOVEL SPEECH, VOICE AND COMMUNICATION TREATMENT FOR PARKINSON'S DISEASE

Hanna Steurer[1,2], Ellika Schalling[1,3]
[1]*Karolinska Institutet, Stockholm, Sweden*
[2]*Stockholms Sjukhem, Stockholm, Sweden*
[3]*Uppsala University, Uppsala, Sweden*

Rationale Speech changes are common in Parkinson's disease (PD) and often result in hypokinetic dysarthria, a motor speech disorder where reduced vocal loudness and impaired voice quality (dysphonia) are common symptoms. HiCommunication is a novel group-intervention targeting speech, voice, and communication for people with PD. The program is based on principles driving neuroplasticity and motor learning and high compliance and acceptability has been shown (Schalling et al., 2021). The aim of the present study was to evaluate the effects on voice sound level and present some results of acoustic analysis after intervention with HiCommunication.

Methods Participants with PD (n=95) were recruited within a randomized controlled trial (RCT) with a double-blind design (Franzén et al., 2019) and randomly allocated to intervention with either HiCommunication or HiBalance, a program targeting balance and gait. Acoustic analyses were performed pre and post intervention.

A linear multilevel model was used for analyses of voice sound level including data from all participants with missing values imputed. Between group Cohen's d was calculated. A clinically relevant change was defined as a change pre to post intervention of $\geq$ 2 dBC based on a study in which participants with PD used a voice sound level that was 2-4 dB lower compared to healthy controls (Fox & Ramig, 1997). Acoustic Voice Quality Index (AVQI) (version 01.03, Phonanium, 2021) analyses were performed. The AVQI is a multivariate construct that combines several acoustic parameters to obtain a single score for the estimation of overall voice quality (Maryn et al., 2009). A reliable change (i.e., not due to measurement error or test-retest variability) is defined as $\geq$ 0.54 between two AVQI scores (Barsties & Maryn, 2013). Clopper Pearson exact confidence intervals for the proportion participants with a clinically relevant change in voice sound level and a reliable change in AVQI were calculated.

Results A significant group by time interaction effect on voice sound level was found where the HiCommunication group showed an increase in comparison to the HiBalance group (unstandardized b = -2.1 [95% CI = -3.4, -0.9], p = 0.0012) with the between group Cohen's d estimated to -0.54 (Freidle et al., 2022). The percentage of participants with a clinically relevant increase in voice sound level ($\geq$2 dBC) after intervention with HiCommunication was 59%, CI [41%, 75%]) and after intervention with HiBalance 13%, CI [4%, 28%]). The AVQI average of all participants at baseline was 4.1 which exceeds the cut-off value for dysphonia (2.95). The percentage of participants with a reliable decrease in AVQI ($\geq$ 0.54) after HiCommunication was 43%, CI [26%, 61%] and after HiBalance 21%, CI [10%, 37%].

Discussion This RCT shows positive effects on voice sound level after intervention with HiCommunication. Notably, the positive response differs between individuals and further studies will investigate what factors are associated with a clinically relevant positive change following HiCommunication. The study further indicates potential for AVQI to be used to analyze dysphonia in people with PD.

References

Barsties, B., & Maryn, Y. (2013). Test-retest variability and internal consistency of the acoustic voice quality index. *HNO, 61*(5), 399-403. https://doi.org/10.1007/s00106-012-2649-0

Fox, C. M., & Ramig, L. O. (1997). Vocal sound pressure level and self-perception of speech and voice in men and women with idiopathic Parkinson disease. *American Journal of Speech-Language Pathology, 6*(2), 85-94. https://doi.org/10.1044/1058-0360.0602.85

Franzén, E., Johansson, H., Freidle, M., Ekman, U., Wallén, M. B., Schalling, E., . . . Hagströmer, M. (2019). The expand trial: Effects of exercise and exploring neuroplastic changes in people with Parkinson's disease: A study protocol for a double-blinded randomized controlled trial. *BMC Neurol, 19*(1), 280. https://doi.org/10.1186/s12883-019-1520-2

Freidle, M., Johansson, H., Ekman, U., Lebedev, A. V., Schalling, E., Thompson, W. H., . . . Franzén, E. (2022). Behavioural and neuroplastic effects of a double-blind randomised controlled balance exercise trial in people with Parkinson's disease. *NPJ Parkinson's Disease, 8*(1), 12-12. https://doi.org/10.1038/s41531-021-00269-5

Maryn, Y., De Bodt, M., & Roy, N. (2009). The acoustic voice quality index: Toward improved treatment outcomes assessment in voice disorders. *Journal of Communication Disorders, 43*(3), 161-174. https://doi.org/10.1016/j.jcomdis.2009.12.004

Schalling, E., Winkler, H., & Franzén, E. (2021). Hicommunication as a novel speech and communication treatment for Parkinson's disease: A feasibility study. *Brain and behavior*, e02150-e02150. https://doi.org/10.1002/brb3.2150

PROGRESS TOWARD ESTIMATING THE MINIMAL CLINICALLY IMPORTANT DIFFERENCE OF SPEECH INTELLIGIBILITY: A CROWDSOURCED PERCEPTUAL EXPERIMENT

Kaila L. Stipancic[1], Frits van Brenk[1,2], Mengyang Qiu[1], Kris Tjaden[1]
[1]*University at Buffalo, Buffalo, United States of America*
[2]*Utrecht University, Utrecht, The Netherlands*

Purpose The purpose of this study was to estimate the minimal clinically important difference (MCID) of sentence intelligibility in neurologically healthy control speakers and speakers with dysarthria due to multiple sclerosis (MS) and Parkinson's disease (PD).

Methods Speakers were part of a larger study examining the acoustic and perceptual consequences of MS and PD. 48 speakers were included in the current study (16 healthy control speakers, 16 speakers with MS, and 16 speakers with PD). Speakers were recorded reading Harvard psychoacoustic sentences in speaking conditions purported to modulate intelligibility: habitual, clear, fast, loud, and slow. 240 listeners were recruited using the crowdsourcing website Prolific (prolific.co). A global ratings of change (GROC) scale developed by Jaeschke et al. (1989) was used as an 'external anchor of meaningfulness'. Listeners heard a given speaker in one condition followed immediately by the same speaker in another condition. Listeners used Jaeschke's GROC scale to indicate "how much more understandable?" one condition was over the other and were given response options on the following seven-point scale: (1) Almost the same, hardly any better at all; (2) A little better; (3) Somewhat better; (4) Moderately better; (5) A good deal better; (6) A great deal better; (7) A very great deal better. Previously published transcription intelligibility data for these speakers and stimuli were employed in calculations. Receiver operating characteristics (ROC) curves were used to determine how well the change in intelligibility scores between conditions differentiated between those speakers for whom listeners identified a change in understandability and those for whom listeners did not identify a change in understandability (Stipancic et al., 2018). The average difference in intelligibility for each response option on the GROC scale was defined as the MCID.

Results & Discussion The MCID for a small change in intelligibility was determined to be 7% and for a moderate to large change in intelligibility was determined to be 15%. This work demonstrates feasibility of the novel experimental paradigm for collecting crowdsourced perceptual data for estimating MCIDs of speech outcomes. Findings are a critical step toward development of a universal language with which to evaluate changes in intelligibility as a result of speech-language therapy and disease progression.

References

Jaeschke, R., Singer, J., & Guyatt, G. H. (1989). Measurement of health status. Ascertaining the minimal clinically important difference. *Controlled Clinical Trials, 10*(4), 407–415. `https://doi.org/10.1016/0197-2456(89)90005-6`

Stipancic, K. L., Yunusova, Y., Berry, J. D., & Green, J. R. (2018). Minimally detectable change and minimal clinically important difference of a decline in sentence intelligibility and speaking rate for individuals with amyotrophic lateral sclerosis. *Journal of Speech, Language, and Hearing Research, 61*(11), 2757–2771. `https://10.1044/2018_AJSLP-17-0074`

INFLUENCE OF STIMULI LENGTH ON TONGUE AND LIP MOVEMENT PATTERN STABILITY IN ALS

Kristin J. Teplansky[1], Alan Wisler[2], Daragh Heitzman[3], Sara G. Austin[1], Jun Wang[1]
[1] *University of Texas at Austin, Austin, Texas, United States of America*
[2] *Utah State University, Logan, Utah, United States of America*
[3] *Texas Neurology, Dallas, Texas, United States of America*

Introduction The spatiotemporal index (STI) is a widely used measure of speech movement pattern stability during multiple repetitions a single motor task. Well-controlled articulatory movement patterns are believed to be highly consistent, whereas deviations in neuromuscular activity decreases movement stability. Although prior work has primarily attributed instability to increased demands on the motor system (e.g., task complexity, motor learning, disordered speech; McHenry, 2003; Saletta et al., 2018), it has recently been shown that in absence of behavioral changes, increased length (in syllables) in stimuli increases STI values in simulated data (Wisler et al., 2022), suggesting STI is more sensitive to temporal instability in longer stimuli. Less is known about how articulatory variability is impacted by syllable length in individuals with dysarthria due to ALS. A better understanding of potential influences on spatiotemporal stability has relevance to methodological interpretation and will help us interpret the results from different studies with different stimuli. The purpose of this study was to investigate syllable related changes in articulatory variability in healthy and disordered speech during the production of phrases.

Methods This study included six participants with ALS (4 females, 2 males) and ten age-matched healthy controls (4 females, 6 males) with an average speaking rate of 91.17 wpm (SD = 27.50) and 190.61 (SD = 16.85), respectively. Electromagnetic articulography (Wave System, NDI Inc.) was used to collect positional data from four sensors attached to the tongue tip, tongue back, upper lip and lower lip. Participants produced 3-, 5-, and 8-syllable phrases ('I love you,' 'Thanks for stopping by,' and 'I need to make an appointment') four times amongst other speech tasks. SMASH was used to segment the speech data (Green et al., 2013). To assess tongue and lip movement stability we derived the spatiotemporal index of the displacement signals from each sensor using the approach proposed by (Smith et al., 1995). Descriptive statistics for ALS and healthy controls were analyzed.

Results & Discussion Descriptive statistics are provided in Table 1. Preliminary results showed that on average, healthy controls exhibit greater variability in the 8-syllable stimuli than the 3-syllable. This finding aligns with prior research which showed that more syllables are associated with greater STI values, despite the length of the target phrase in simulated data (Kleinow & Smith, 2006). For speakers with ALS, group mean variability was consistently greater for the 3-syllable stimulus than the 5- and 8- syllable stimuli for the tongue tip, tongue back, and upper lip. Interestingly, there was no change in variability for the lower lip sensor across the phrases.

Our findings provide preliminary evidence that variability differences due to stimuli length in ALS may exist and manifest differently than in healthy speakers. This finding may be due to difficulties initiating motor sequences due to neurodegeneration. We must note that the presented results cannot ambiguate if differences in stability are strictly due to behavioral changes.

Mean ± standard deviation of STI values for 3-, 5-, and 8-syllable phrases separated by articulator.

	Descriptive Statistics				
Group	Number of Syllables	Tongue Tip	Tongue Back	Upper Lip	Lower Lip
Healthy	3	17.59 ± 7.91	14.43 ± 7.11	26.85 ± 12.96	13.60 ± 6.40
	5	14.30 ± 5.09	16.34 ± 7.26	20.67 ± 6.94	18.29 ± 5.63
	8	19.22 ± 5.19	20.88 ± 5.54	30.30 ± 9.99	22.69 ± 7.20
ALS	3	26.47 ± 10.71	22.01 ± 10.02	33.62 ± 15.86	21.48 ± 10.49
	5	18.18 ± 5.04	20.78 ± 9.73	28.75 ± 12.59	21.29 ± 5.81
	8	19.25 ± 6.23	18.83 ± 6.93	25.40 ± 10.90	21.12 ± 7.10

References

Green, J. R., Wang, J., & Wilson, D. L. (2013). SMASH: A tool for articulatory data processing and analysis. *Interspeech*, 1331–1335.

Kleinow, J., & Smith, A. (2006). Potential interactions among linguistic, autonomic, and motor factors in speech. *The Journal of the International Society for Developmental Psychobiology, 48*(4), 275–287. https://doi.org/doi/10.1002/dev.20141

McHenry, M. A. (2003). The effect of pacing strategies on the variability of speech movement sequences in dysarthria. *Journal of Speech, Language, and Hearing Research, 46*(3), 702–710. https://doi.org/10.1044/1092-4388(2003/055)

Saletta, M., Goffman, L., Ward, C., & Oleson, J. (2018). Influence of language load on speech motor skill in children with specific language impairment. *Journal of Speech, Language, and Hearing Research, 61*(3), 675–689. https://doi.org/10.1044/2017_JSLHR-L-17-0066

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). https://doi.org/10.1007/BF00231983

Wisler, A., Goffman, L., Zhang, L., & Wang, J. (2022). Influences of methodological decisions on assessing the spatiotemporal stability of speech movement sequences. *Journal of Speech, Language, and Hearing Research, 65*(2), 538–554. https://doi.org/10.1044/2021_JSLHR-21-00298

AFFECTED TONGUE BODY MOVEMENTS IN IDIOPATHIC REM SLEEP BEHAVIOR DISORDER

Tabea Thies[1], Nuria Geerts[1], Doris Mücke[1], Aline Seger[1,2], Michael T. Barbe[1], Michael Sommerauer[1,2]
[1] *University of Cologne, Cologne, Germany*
[2] *Forschungszentrum Jülich,Jülich, Germany*

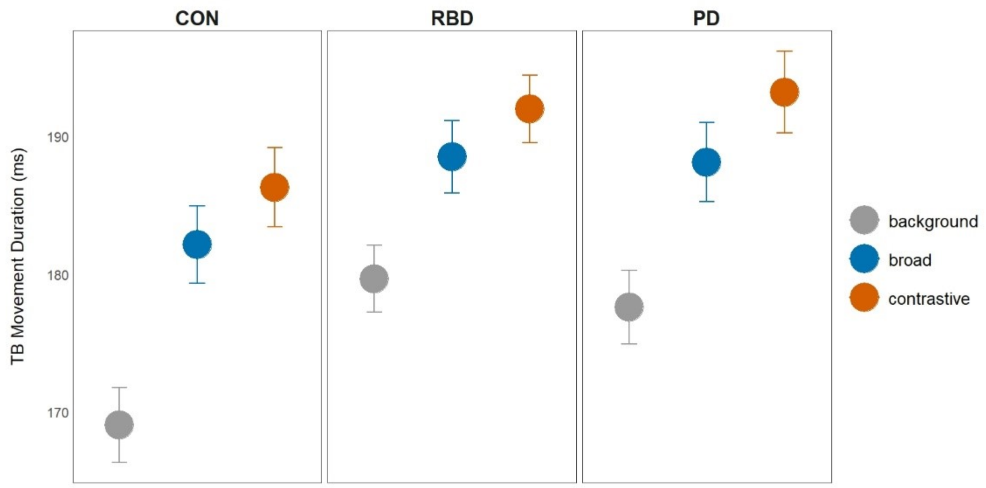
Introduction Patients with isolated REM sleep behavior disorder (iRBD) have an increased risk of developing Parkinson's disease (PD) in the future. In PD, recognition of speech deviations often follows clinical diagnosis. But what if speech might function as a prodromal biomarker for PD? Speech production is a sensitive domain that can provide early information about deficits in motor control. Previous acoustic studies have shown that articulatory deficits (Rusz et al., 2016) and a reduced pitch range can be found in patients with iRBD (Rusz et al., 2021) compared to healthy control speakers. Those two speech areas are also impaired in patients with PD indicating a hypokinetic dysarthria (Duffy, 2019). The present study examines if speech movements of patients with iRBD differ from patients with PD.

Methods Data were collected from 66 age- and sex-matched subjects which were divided into three groups: healthy control speakers (n=22), patients with iRBD (n=22), and patients with PD (n=22). Patients with PD were assessed in medication-OFF condition after suspending dopaminergic drugs for at least 12 hours.

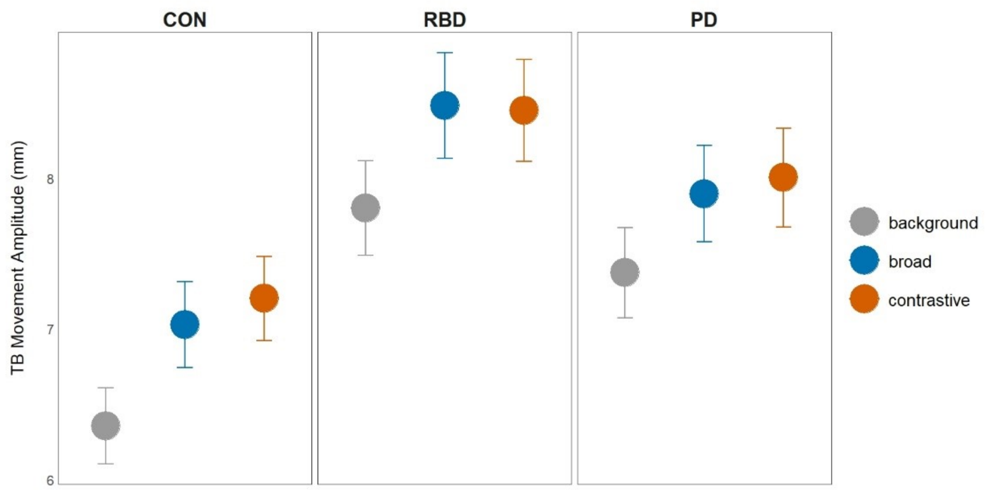
All participants were recorded with an electromagnetic articulograph (AG 501) to capture tongue and lip movements. Ten different target words, such as Mila or Lina (C1V1.C2V2-structure), embedded in a carrier sentence were produced by the speakers in three different controlled focus conditions. The target syllables contain peripheral vowels of German /i, e, a, o, u/ flanked by alveolar and labial consonants (/m, l/). In total 1980 items went into the analysis (66 speaker x 10 words x 3 conditions). Kinematic tongue and lip measures (durations, amplitudes, peak velocities) were calculated. In addition, the motor status of all subjects were determined using the motor part III of the Unified Parkinson's Disease Rating Scale (UPDRS III, Fahn & Elton, 1987).

Results Results indicate that both, patients with iRBD and with PD need more time to achieve the articulatory target of the tongue body during vowel production compared to healthy controls (Figure 1). Interestingly, when looking at maximal velocities, tongue body movements are faster and larger in patients with iRBD compared to controls (Figure 2). Patients with PD show overall increased movement durations and slower velocities of the tongue tip and the lip. The distinguishing feature between iRBD and PD are longer and slower tongue tip movements which are present in the PD group but not in the iRBD group.

Conclusion In patients with iRBD only the tongue body is affected, whereas in patients with PD the whole tongue is affected. This might indicate an evolution towards dysarthria from dorsal to ventral parts of the tongue. Patients with iRBD seem to compensate for longer movements with faster peak velocities and larger amplitudes. Affected tongue body movements in patients with iRBD might indicate articulatory distortion as an early sign of hypokinetic dysarthria (e.g., imprecise consonants, articulated with the tongue tip; Duffy, 2019).



Duration of tongue body (TB) movement during vowel production in three different focus conditions. Means and standard errors are presented. CON = healthy control speakers.



Amplitude of tongue body (TB) movement during vowel production in three different focus conditions. Means and standard errors are presented. CON = healthy control speakers.

References

Duffy, J. R. (2019). *Motor Speech Disorders: Substrates, Differential Diagnosis, and Management*, 4th ed. Edinburgh: Elsevier.

Fahn, S., & Elton, R. L. (1987). UPDRS program members. Unified Parkinsons disease rating scale. *Recent developments in Parkinson's disease, 2*, 153-163.

Rusz, J., Hlavnička, J., Tykalová, T., Bušková, J., Ulmanová, O., Růžička, E., & Šonka, K. (2016). Quantitative assessment of motor speech abnormalities in idiopathic rapid eye movement sleep behaviour disorder. *Sleep medicine, 19*, 141-147.

Rusz, J., Hlavnička, J., Novotný, M., Tykalová, T., Pelletier, A., Montplaisir, J., ... & Šonka, K. (2021). Speech Biomarkers in Rapid Eye Movement Sleep Behavior Disorder and Parkinson Disease. *Annals of Neurology, 90*, 62-75.

QUANTIFYING CHANGES IN ARTICULATORY WORKING SPACE FOLLOWING ORAL CANCER TREATMENT

Thomas B. Tienkamp[1,5], Teja Rebernik[1,3], Bence M. Halpern[2,3,4], Defne Abur[1], Rob, J.J.H. van Son[2,3], Sebastiaan A.H.J. de Visscher[5], Max J.H. Witjes[5], and Martijn, B. Wieling[1,6]

[1] *University of Groningen, Groningen, Netherlands*

[2] *University of Amsterdam, Amsterdam, Netherlands*

[3] *Netherlands Cancer Institute, Amsterdam, Netherlands*

[4] *Delft University of Technology, Delft, The Netherlands*

[5] *University Medical Centre Groningen, Groningen, The Netherlands*

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

Background Surgical treatment of tumours in the oral cavity results in physiological changes which may complicate articulation. The tongue becomes less mobile due to scar tissue and potential postoperative radiation therapy and it has been shown that a less mobile tongue is associated with reduced intelligibility (Bressmann et al., 2004; Jacobi et al., 2013). Moreover, tissue loss may create a bypass for airflow or limit constriction possibilities. Yet, the articulatory consequences of treatment have only received scant attention. In general, patients exhibit less complex lingual movements and reduced lateral Range of Motion (ROM) post-surgery (Hagedorn et al., 2021; Kappert et al., 2019). However, no study has investigated the overall Articulatory Working Space (AWS), defined here as a convex hull area in mm2, or anterior-posterior and superior-inferior ROM of oral cancer patients during speech, even though this type of data may serve as a direct evaluation of treatment-induced changes in lingual movement (Lee & Bell, 2018). The objective of this study is to investigate the AWS and ROM of Dutch individuals who were surgically treated for oral cancer, in order to assess whether the AWS and ROM decrease post-treatment as compared to control speakers.

Design and method Data collection and analysis are still ongoing. So far, nine patients (five female, mean age = 61.6 years) who were surgically treated for a stage 3 or 4 squamous cell carcinoma of the oral cavity have taken part in this study, together with five control speakers (three female, mean age = 60.6 years). Patients varied between 1 and 14 years post-treatment and were all native speakers of Dutch. Data is collected using a VOX-EMA electromagnetic articulograph (EMA) and a Sennheiser ME 66 shotgun microphone. Stimuli include a wide array of unique sentences, varying from the North Wind and the Sun passage to news items. In our ongoing analysis, we measure the AWS (convex hull in mm2), anterior-posterior and superior-inferior ROM patterns in mm of the tongue-tip, tongue-back, and jaw sensor. Data is normalised in order to correct for differences in oral cavity size.

Hypotheses and predictions We hypothesise that surgical treatment affects the AWS and ROM patterns in patients post-treatment (Bressmann et al., 2004; Hagedorn et al., 2021). Specifically, we predict that the overall AWS as well as the posterior-anterior ROM decrease post-treatment. Conversely, we do not expect significant changes in the superior-inferior ROM as a result of treatment. This is because prior studies on vowel formants in this population have shown that F2 (the approximate acoustic correlate of posterior-anterior movement) is more affected than F1 (reflecting superior-inferior movement) (de Bruijn et al., 2009; van Son et al., 2018).

References

Bressmann, T., Sader, R., Whitehill, T. L., & Samman, N. (2004). Consonant intelligibility and tongue motility in patients with partial glossectomy. *Journal of Oral and Maxillofacial Surgery, 62*(3), 298–303. https://doi.org/10.1016/j.joms.2003.04.017

de Bruijn, M. J., ten Bosch, L., Kuik, D. J., Quené, H., Langendijk, J. A., Leemans, C. R., & Verdonck-de Leeuw, I. M. (2009). Objective Acoustic-Phonetic Speech Analysis in Patients Treated for Oral or Oropharyngeal Cancer. *Folia Phoniatrica et Logopaedica, 61*(3), 180–187. https://doi.org/10.1159/000219953

Hagedorn, C., Kim, J., Sinha, U., Goldstein, L., & Narayanan, S. S. (2021). Complexity of vocal tract shaping in glossectomy patients and typical speakers: A principal component analysis. *The Journal of the Acoustical Society of America, 149*(6), 4437–4449. https://doi.org/10.1121/10.0004789

Lee, J., & Bell, M. (2018). Articulatory range of movement in individuals with dysarthria secondary to amyotrophic lateral sclerosis. *American journal of speech-language pathology, 27*(3), 996-1009. https://doi.org/10.1044/2018_AJSLP-17-0064

Jacobi, I., van Rossum, M. A., van der Molen, L., Hilgers, F. J., & van den Brekel, M. W. (2013). Acoustic analysis of changes in articulation proficiency in patients with advanced head and neck cancer treated with chemoradiotherapy. *Annals of Otology, Rhinology & Laryngology, 122* (12), 754-762. https://doi.org/10.1177/000348941312201205

Kappert, K. D. R., van Alphen, M. J. A., Smeele, L. E., Balm, A. J. M., & van der Heijden, F. (2019). Quantification of tongue mobility impairment using optical tracking in patients after receiving primary surgery or chemoradiation. *PloS one, 14*(8), e0221593. https://doi.org/10.1371/journal.pone.0221593

van Son, R., Middag, C., & Demuynck, K. (2018). Vowel Space as a Tool to Evaluate Articulation Problems. *Interspeech 2018*, 357–361. https://doi.org/10.21437/Interspeech.2018-68

SPEECH IN PREMANIFEST AND EARLY-STAGE HUNTINGTON'S DISEASE

Adam P. Vogel[1], Jess Chan[1], Geoff Stuart[2],Paul Maruff[3],Yenni Lie[4],Jullie Stout[5]
[1]*The University of Melbourne, Melbourne, Australia*
[2]*Redenlab, Melbourne, Australia*
[3]*Cogstate Ltd., Australia*
[4]*Calvary Health Care Bethlehem, Parkdale, Australia*
[5]*Monash University, Clayton, Australia*

Background Clinical markers that show change in performance in people with Huntington's disease (HD) during the presymptomatic and prodromal stages remain a target of investigation in clinical medicine. Alongside genetic and neuroimaging initiatives, digital speech analytics has shown promise as a sensitive clinical marker of premanifest HD.

Objective To investigate the sensitivity of digital speech measures for detecting subtle cognitive-linguistic and fine motor features in people carrying the expanded HD gene, with and without symptoms.

Methods Speech data were acquired from 110 participants (55 people with the expanded HD gene including 16 presymptomatic HD; 16 prodromal HD; 14 early-stage HD; 9 mid-stage HD; and 55 matched healthy controls). Objective digital speech measures were derived from speech tasks that fit along a continuum of motor and cognitive complexity. Acoustic features quantified speakers' articulatory agility, voice quality and speech-timing. Subjects also completed the tests of cognition and upper limb motor function.

Results Some presymptomatic HD (furthest from disease onset) differed to healthy controls. Prodromal HD presented with reduced articulatory agility, reduced speech rate and longer and variable pauses. Speech agility correlated with poorer performance on the upper limb motor test.

Conclusion Tasks with a mix of cognitive and motor demands differentiated premanifest HD from their matched control groups. Motor speech tasks alone did not differentiate groups until participants were relatively closer to disease onset or symptomatic. Data demonstrated how ubiquitous behaviors like speech, when analyzed objectively, provide insight into disease related change.

ANALYSIS OF SPEECH MOVEMENTS DURING METRONOME-TIMED SPEECH IN PEOPLE WHO STUTTER USING VOCAL-TRACT IMAGING

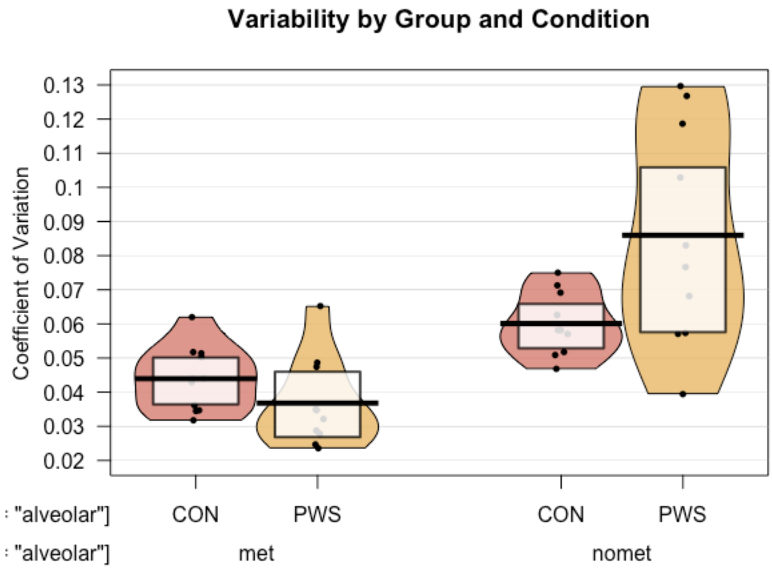
Charlotte E.E. Wiltshire[1,2], Gabriel J. Cler[1,3], Mark Chiew[1], Jennifer Chesters[1,4], Philip Hoole[2], Kate E. Watkins[1]
[1] *University of Oxford, United Kingdom*
[2] *Ludwig-Maximilians-University, Munich, Germany*
[3] *University of Washington, Seattle, Washington, United States of America*
[4] *Southmead Hospital, Bristol, United Kingdom*

Several studies indicate that people who stutter show greater variability in speech movements than people who do not stutter, even when the speech produced is perceptually fluent (Jackson, Tiede, Beal, & Whalen, 2016; Smith, Sadagopan, Walsh, & Weber-Fox, 2010; Wiltshire, Chiew, Chesters, Healy, & Watkins, 2021). Speaking to the beat of a metronome reliably increases fluency in people who stutter, regardless of the severity of stuttering (Boutsen, Brutten, & Watts, 2000; Chesters, Watkins, & Möttönen, 2017; Frankford et al., 2021). Here, we aimed to test whether this fluency-enhancer also reduces articulatory variability.

We scanned the vocal tracts of 34 people who stutter and 22 controls using MRI while participants repeated sentences with or without a metronome. Sentences contained the target words: artillery, catastrophe, impossibility (Ogar et al., 2006). For this analysis, we compared data from 10 people per group. Midsagittal images of the vocal tract from lips to larynx were reconstructed at 33.3 frames per second (Wiltshire et al., 2021). Any utterances containing dysfluencies or other non-speech movements were excluded. For each participant, we measured the variability of movements from the alveolar and palatal regions of the vocal tract. In previous studies, variability has been measured using two contrasting methods: (1) The coefficient of variation (CoV; Wiltshire et al., 2021), calculated by dividing the standard deviation of the summed amplitude of the raw signal, by the mean and (2) the Spatial Temporal Index (STI; Smith, Goffman, Zelaznik, Ying, & McGillem, 1995), calculated by normalizing the movement traces in time and amplitude before summing the standard deviation at a sample of 50 time points (Smith et al., 1995). Here, we calculate both measures and compare them.

In line with our predictions, there was an interaction between group and condition ($f(1)=8.6$, $p=.003$) such that people who stutter had more variability than control speakers during no-metronome speech, which was then reduced to the same level as controls when speaking with the metronome (fig 1).

In addition, there was a strong correlation between STI and CoV measures of variability ($r(128) = .66$, $p < .0001$). There was no difference in variability between alveolar and palatal regions of the vocal tract ($p = .41$). These results replicate previous findings of greater variability in the movements of people who stutter compared with controls during normal speaking (no-met condition). Furthermore, we show that metronome timed speech reduces variability in people who stutter to same level as control speakers.



Variability (coefficient of variation) of movements in the alveolar region for control speakers and people who stutter during metronome and no-metronome speaking conditions. Horizontal line shows mean, error shows inter-quartile range.

References

Boutsen, F. R., Brutten, G. J., & Watts, C. R. (2000). Timing and Intensity Variability in the Metronomic Speech of Stuttering and Nonstuttering Speakers. *Journal of Speech, Language, and Hearing Research, 43*(2), 513–520. https://doi.org/10.1044/jslhr.4302.513

Chesters, J., Watkins, K. E., & Möttönen, R. (2017). Investigating the feasibility of using transcranial direct current stimulation to enhance fluency in people who stutter. *Brain and Language, 164*, 68–76. https://doi.org/10.1016/j.bandl.2016.10.003

Frankford, S. A., Heller Murray, E. S., Masapollo, M., Cai, S., Tourville, J. A., Nieto-Castañón, A., & Guenther, F. H. (2021). The Neural Circuitry Underlying the "Rhythm Effect" in Stuttering. *Journal of Speech, Language, and Hearing Research, 64*(6S), 2325–2346. https://doi.org/10.1044/2021_JSLHR-20-00328

Jackson, E. S., Tiede, M., Beal, D., & Whalen, D. H. (2016). The impact of social–cognitive stress on speech variability, determinism, and stability in adults who do and do not stutter. *Journal of Speech, Language, and Hearing Research, 59*(6), 1295–1314. https://doi.org/10.1044/2016_JSLHR-S-16-0145

Ogar, J., Willock, S., Baldo, J., Wilkins, D., Ludy, C., & Dronkers, N. (2006). Clinical and anatomical correlates of apraxia of speech. *Brain and Language, 97*, 343–350. https://doi.org/10.1016/j.bandl.2006.01.008

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. https://doi.org/10.1007/BF00231983

Smith, A., Sadagopan, N., Walsh, B., & Weber-Fox, C. (2010). Increasing phonological complexity reveals heightened instability in inter-articulatory coordination in adults who stutter. *Journal of Fluency Disorders, 35*(1), 1–18. https://doi.org/10.1016/j.jfludis.2009.12.001

Wiltshire, C. E. E., Chiew, M., Chesters, J., Healy, M. P., & Watkins, K. E. (2021). Speech Movement Variability in People Who Stutter: A Vocal Tract Magnetic Resonance Imaging Study. *Journal of Speech, Language, and Hearing Research, 64*(7), 2438–2452. https://doi.org/10.1044/2021_JSLHR-20-00507

THE PREVALENCE OF APRAXIA OF SPEECH IN CHRONIC APHASIA AFTER STROKE

Wolfram Ziegler[1], Anja Staiger[1]*, Ingrid Aichert[1]*, Klaus Willmes[2], Annette Baumgaertner[3], Tanja Grewe[4], Agnes Flöel[5,6], Walter Huber[2], Roman Rocker[7], Catharina Korsukewitz[7], Caterina Breitenstein[7]

[1]*Ludwig-Maximilians-University, Munich, Germany*

[2]*RWTH Aachen University Clinic, Aachen, Germany*

[3]*University of Luebeck, Luebeck, Germany*

[4]*Jade University of Applied Sciences, Oldenburg, Germany*

[5] *University Medicine Greifswald, Greifswald, Germany*

[6]*German Center for Neurodegenerative Diseases (DZNE), Greifswald, Germany*

[7]*University of Muenster, Muenster, Germany*

**equal contribution*

Background and objectives Apraxia of speech (AOS) is a motor speech disorder that occurs after lesions to the left cerebral hemisphere, most often concomitant with aphasia (Ziegler et al., 2012). It requires specific approaches in the study of its theoretical basis and special expertise in clinical care. Knowing its prevalence in patients with aphasia (PWA) after stroke is therefore relevant for planning specific resources in clinical research and in health care provision. This is the first systematic study of the frequency of this condition in chronic post-stroke aphasia.

Methods The frequency of AOS was examined in a representative sample of 156 patients with chronic post-stroke aphasia (Breitenstein et al., 2017). Three experts classified the patients' speech by best-practice auditory-perceptual methods. Bayesian hierarchical models were fitted to obtain probability distributions for prevalence estimates (Lambert, 2018). A prior distribution was calculated in two steps, including Bayesian models for published frequency data (step 1) and prevalence estimates from experienced clinicians (step 2). Separate models were fitted with three different severity filters.

Results The overall prevalence rate was 0.44 (maximum density), with a 90% credible interval of [0.30, 0.58]. When only moderate and severe cases were considered, the rate was 0.35 [0.23, 0.49]. After restriction to only severe impairment, the prevalence dropped to 0.22 [0.12, 0.34]. Patients identified with AOS had suffered more severe strokes according to clinical criteria and had more severe aphasias. The presence of AOS was predicted by the articulation/prosody and syntax rating scales of the Aachen Aphasia Test.

Discussion The estimate that almost half of the patients with aphasia lasting longer than six months have at least mild AOS highlights the need for increased development and planning of specific interventions in the care of this population, as well as regular updating of pertinent guidelines. Lower prevalence estimates published earlier (ASHA, 2019) are probably biased by low sensitivity of assessment instruments for mild speech impairment.

References

American Speech-Language-Hearing Association. (2019). *National Outcomes Measurement System: Adults in Healthcare – Inpatient Rehab National Data Report 2019*. In:National Center for Evidence-Based Practice in Communication Disorders. `https://www.asha.org/noms/noms-data-reports/`

Breitenstein C, Grewe T, Flöel A, Ziegler W, Springer L, Martus P, Huber W, Willmes K, Ringelstein EB, Haeusler KG et al. (2017). Intensive speech and language therapy in patients with chronic aphasia after stroke: a randomised, open-label, blinded-endpoint, controlled trial in a health-care setting. *The Lancet, 389*, 1528-1538.

Lambert B (2018) *A Student's Guide to Bayesian Statistics*. Sage, Los Angeles, CA.

Ziegler W, Aichert I, & Staiger A (2012) Apraxia of speech: Concepts and controversies. *Journal of Speech, Language, and Hearing Research, 55*, S1485-S1501

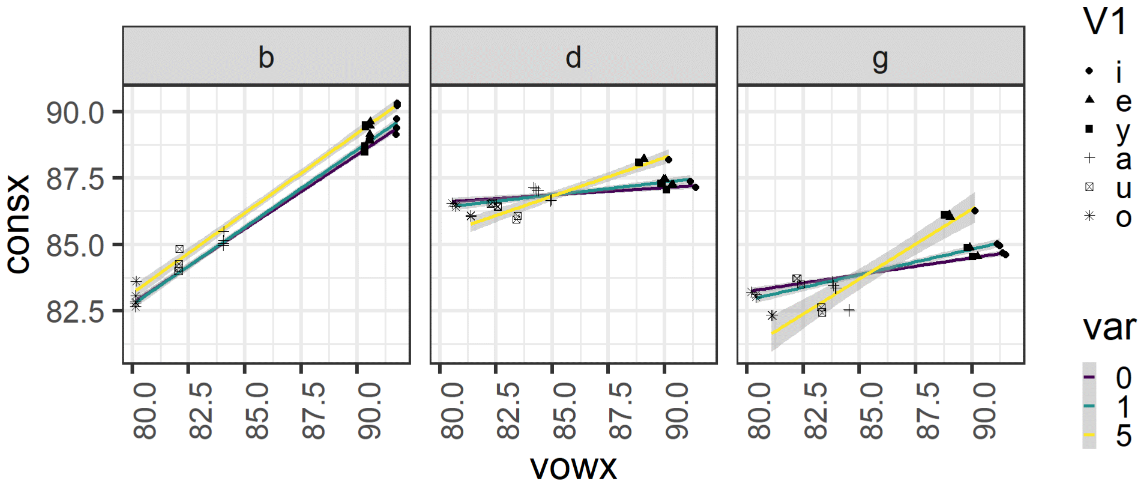
LINKING DIFFERENCES IN PHONOLOGICAL REPRESENTATIONS AND COARTICULATION DEGREE: A MODELLING APPROACH

Dzhuma Abakarova[1,2], Susanne Fuchs[2], Aude Noiray[3,4]
[1]*University of Potsdam, Potsdam, Germany*
[2]*Leibniz-Centre General Linguistics (ZAS), Berlin, Germany*
[3]*Laboratoire Dynamique du Langage, CNRS UMR 5596, Lyon, France*
[4]*Haskins Laboratories, New Haven, CT, United States of America*

There is converging evidence across languages that children show greater coarticulation degrees (CD) i.e., greater spatial overlap between consecutive segments than adults (e.g., Noiray et al., 2013; Noiray et al., 2019a; Nittrouer, 1989; Zharkova et al., 2011). Previous research has suggested that CD reflects the nature of phonological representations, and so related developmental differences in CD to the emergent awareness of phonemes (Nittrouer et al., 1989). Recent findings of a negative correlation between CD and phonological awareness (Noiray et al., 2019b) provide support for this view. However, phonological awareness develops in parallel to motor control, vocabulary and various other cognitive abilities, making it difficult to experimentally control for the independent contribution of each of these factors into CD differences. That is why the present study employs dynamic modelling to address the question of whether developmental differences in coarticulation degree and phonological representations are causally related.

In our understanding, phonological representations are sensorimotor in nature. They are associations between sensory cues and motor plans. Children's initial motor representations are more variable than adults' with less clearly defined boundaries between categories and they lack an abstract level of categorization. With the growing vocabulary and practice, children's motor representations for various segments become more delimited. Over time, the optimal trajectories are selected and the ones that lead to overlap and compromise intelligibility or take more effort are discarded. Therefore, in children, the space of motor realizations for a segment is larger than in adults. We hypothesize that greater variability in phonological representations allows for greater spatial overlap between co-produced segments and results in the observed age differences in CD. To test this hypothesis, we conducted a series of simulations with the Task Dynamics Application (TaDA, Nam et al., 2004). To enable comparison with experimental findings, we simulated the utterances from Noiray et al. (2019) dataset. The dataset consisted in ultrasound tongue imaging data collected in 3- to -7-year-old children and adults producing CV syllables (/b/, /d/, or /g/ in six vowel contexts), and preceded by a carrier word /aɪnə/.

The age differences in variability of representations were simulated by manipulating the variability of the gestural targets in TADA. Gestural targets determine constriction locations within the vocal tract (i.e., further front or back). In the default TADA settings, they are point targets, but here they are modeled as target regions of varying size to simulate the difference between smaller and bigger space of motor realizations for a segment. Preliminary results suggest that the simulations with higher variability in gestural targets result in higher CD, as was observed for children in the experimental data (Figure 1). The effect is rather marginal for /b/, which does not involve a tongue gesture, while variability of /d/ and /g/ has a more pronounced effect on CD. This work contributes to an improved understanding of phonological representations and their development. It also helps dissect the relative role of developing phonological representations with respect to the experimentally observed age-related changes in coarticulation patterns.



The linear regression of the horizontal position of the tongue body at the acoustically determined temporal midpoint of the consonant (consx) on the position of tongue body at acoustically determined temporal midpoint of the subsequent vowel (vowx). The regressions were fit for three consonants (/b, d, g/), each followed by one of the six vowels (/i:, y:, u:, a:, e:, o:/), for three simulated variability values (labelled "var"). A steeper slope of the line corresponds to a higher coarticulation degree.

References

Browman, C. P., & Goldstein, L. (1989). Articulatory gestures as phonological units. *Phonology, 6*(2), 201–251. https://doi.org/10.1017/S0952675700001019

Goldstein, L., Byrd, D., & Saltzman, E. (2006). The role of vocal tract gestural action units in understanding the evolution of phonology. In M. A. Arbib (Ed.), *Action to Language via the Mirror Neuron System* (pp. 215–249). Cambridge University Press. https://doi.org/10.1017/CBO9780511541599.008

Nam, H., Goldstein, L., Saltzman, E., & Byrd, D. (2004). TADA: An enhanced, portable Task Dynamics model in MATLAB. *The Journal of the Acoustical Society of America, 115*(5), 2430. https://doi.org/10.1121/1.4781490

Nittrouer, S., Studdert-Kennedy, M., & McGowan Richard, S. (1989). The emergence of phonetic segments. *Journal of Speech, Language, and Hearing Research, 32*(1), 120–132. https://doi.org/10.1044/jshr.3201.120

Noiray, A., Popescu, A., Killmer, H., Rubertus, E., Krüger, S., & Hintermeier, L. (2019). Spoken Language Development and the Challenge of Skill Integration. *Frontiers in Psychology, 10*. https://doi.org/10.3389/fpsyg.2019.02777

Noiray, A., Wieling, M., Abakarova, D., Rubertus, E., & Tiede, M. (2019). Back from the future: Nonlinear anticipation in adults' and children's speech. *Journal of Speech, Language, and Hearing Research: JSLHR, 62*(8S). https://doi.org/10.1044/2019_JSLHR-S-CSMC7-18-0208

Redford, M. A. (2019). Speech Production From a Developmental Perspective. *Journal of Speech, Language, and Hearing Research, 62*(8S), 2946–2962. https://doi.org/10.1044/2019_JSLHR-S-CSMC7-18-0130

Vihman, M. M., & Keren-Portnoy, T. (2013). T*he Emergence of Phonology: Whole-word Approaches and Cross-linguistic Evidence*. Cambridge University Press.

PERCEPTUAL LEARNING OF DYSARTHRIC SPEECH WITH ADOLESCENT POPULATIONS

Stephanie A. Borrie[1], Taylor Hepworth[1], Kaitlin L. Lansford[2], Camille J. Wynn[1], Katherine Hustad[3]
[1]*Utah State University, Logan, Utah, United States of America*
[2]*Florida State University, Tallahassee, Florida, United States of America*
[3]*University of Wisconsin-Madison, Madison, United States of America*

Background Perceptual learning describes the idea that, with experience, listeners can learn to better understand degraded or otherwise noncanonical speech. Theoretical models posit that experience affords the listener an opportunity to acquire distributional knowledge of the degraded speech signal (Kleinschmidt & Jaeger, 2015). A large body of literature has experimentally examined perceptual learning of dysarthric speech in adult populations: neurotypical adult listeners familiarized with the speech of adults with dysarthria show significant improvements in intelligibility performance relative to those familiarized with neurotypical speech (see Borrie & Lansford, 2021, for a review). Collectively, the results support a novel approach for adult speech remediation: listener-targeted remediation to improve intelligibility of speakers with dysarthria. Up until now, this work has not been extended to adolescent populations. In the current study, we examine whether adult and adolescent listeners can learn to better understand an adolescent speaker with dysarthria. Given theoretical and empirical evidence of a dynamic and adaptable speech perception system, we hypothesize intelligibility benefits for all listeners (adults and adolescents) familiarized with adolescent dysarthric speech. However, given the still-developing speech perception systems of adolescents (Huyck, 2018), we hypothesize that intelligibility improvements will be smaller for adolescent listeners relative to adult listeners.

Methods Eighty-two normal-hearing, neurotypical listeners (42 adults, 40 adolescents) completed a standard three-phase perceptual training protocol (pretest, familiarization, posttest). During the pretest and posttest phases, listeners heard audio recordings of stimuli produced by a 13-year-old speaker with cerebral palsy and spastic dysarthria and reported what they thought was being said. During the familiarization phase, listeners were presented with audio recordings of the same speaker with dysarthric (training) or neurotypical speech (control) while following along with lexical transcripts of the spoken targets. Orthographic transcriptions of listener responses from the pretest and posttest phases were scored for an objective measure of intelligibility, percent words correct, using Autoscore, a validated open-source tool for automated scoring of words correct (Borrie et al., 2019).

Results & Conclusions Linear regression models revealed no significant difference in pretest intelligibility scores by condition (training vs. control) and age (adult vs. adolescent). Thus, listeners were statistically comparable at initially understanding the adolescent speaker with dysarthria. Contrastingly, linear regression models revealed significantly higher posttest scores for training vs. control conditions. Thus, listeners familiarized with adolescent dysarthric speech achieved greater intelligibility improvements than listeners familiarized with neurotypical speech. This was true of both adult and adolescent groups, though adults achieved significantly higher posttest scores. The study extends previous findings of improved intelligibility following dysarthria familiarization

to adolescent populations, namely adolescents with dysarthria and their adolescent peers. Results will be discussed relative to learning theory, adolescent development, and implications for dysarthria management.

References

Borrie S.A. & Lansford, K.L. (2021). A perceptual learning approach for dysarthria remediation: An updated review of experimental studies. *Journal of Speech, Language, and Hearing Research, 6*, 3060–3073. https://doi.org/10.1044/2021_JSLHR-21-00012

Borrie, S.A., Barrett, T.S., & Yoho, S.E. (2019). Autoscore: An open-source automated tool for scoring listener perception of speech. *Journal of Acoustical Society of America, 145*, 392–399. https://doi.org/10.1121/1.5087276

Huyck, J.J. (2018). Comprehension of degraded speech matures during adolescence, *Journal of Speech, Language, and Hearing Research, 61*, 1012–1022. https://doi.org/10.1044/2018_JSLHR-H-17-0252

Kleinschmidt, D.F., & Jaeger, T.F. (2015). Robust speech perception: Recognize the familiar, generalize to the similar, and adapt to the novel. *Psychological Review, 122*(2), 148–203. https://doi.org/10.1037/a0038695

ACOUSTIC MEASURES OF ONLINE AND OFFLINE COLLECTED SPEECH PRODUCED BY CHILDREN WITH CP AND DYSARTHRIA: STEPS TOWARDS VALIDATION

Frits van Brenk[1], Kyung Hae Hwang[2], Jiyoung Choi[2], Rachael Brisman[2], Eunice Hong[2], Erika S. Levy[2]
[1]*Utrecht University, Utrecht, The Netherlands*
[2]*Columbia University, New York, United States*

Purpose More than 50 percent of children with cerebral palsy (CP) present with dysarthria, affecting their functional communication. Access to clinical assessment and treatment and to research participation might be obstructed due to reduced mobility, geographical distance and pandemic-related restrictions on movement. Online organized treatment programs may overcome these obstacles. The current study aimed to take initial steps towards validating the analysis of a series of acoustic parameters pertinent to describing dysarthric speech in CP. Outcome measures of speech collected online (with varying microphone quality and transmission bandwidth) and offline (research-grade quality) equipment were compared to determine validity of online data collection.

Method
Participants: Fifteen American-English children with dysarthria due to CP (mean age = 10;8; range 7;6–14;0; 7 females, 8 males) participated in the study. CP subtypes included spastic and mixed, and dysarthria severity ranged from mild to severe.

Data collection: Speech data were collected at the child's home. Research-grade recording equipment and a sound level meter were forwarded to each child's caregiver. The researcher met with the child and caregiver remotely via the online conference platform Zoom, and guided them in setting up equipment and completing the testing protocol. Children repeated pre-recorded utterances produced by an adult native-speaker of American English. At word level, the children repeated 12 contrastive words in a carrier phrase. At sentence level, they repeated four phrases from the Test of Children's Speech (Hodge et al., 2009). Speech productions were recorded simultaneously using 1) the microphone available through the tablet or computer they used for Zoom conferencing and 2) the research-grade audio-recording equipment that had been sent to the parents.

Acoustic measures and analysis: Acoustic measures were quasi-automatically extracted using custom Praat scripts (Boersma & Weenink, 2020), and included articulation rate, sound pressure level (SPL), fundamental frequency variation, and second formant (F2) values of corner vowels for measuring vowel space areas. Manually obtained measures included F2 range of diphthongs and fricative-affricate duration differences (Lee et al., 2014; Levy et al., 2017). Pearson correlations were employed to compare outcome measures across recording devices, with correlations $r > .7$ as cut-off for validity.

Results Preliminary results indicate that at both word and sentence level, articulation rate, SPL, F2 range of diphthongs, and fricative-duration difference measures showed reliably strong positive correlations between recording devices. F0 variation failed to show reliable agreement between devices. Data with respect to vowel space areas are currently being analyzed, with anticipated data collection and analysis of these measures to have been completed by Spring 2022.

Discussion Findings from the current study revealed a number of strong associations between acoustic measures obtained from online and offline recording equipment. These results indicate that online speech recording may have potential as valid alternative to in-person recording methods for children with CP. These findings will be an important step towards building the evidence base of on-line speech evaluation and, potentially, treatment, supporting clinicians and researchers to consider employing methods of remote data collection for clinical research, assessment or monitoring treatment progress.

References

Boersma, P., & Weenink, D. (2020). *Praat: Doing Phonetics by Computer* (6.1.09) [Computer software]. Phonetic Sciences, University of Amsterdam. http://praat.org

Hodge, M., Daniels, J., & Gotzke, C. L. (2009). *TOCS+ intelligibility measures (Version 5.3)*[Computer software]. Edmonton, Canada: University of Alberta.

Lee, J., Hustad, K. C., & Weismer, G. (2014). Predicting speech intelligibility with a multiple speech subsystems approach in children with cerebral palsy. *Journal of Speech, Language, and Hearing Research, 57*(5), 1666–1678.

Levy, E. S., Chang, Y. M., Ancelle, J. A., & McAuliffe, M. J. (2017). Acoustic and perceptual consequences of speech cues for children with dysarthria. *Journal of Speech, Language, and Hearing Research, 60*(6S), 1766–1779.

VOWEL DISTINCTIVENESS IS RELATED TO EXPRESSIVE LANGUAGE IN LOW- AND MINIMALLY VERBAL AUTISTIC CHILDREN

Karen V. Chenausky[1,2,3], Helen Tager-Flusberg[3], Jordan Green[1]
[1]*MGH Institute of Health Professions, Boston, Massachusetts, USA*
[2]*Harvard Medical School, Boston, Massachusetts, USA*
[3]*Boston University, Boston, Massachusetts, USA*

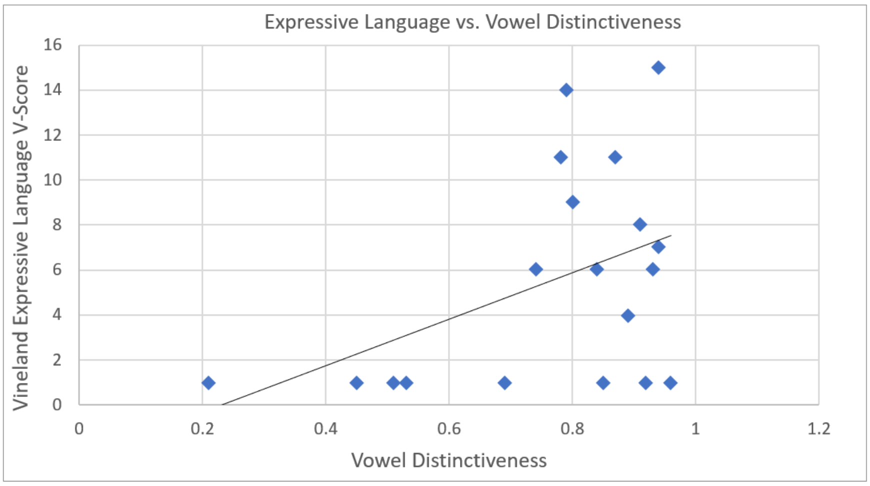
Up to one-quarter of low- and minimally verbal (LMV) autistic children experience comorbid childhood apraxia of speech, and their speech production ability significantly predicts concurrent expressive language (EL) ability (Chenausky et al., 2019). We investigated whether an acoustic measure of vowel production, obtained via remote data collection, can predict concurrent EL scores. Vowels are compelling candidates for this purpose because they develop earlier than consonants and may therefore function as an early indicator of risk for future EL impairment. We employed a normalized acoustic measure of vowel distinctiveness (Fourakis et al., 1993), making it applicable to children of any age. Our language measures were from the Vineland Adaptive Behavior Scales, a parent questionnaire appropriate for children who produce very little speech.

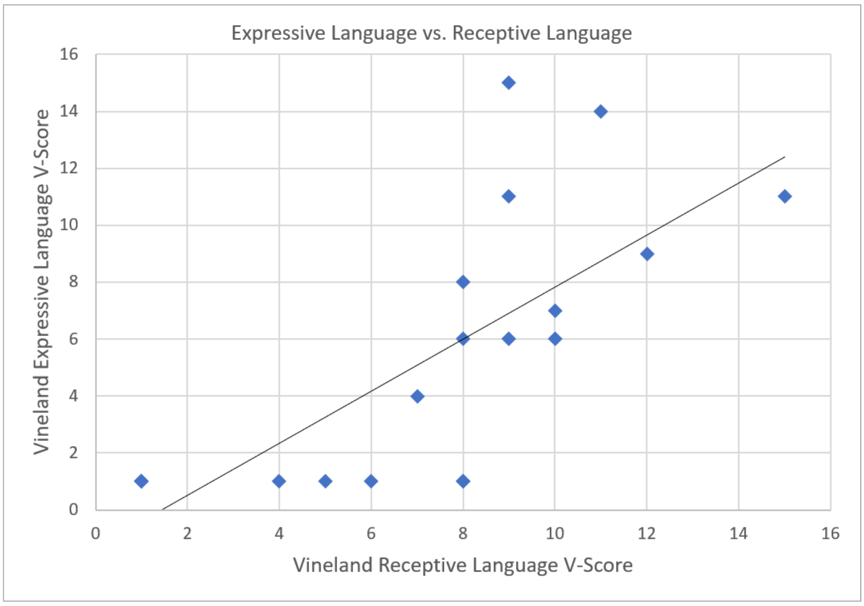
19 children with LMV ASD (6 F) aged 4;2-7;10 (mean 6;4, SD 1;0) repeated the syllables [bi], [ba], [bu], [bæ], and [bʌ] at least three times each. Parents recorded children's speech in WAV format using a smartphone app and uploaded the audio to a secure server. Files were downsampled to 16 kHz/16 bits using Audacity. F0, F1, F2,and F3 were manually measured in the middle third of each vowel using Wavesurfer.

For each vowel target, mean values for F0 and F1-3 were calculated. Then, each mean vowel was represented as a point in 3-space using a measure that normalizes for vocal tract size differences. The distance from /i/ to /a/ was then calculated to produce a Vowel Distinctiveness score for each child. Vowel Distinctiveness for typically developing children was also calculated from normative figures for 7-year-olds (Assmann & Katz, 2000), for comparison. Pearson's correlations were used to assess the relationship of five descriptive variables to EL: Vowel Distinctiveness, chronological age, nonverbal IQ, receptive language (RL) score, EL score, and ASD severity. Variables with significant correlations were then regressed on EL score.

Mean EL score for our sample was 10.3 (SD 21.6) and mean RL score was 12.2 (SD 21.0), vs. the typical means of 15 (SD 3). Mean Vowel Distinctiveness was 0.8 (SD 0.2), vs. 0.9 for typical children. EL score was significantly correlated with RL score ($r = 0.64$, $p = 0.002$) and Vowel Distinctiveness ($r = 0.49$, $p = 0.016$) but not with age, nonverbal IQ, or autism severity. RL score and Vowel Distinctiveness significantly predicted concurrent EL score ($F(2,18) = 15.610$, $p < 0.001$) and accounted for 40.6% and 25.5% of the variance in EL score respectively (adjusted R^2).

The finding that vowel distinctiveness contributes to expressive language in at least some autistic children is consistent with previous work relating speech repetition more generally to expressive language. The fact that this relationship still holds for remotely-collected data suggests that remote data collection paradigms are feasible and may yield larger sample sizes than is possible for in-person data. Future work will compare acoustic to perceptual measures of vowel distinctiveness as predictors of concurrent expressive language and investigate the utility of both types of vowel distinctiveness measures for predicting future expressive language ability.





References

Assmann, P. & Katz, W. (2000). Time-varying spectral change in the vowels of children and adults. *JASA 108*(4), 1856–1866.

Chenausky K, Brignell A, Morgan A, & Tager-Flusberg H. (2019) Motor speech impairment predicts expressive language in minimally verbal, but not low verbal, individuals with autism spectrum disorder. *ADLI 4*:1-12.

Fourakis M, Geers A, & Tobey E. (1993) An acoustic metric for assessing change in vowel production for profoundly hearing-impaired children. *JASA 94*(5), 2544-2552.

USING THE APRAXIA OF SPEECH RATING SCALE TO CHARACTERIZE RESOLVING CHILDHOOD APRAXIA OF SPEECH

Heather M. Clark, Becky S. Baas, Taylor M. Brown, Tess M. Hansen
Mayo Clinic, Rochester, Minnesota, United Stated of America

Background The Apraxia of Speech Rating Scale (ASRS) in its previous (Clark et al., 2016; Josephs et al., 2012) and current (Strand et al., 2014) forms is reliable and valid in the diagnosis and description of adult speakers with apraxia of speech (AOS) (Josephs et al., 2012; Utianski et al., 2018; Duffy et al., 2017; Josephs et al., 2013). The Dynamic Evaluation of Motor Speech Skill (DEMSS) (Basilakos et al., 2015) is validated in children, and uses a cut-off score to identify severe CAS or CAS in children with limited speech (CAS). This project explored the usefulness of the ASRS for characterizing motor speech impairment in children with a history of CAS.

Method Participants in this IRB-approved retrospective study were children with a diagnosis of CAS who were administered the ASRS during their clinical visit. The ASRS involves rating 13 features during connected speech (conversation, picture description, word and sentence repetition), and diadochokinetic repetition of syllables – alternate motion rates, AMRs – and syllable sequences, SMRs. Along with standard demographic information, ASRS scores were abstracted from the clinical record. If available from the same visit, DEMSS scores were recorded as well. Descriptive analysis was adopted for this small data set.

Results Thirteen children (8 male), mean age 8.4 years, were identified. DEMMS scores were available for 8 participants. No participants exhibited dysarthria, but many had developmental phonologic speech sound disorders and/or language disorders. All but 3 children had Total Scores on the DEMMS were above the cutoff score for "little or no evidence for CAS," although speech sound and prosodic errors were evident in each child's exam. On the ASRS, the item "Increased sound distortions or distorted sound substitutions with increased utterance length or complexity" elicited the highest scores (indicating the greatest level of impairment) across participants. More phonetic features (sound distortions, distorted substitutions, and distorted additions) were judged to be present than prosodic features (syllable segmentation was the only prosodic item rated as present in at least half the children). This was further reflected in the relative size of the mean phonetic subscore (6.9) versus prosodic subscore (2.4). Participants had no difficulty with AMRs but at least mild impairment with SMRs was not uncommon. Also noted for just over half the participants was audible groping and false starts.

Discussion The ASRS proved useful for documenting features of motor speech impairment evident in the speech of children with resolving CAS. It is interesting that the overall scores exhibited on the ASRS by these children were relatively low, with the scores of two children falling below the cut-off score of "8" that is recommended for discriminating adults with aphasia with concomitant AOS from those without AOS (Josephs et al., 2012). A larger sample of children with resolving CAS as well as a control sample without motor speech impairment is needed to determine if the ASRS total score can be put to such use. Regardless, the current study suggests profiling speech features during phonetically complex tasks may help characterize resolving motor speech impairments in children with a history of CAS when other tools, designed for children with higher levels

of severity, are no longer sensitive to the subtleties of their impairment.

DEMMS	**Range of possible scores (Higher scores = better performance)**	**Mean Score**
Vowel accuracy	(0-120)	109.8 (92-20)
Prosodic Accuracy	(0-24)	20.3 (12-24)
Overall Accuracy	(0-240)	207 (161-233)
Consistency	(0-42)	31.6 (17-39)
Mean total score	Total possible score = 426 0-323 Significant evidence for CAS 324 – 373 Some evidence for at least mild CAS >374 Little or no evidence for CAS	368.6 (282-412)
ASRS	Scoring Guidelines 0 Not observed in any task 1 Infrequent Frequent but not pervasive 2 Very often evident but not marked in severity 3 nearly always evident but not marked in severity 4 Nearly always evident and/or marked in severity	Mean Score (% of participants demonstrating feature)
Phonetic Features (Mean Phonetic Subscore: 6.9)		
Sound distortions		1.6 (91%)
Distorted sound substitutions		1.3 (91%)
Distorted sound additions (including intrusive schwa)		0.8 (82%)
Increased sound distortions or distorted sound substitutions with increased utterance length or increased complexity		2.3 (91%)
Prosodic Features (Mean Prosodic Subscore: 2.4)		
Syllable segmentation within words >1 syllable		0.4 (36%)
Syllable segmentation across words in phrases/sentences		0.8 (64%)
Slow overall speech rate (apart from pauses for word retrieval)		0.5 (27%)
Lengthened vowel &/or consonant segments		0.5 (36%)
Other		
RATE ONLY FOR AMRs Slow and/or off-target		0 (0%)
RATE ONLY FOR SMRs (Slow, segmented, incorrectly sequenced, and/or off-target)		0.5 (45%)
One or both of the following: Consistently reduced words per breath group		0.1 (9%)
Silent articulatory groping		0.4 (36%)
Audible false starts/restarts or groping including sound repetitions		0.8 (55%)
Mean Total Score		10.9

References

Clark, H.M., et al., Revisions to the Apraxia of Speech Rating Scale, in Conference on Motor Speech. 2016: Newport Beach, CA.

Strand, E.A., et al., The Apraxia of Speech Rating Scale: a tool for diagnosis and description of apraxia of speech. Journal of Communication Disorders, 2014. 51: p. 43-50.

Utianski, R.L., et al., Prosodic and phonetic subtypes of primary progressive apraxia of speech. Brain Lang, 2018. 184: p. 54-65.

Basilakos, A., et al., Patterns of poststroke brain damage that predict speech production errors in apraxia of speech and aphasia dissociate. *Stroke, 2015. 46*(6): p. 1561-6.

Haley, K.L., M. Smith, and J.L. Wambaugh, Sound Distortion Errors in Aphasia With Apraxia of Speech. *Am J Speech Lang Pathol, 2019. 28*(1): p. 121-135.

Wambaugh, J.L., et al., Interrater Reliability and Concurrent Validity for the Apraxia of Speech Rating Scale 3.0: Application With Persons With Acquired Apraxia of Speech and Aphasia. *American Journal of Speech-Language Pathology, 2019. 28*(2S): p. 895-904.

Strand, E. and R. McCauley. 2019. *Dynamic Evaluation of Motor Speech Skill* (DEMSS) Manual. 2019, Baltimore, MD: Brookes.

DIFFERENCES BETWEEN DIADOCHOKINESIS RATES IN CHILDREN ACROSS THREE EUROPEAN LANGUAGES

Sanne Diepeveen[1], Dora Knežević[2], Ben Maassen[3]
[1] *HAN University of Applied Sciences, Nijmegen, The Netherlands*
[2] *University of Zagreb, Zagreb, Croatia*
[3] *University of Groningen, Groningen, The Netherlands*

Speech is a complex motor skill; its assessment carries clinical importance (Van Haaften et al., 2019). One of the few objective assessments of motor speech performance is diadochokinesis (DDK). The DDK-task involves the repetition of syllables composed of a consonant and a vowel as quickly as possible (Diepeveen et al., 2019; Ziegler, 2002). Two types of DDK-tasks are common: Alternating Motion Rates (AMR) - repetition of monosyllabic sequences (/pʌ/, /tʌ/, /kʌ/) and Sequential Motion Rates (SMR) - repetition of multiple syllabic sequence (/pʌtʌkʌ/). These stop-consonants /p/, /t/, and /k/ occur in most languages (Schwartz et al., 2012) and are among the first consonants acquired (McLeod & Crowe, 2018), their use in DDK allows a nearly universal application (Kent et al., 2022). In typically developing (TD) children, the DDK-rate is correlated with age and is viewed as an index of oral motor development (Van Haaften et al., 2019; Diepeveen et al., 2021). There is no uniform method of administrating the DDK or measuring the rate of the sequence(s) which leads to large variability of norm data and makes it difficult to compare between studies (Diepeveen et al., 2019). In addition, cross-linguistic studies of speech rate show that German is typically produced with an average syllable rate of approximately 6 syll/s (Pellegrino et al., 2011) compared to Dutch and Croatian with an average of 5 syll/s (Verhoeven et al., 2004; Horga, 1988), but in different speech elicitation tasks. Although different reference values have been reported (Alshahwan et al., 2020; Icht & Ben-David, 2014; Jothi & Amritha, 2019), other research suggests that the DDK-task could be language-independent (kent et al., 2022) and that it could serve as a universal measure (Moro-Velazquez et al., 2021; Näsström & Schalling, 2020). Several studies have shown gender differences for the DDK-speech rate (Diepeveen et al., 2021). The differences between studies suggest the need to compare DDK-results using the same protocol for multiple languages. Therefore, the goal is to analyze DDK-data collected from Croatian, Dutch, and German-speaking children.

The DDK-task was conducted for a sample of around 50 TD children (4;0 and 6;11) per language. Children were seen by a speech and language pathologist (SLP)/SLPstudent. Caregivers gave informed consent. We applied the protocol described by Diepeveen et al. (2019). The Maximum Repetition Rate (MRR) was defined as the number of syllables uttered per second in three mono-, two bi-, and one trisyllabic sequence. The DDK-task was elicited via a computerized protocol consisting of spoken instructions and examples. For each sequence, at the end of the protocol, two different instructions were given for the children to pronounce the sequence as fast as possible without any example. The fastest sequence of these two latter uttered sequences was included in the analysis. We compared the outcome of the six sequences of the MRR between the three languages with an ANOVA per age group and gender.

We are collecting data until March 2022. Considering the uniform method of conducting the DDK-task we expect no significant differences between DDK data between the Croatian, Dutch, and German-speaking children.

References

Alshahwan, M. I., Cowell, P. E., & Whiteside, S. P. (2020). Diadochokinetic rate in Saudi and Bahraini Arabic speakers: Dialect and the influence of syllable type. *Saudi Journal of Biological Sciences, 27*(1), 303–308.

Diepeveen, S., van Haaften, L. Terband, H. de Swart, B., Maassen., B. (2019). A standardized protocol for Maximum Repetition Rate assessment in children. *Folia Phoniatrica Logopaedica 1*(5-6):238-50

Diepeveen, S., van Haaften, L., Terband, H., de Swart, B., van den Engel-Hoek, L., Maassen, B. (2021). Maximum Repetition Rate in a large cross-sectional sample of typically developing Dutch-speaking children. *International Journal of Speech-Language Pathology, 23*(5), 508-518

Horga, D. (1988). Latentna struktura brzine izgovora. Govor, 5, 2, 129–146

Icht, M., Ben-David, B. M. (2014). Oral-diadochokinesis rates across languages: English and Hebrew norms. *Journal of Communication Disorders, 48*, 27-37.

Jothi, S., & Amritha, M. L. (2019). Comparison of diadochokinetic rate between Malayalam and Tamil native speakers. *International Journal of Research and Review, 6*(5), 144–148

Kent, R. D., Kim, Y., & Chen, L. M. (2022). Oral and Laryngeal Diadochokinesis Across the Life Span: A Scoping Review of Methods, Reference Data, and Clinical Applications. *Journal of Speech, Language, and Hearing Research, 65*(2), 574-623.

McLeod, S., & Crowe, K. (2018). Children's consonant acquisition in 27 languages: A cross-linguistic review. *American Journal of Speech-Language Pathology, 27*(4), 1546–1571.

Moro-Velazquez, L., Gomez-Garcia, J. A., Arias-Londoño, J. D., Dehak, N., & Godino-Llorente, J. I. (2021). Advances in Parkinson's disease detection and assessment using voice and speech: A review of the articulatory and phonatory aspects. *Biomedical Signal Processing and Control, 66*, 102418.

Näsström, A. K., & Schalling, E. (2020). Development of a method for assessment of dysarthria in a foreign language: A pilot study. *Logopedics Phoniatrics Vocology, 45*(1), 39–48.

Pellegrino F., Coupé C., Marsico E. (2011). Across-language perspective on speech information rate. *Language 87* 539–558.

Schwartz, J.-L., Boë, L.-J., Badin, P., & Sawallis, T. R. (2012). Grounding stop place systems in the perceptuo-motor substance of speech: On the universality of the labial–coronal– velar stop series. *Journal of Phonetics, 40*(1), 20–36.

Van Haaften, L., Diepeveen, S., Terband, H., Vermeij, B., van den Engel-Hoek, L., de Swart, B., & Maassen, B. (2019). Profiling Speech Sound Disorders for Clinical Validation of the Computer Articulation Instrument. *American Journal of Speech-Language Pathology, 28*(2S), 844–856.

Verhoeven, J., De Pauw, G., & Kloots, H. (2004). Speech Rate in a Pluricentric Language: A Comparison Between Dutch in Belgium and the Netherlands. *Language and Speech, 47*(3), 297–308.

Ziegler, W. (2002). Task-related factors in oral motor control: Speech and oral diadochokinesis in dysarthria and apraxia of speech. *Brain and Language, 80*, 556–575.

TONGUE SHAPE COMPLEXITY IN CHILDREN WITH SPEECH SOUND DISORDERS

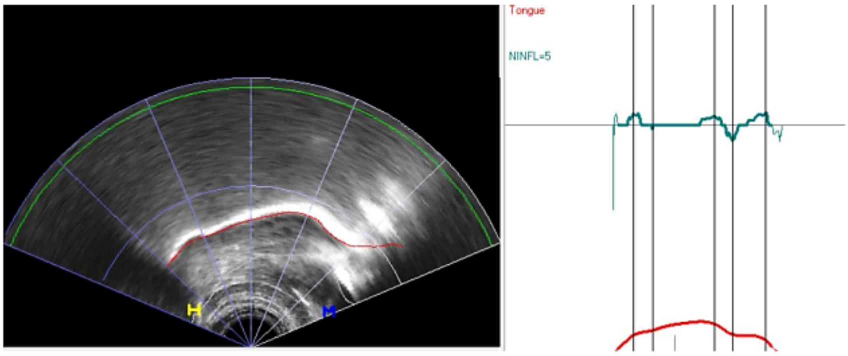
Marie Dokovova[1], Eleanor Sugden[2], Gemma Cartney[3], Joanne Cleland[1]
[1]*University of Strathclyde, Glasgow, United Kingdom*
[2]*Western Sydney University, Sydney, Australia*
[3]*Queen Margaret University, Edinburgh, United Kingdom*

Purpose This study compares lingual complexity in children with idiopathic Speech Sound Disorders (SSD) and typical development (TD) using the measure 'number of tongue inflections' (NINFL) (Preston et al., 2019) focusing on a wider range of consonants, SSD severity, and ages than previously studied. It has been hypothesised that children with speech sound disorders (SSD) and younger speakers are more likely to use undifferentiated tongue gestures due to immature speech motor control (Gibbon 1999). Kabakoff et al. (2021) showed that children with SSD have lower NINFL when producing /ɹ/ than TD children and younger children had higher complexity for /t/ than older children, contrary to expectations.

Method Children with idiopathic SSD (n=30, aged 5;0-12;11) and TD children (n=29, aged 5;8-12;11) were recorded with high-speed ultrasound and audio by a Speech and Language Therapist. The children with SSD produced 10 repetitions of /p, t, k, ɹ, l, w, s, θ, ʃ, j/ in an /aCa/ environment and TD children produced one repetition (there were no /j/ productions in the TD samples). Percent consonants correct (PCC) were measured by a Speech and Language Therapist as a measure of SSD severity. NINFL was measured using AAA software after fitting tongue splines at the maximal lingual gesture. /p/ was the baseline for comparing consonants. As a lingual consonant surrounded by /a/, /p/ functions like Kabakoff et al.'s baseline, and its spline was fitted using the same protocol as the rest of the consonants. Two mixed effects ordinal regression models were used for analysis. The first one used the predictors and interactions between Diagnosis, Age and Consonant to predict NINFL of TD and SSD children. The second one used the predictors and interactions between PCC, Age and Consonant to predict NINFL in SSD children. In both models Age and PCC were used as z-scores.

Results In the first model there were no significant interactions between age, consonant, and diagnosis. TD speakers' /ɹ/ had significantly higher NINFL than /p/, keeping age constant at baseline. In the second model, there was a significant interaction between PCC and age for /l/ and /ɹ/. With increase in PCC and age /l/'s NINFL decreased, relative to /p/, while /ɹ/'s NINFL increased. There were no other significant interactions, however, keeping age and PCC constant at baseline, /l/ and /ɹ/'s NINFL was higher than /p/'s and /θ/ and /k/'s NINFL was lower than /p/'s.

Conclusions The results suggest that an increase in age and a decrease in SSD severity, which are both theoretically linked to more mature motor control, can lead to either an increase or a decrease of NINFL. The lack of systematic difference between children with SSD or TD is also reminiscent of Kabakoff et al. (2021), who only observed differences for /ɹ/ compared to other consonants. The results suggest a complex relationship between lingual complexity and the development of speech motor control skills.



Example of a consonant /l/ produced with five inflections.

References

Gibbon, F. E. (1999). Undifferentiated Lingual Gestures in Children With Articulation/Phonological Disorders. *Journal of Speech, Language, and Hearing Research, 42*(2), 382–397. `https://doi.org/10.1044/jslhr.4202.382`

Kabakoff, H., Harel, D., Tiede, M., Whalen, D. H., & McAllister, T. (2021). Extending Ultrasound Tongue Shape Complexity Measures to Speech Development and Disorders. *Journal of Speech, Language, and Hearing Research, 64*(7), 2557–2574. `https://doi.org/10.1044/2021_JSLHR-20-00537`

Preston, J. L., McCabe, P., Tiede, M., & Whalen, D. H. (2019). Tongue shapes for rhotics in school-age children with and without residual speech errors. *Clinical Linguistics & Phonetics*, 33(4), 334–348. `https://doi.org/10.1080/02699206.2018.1517190`

CO-EXISTING DIFFICULTIES IN CHILDREN WITH PERSISTENT SPEECH SOUND DISORDERS (SSD) AND MOTOR SPEECH INVOLVEMENT

Åsa Mogren
Karolinska Institutet, Stockholm, Sweden

Speech sound disorders (SSD) is one of the most common neurodevelopmental disorders and often co-exist with other disorders in children, such as fine-, gross- and oral motor difficulties (Redle et al., 2015). Oral motor function is rarely assessed in clinic (Wikse Barrow et al. 2021) or in research (Braden et al. 2019). It is important to assess and describe orofacial function in children with SSD, as it may be relevant in differential diagnostics of speech disorders. The overall aim of this project was to investigate and describe orofacial function, speech characteristics, malocclusion, and other co-existing symptoms in children with SSD persisting after the age of six years.

The participants included 61 children with SSD aged 6:0-16:7 years (mean age 8:5), 14 girls and 47 boys, and 44 children with typical speech development (TSD) aged 6:0-12:2 years (mean age, 8:8), 19 girls and 25 boys. Speech was assessed in children with SSD by phonetic transcription of consonant and vowel production in a word naming test. Perceptual ratings of nasality were also performed. Differential diagnostics of SSD was made using the operationalised 12 Childhood Apraxia of Speech (CAS) features list by Iuzzini-Seigle & Murray (2017) and Shriberg's classification system (Shriberg et al., 2010). Parents completed a questionnaire including anamnestic information. The prevalence, type, and severity of malocclusion in children with SSD and TSD was registered. Assessment of orofacial function consisted of a screening test together with measurements of biteforce, chewing efficiency, jaw stability and sensory function. The results from the two groups were compared and related to malocclusions in the SSD group.

The results showed that all participants with SSD had impaired consonant production to a varying degree. Many participants also had impaired vowel production. Half of the participants were found to have deviant nasality. All children with SSD were assessed as having motor speech involvement (34 CAS, 23 Speech Motor Delay, 3 Articulation Impairment and 1 Developmental dysarthria). Children with SSD had worse performance on all orofacial function assessments than children with TSD, especially regarding assessments involving jaw stability and sensory function. In addition, children with SSD had a higher prevalence of malocclusions (61% vs 29%) and the malocclusions were also rated as more severe. In children with SSD, those with poorer orofacial function were at greater risk of malocclusion. General motor difficulties and other neurodevelopmental disorders were reported in children with SSD.

The findings from this project suggest that children with persistent SSD and motor speech involvement are at risk of orofacial dysfunction, malocclusions, general motor difficulties and other neurodevelopmental disorders, and should therefore be screened for co-occurring disorders. Children with SSD and poor orofacial function are at greater risk of malocclusion. An assessment of orofacial function is important when describing the characteristics of children with SSD, as it adds valuable information in differential diagnostics and also to ensure appropriate care.

References

Iuzzini-Seigel, J., Murray, E. (2017). Speech Assessment in Children With Childhood Apraxia of Speech. *Perspectives of the ASHA Special Interest Groups, 2*, 47-60. `https://doi.org/10.1044/persp2.SIG2.47`

Redle, E., Vannest, J., Maloney, T., Tsevat, R. K., Eikenberry, S., Lewis, B., . . . & Holland, S. K. (2015). Functional MRI evidence for fine motor praxis dysfunction in children with persistent speech disorders. *Brain Research, 1597*, 47-56. `https://doi.org/10.1016/j.brainres.2014.11.047`

Shriberg, L. D., Fourakis, M., Hall, S. D., Karlsson, H. B., Lohmeier, H. L., McSweeny, J. L., Wilson, D. L. (2010). Extensions to the Speech Disorders Classification System (SDCS). *Clinical Linguistics & Phonetics, 24* (10), 795-824.

Wikse Barrow, C., Körner, K., & Strömbergsson, S. (2021). A survey of Swedish speech-language pathologists' practices regarding assessment of speech sound disorders. *Logoped Phoniatr Vocol*, 1-12. `https://doi.org/10.1080/14015439.2021.1977383`

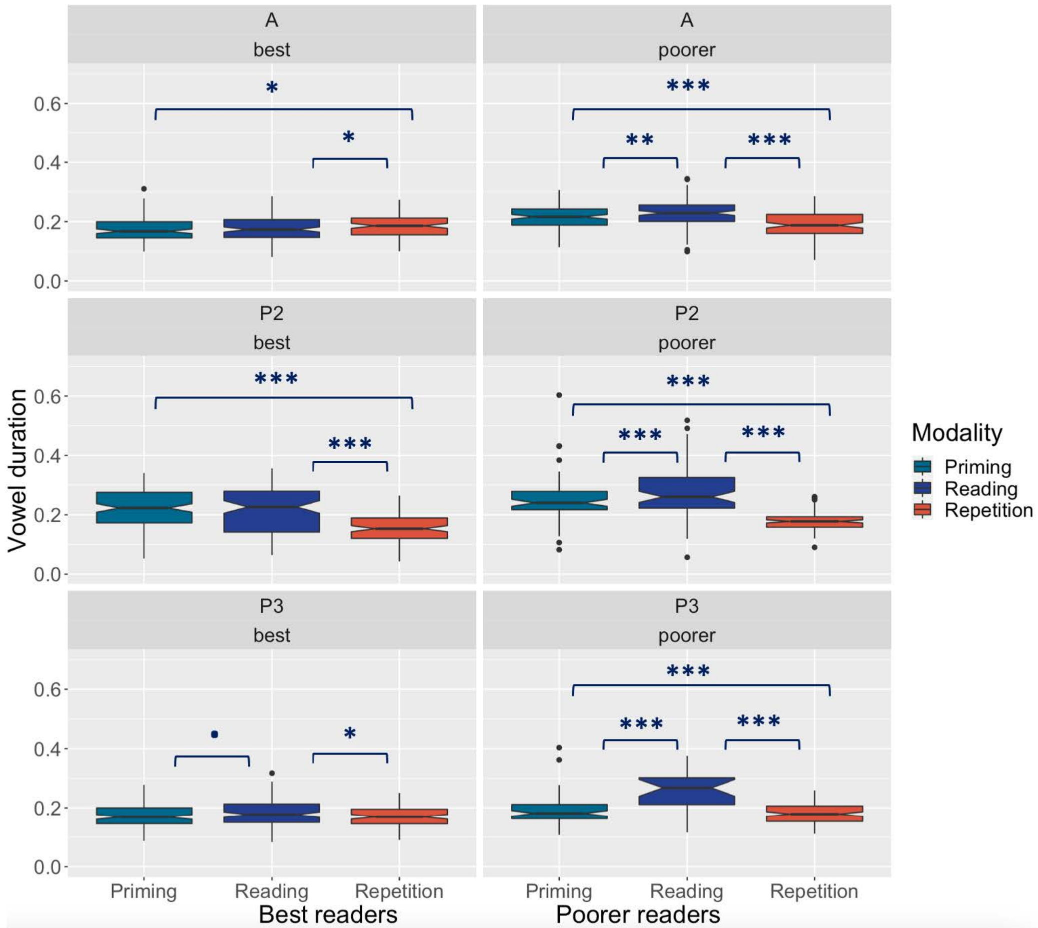
DIFFERENCES IN READING PROFICIENCY CORRELATE WITH VARIATIONS IN VOWEL DURATION AND DYNAMICS

Anisia Popescu[1], Elina Rubertus[1], Aude Noiray[2]
[1] *University of Potsdam, Potsdam, Germany*
[2] *Laboratoire Dynamique du Langage, Lyon, France*

The present study investigates the production of vowels in three speech modalities (repeated vs. read aloud vs. phonics style primed read speech) in German children (second and third graders) and adults with varying levels of reading proficiency. Recent studies have shown that speech production interacts with metalinguistic skills, such as phonological awareness (Noiray et al., 2019) and reading proficiency (Popescu & Noiray, 2021). Beginning readers of transparent alphabetical scripts use grapheme-phoneme correspondences to decode (non)words one grapheme at a time. With practice phonological decoding becomes more fluent and automatic. In this study we investigate to what extent reading proficiency (i.e., decoding capabilities) influences vowel duration and formant trajectories in different modalities (i.e., read vs. repeated speech). In addition to the classic reading and repetition modalities we add a third, inspired by synthetic phonics, a reading instruction method which uses phoneme decomposition and spelling out letters before putting them together to form a word (e.g., d-o-g – dog). This is a standard method used for transparent languages such as German (Landerl, 2000). We expect poorer readers to show greater variability in vowel duration and formant trajectories across modalities. In more proficient readers vowel production should not vary as a function of modality, indicating a more automated decoding process.

A total of 48 native German participants were recorded: 16 monolingual 2nd graders (age span: 7;03 (Y;MM) – 8;03), 16 monolingual 3rd graders (age span: 8;03 – 9;05) and 16 adults were recorded. Stimuli for all three modality blocks consisted of C_1VC_2ə pseudowords (C=/b,d,g/, V=/i,e,a,o,u/, C1≠C2) following the indefinite article /aɪnə/. The production task was split into three sequential sections corresponding to the three modalities: reading (stimuli appeared on screen), repetition (pre-recorded audio stimuli) and primed reading (C1 and V of each stimulus appeared separately before the word). Reading proficiency was evaluated using the standard SLRT-II test for German (Moll & Landerl, 2010). Measures of vowel duration and spectral dynamics (F2 trajectories) were carried out in Praat. Vowel duration was evaluated using linear mixed models with modality, proficiency, cohort, and their interaction as predictors. GAMM binary difference smooths (Wieling et al., 2018) were used to analyze differences in formant trajectories between modalities for different proficiency levels.

We present preliminary findings from 12 speakers, the two best and poorest readers from each cohort (adults – A, 2nd graders – P2, 3rd graders - P3). Results indicate that vowel duration is significantly longer in the reading compared to priming and shortest for the repetition modality in poorer readers, including adults. This is not the case for the best readers for whom vowels have comparable duration in the reading and primed-reading conditions (Figure 1). This suggests that poorer readers benefit from seeing the letters of the first syllable prior to decoding the printed word. These patterns are reflected in spectral dynamics – with significant differences in formant trajectories based on modality noted only for poorer readers. These findings indicate unskilled phoneme decoding is reflected in segmental duration and formant transitions.



Vowel duration (s) as a function of modality (repeated, read and primed-read speech), cohort (adults – A, 2nd graders – P2, 3rd graders P3) and reading proficiency (best vs. poorer readers). Data from 12 speakers.

References

Landerl, K. (2000) Influences of orthographic consistency and reading instruction on the development of nonword reading skills, *European Journal of Psychology of Education, 15*, 239–257

Moll, K., & Landerl, K. (2010). *SLRT-II: Lese-und Rechtschreibtest; Weiterentwicklung des Salzburger Lese-und Rechtschreibtests (SLRT)*. Huber.

Noiray, A., Popescu, A., Killmer, H., Rubertus, E., Krüger, S., & Hintermeier, L., (2019). Spoken language development and the challenge of skill integration, *Frontiers in Psychology, Language Sciences, 10*.

Popescu, A., & Noiray, A., (2021). Learning to Read Interacts with Children's Spoken Language Fluency, Language Learning and Development.

Wieling, M. (2018). Analyzing dynamic phonetic data using generalized additive mixed modeling: A tutorial focusing on articulatory differences between L1 and L2 speakers of English. *Journal of Phonetics, 70*, 86–116.

PRODUCTION OF VOCALIZED LATERALS IN WEST CENTRAL BAVARIAN - AN ARTICULATORY ANALYSIS OF PRIMARY SCHOOL CHILDREN

Katrin Wolfswinkler, Lia Saki Bucar Shigemori
Ludwig-Maximilians-University, Munich, Germany

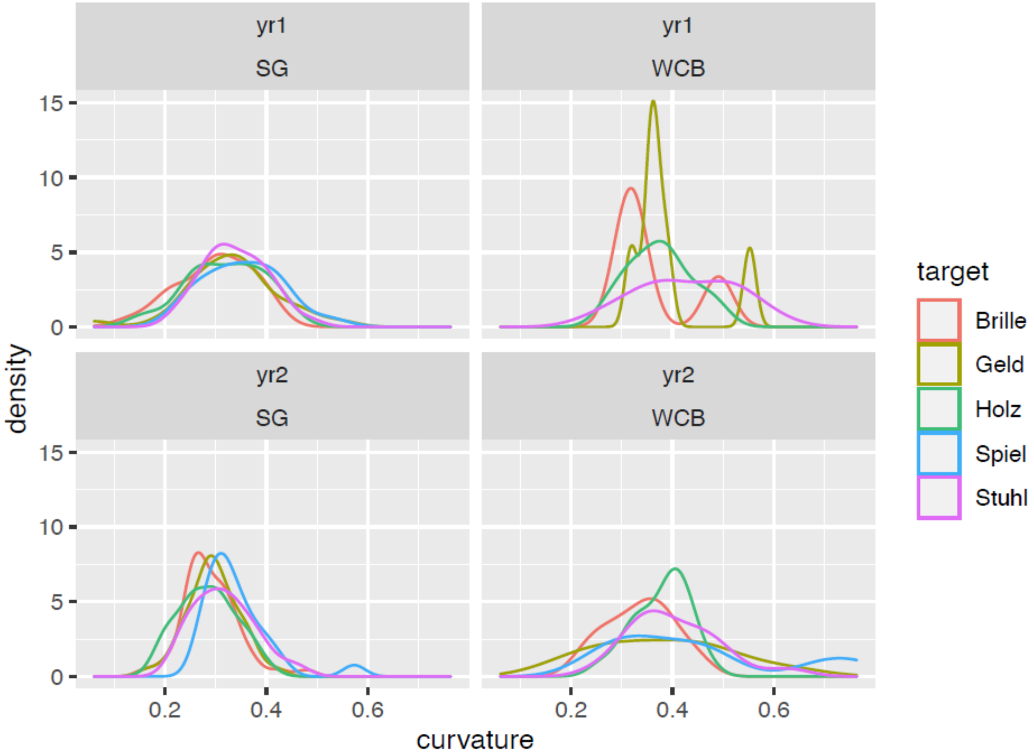
The present study is concerned with an articulatory investigation of child speakers of the West Central Bavarian (WCB) dialect, which is spoken in the south of Germany. The focus is on the production of diphthongs deriving from a historic sound change in which a post-vocalic lateral was vocalized towards a high vowel (Rein, 1974). From this process new diphthongs emerged that were until then not part of the WCB diphthong inventory (Bannert, 1976). In contrast, standard German (SG) still maintains the lateral.

We recorded children in whose parental background WCB was spoken. From the start of schooling children are exposed to SG to a much greater accent. That is, while WCB speaking children grow up with the vocalized form, the influence of SG - where the lateral is fully present - increases as soon as they enter primary school. It is assumed that the first years of schooling are particularly interesting to study shifts in spoken accent as the new peer-group increases in influence relative to the parental environment. Also, from an articulatory point of view the beginning of primary school is highly interesting. While changes in articulation are often linked to the maturation of the speech motor system, it has been shown that there is a relationship between reading proficiency and coarticulatory organization (Popescu & Noiray, 2021). This suggests that improving experience with grapheme to phoneme to speech motor correspondences that comes along with the acquisition of reading might also promote children's spoken language competence (Popescu & Noiray, 2021).

The focus of the present study was on children's abilities to differentiate the lateral (in SG pronunciation) and its vocalized counterpart (in WCB pronunciation) articulatory. A further goal was to investigate to what point children are able to form clear SG vs. WCB categories and whether this ability might improve as they grow older.

Ultrasound recordings of 14 WCB speaking primary school children, conducted in their first year of primary school and again one year later, were analyzed. They each produced four repetitions of five isolated words with V+/l/ in SG and a diphthong in WCB as target segments and of two words with either /i/ or /l/ in both SG and WCB as controls. The data was obtained via a picture naming task where the children had to name pictures appearing on a screen. The children switched rather unconsciously between the two varieties, which allowed us to investigate WCB and SG productions within the same children.

Quantifying the tongue shape in terms of values for tongue curvature and tongue curvature position (Ménard et al., 2012), the main difference between the SG lateral and the WCB vocalic production was found in tongue curvature. Figure 1 shows that there was a greater variability in the articulation of both SG /l/ and WCB /i/ in the children in their first as compared to their second year of primary school. These results suggest improving articulatory accuracy with increasing age, possibly stimulated by increasing orthographic experience and the concomitant enhanced awareness of phonemic units that starts to evolve between the investigated timepoints.



Distribution of tongue curvature values at the relative timepoint 0.8 (on a scale from 0 to 1, hence towards the end of either the V+/l/ segment or the diphthong in order to capture the shape of the /l/ in SG pronunciation or /i/ in WCB pronunciation), separated by recording timepoint (yr1 = 1st year in primary school; yr2 = 2nd year in primary school) and produced variety.

References

Bannert, R. (1976). *Mittelbairische Phonologie auf akustischer und perzeptorischer Grundlage.* Ph.D. dissertation, Lund University.

Ménard, L., Aubin, J., Thibeault, M., & Richard, G. (2012). Measuring tongue shapes and positions with ultrasound imaging: A validation experiment using an articulatory model. *Folia Phoniatrica et Logopaedica, 64*(2), 64-72.

Popescu, A., & Noiray, A. (2021). Learning to Read Interacts with Children's Spoken Language Fluency. *Language Learning and Development*, 1-20.

Rein, K. (1974). Die mittelbairische Liquiden-Vokalisierung. *Zeitschrift für Dialektologie und Linguistik, 41*(1), 21-37.

Poster Session III

THE CROOKED RELATIONSHIP BETWEEN TONGUE SHIFT AND F2 IN /U/-/Y/-TRANSITIONS

Rosa Franzke, Lia Saki Bučar Shigemori, Phil Hoole, Jonathan Harrington
Ludwig-Maximilians-University, Munich, Germany

In this study, we analyse speaker-dependent differences in the mapping between articulation and acoustics of /u/-/y/ transitions and consider possible systematic differences between female and male speakers.

It is known that acoustic variation between speakers can be systematic and linked to anatomical differences between biological females and males (Fitch & Giedd, 1999). Many studies have focused on mean fundamental frequencies and the distribution of phoneme targets in a F1-F2 space, but less is known about possible differences emerging in coarticulated segments or transitions within diphthongs.

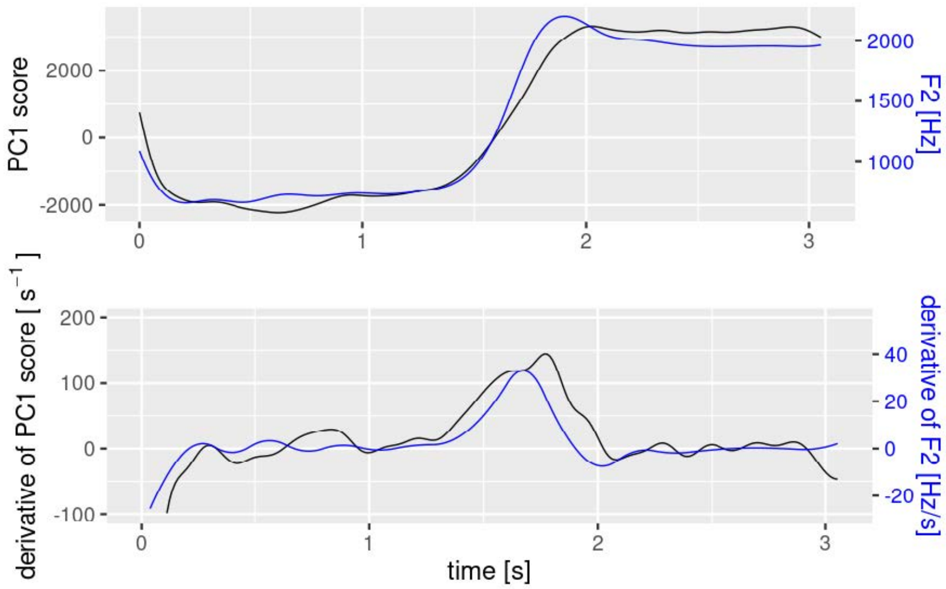
According to the quantal theory of speech (Stevens, 1989), the relationship between articulation and acoustics is also not linear within a speaker: In stable regions, a certain amount of articulatory movement results in less change in the acoustic output than in unstable regions, where a similar amount of articulatory movement results in a much greater change in the same acoustic parameter. Studies suggest that the region of the sub-glottal resonance frequencies can be viewed as one such unstable region (Stevens & Keyser, 2010). The second subglottal resonance which is robust within a speaker lies between 1350 and 1600 Hz. In the vicinity of the subglottal resonance frequency, a sudden jump in F2 occurs. In particular, this phenomenon has been observed in the spectrum during the transition of a diphthong with tongue movement from back to front in the vocal tract (Lulich et al., 2012). The region of subglottal resonance might overlap with F2 in fronted /u/ in alveolar consonant context especially for female speakers, magnifying the acoustic coarticulatory effect.

By analysing continuous articulatory transitions from /u/ to /y/ and vice versa we investigate the direct dependency between the articulatory movement and resulting acoustic output. We aim to detect regions of linear and non-linear acoustic-articulatory relationships and investigate to what extent speaker-dependent patterns can be observed.

19 German speakers (11 females) produced 4 to 6 repetitions of such /u/-/y/ and /y/-/u/continua. To quantify the articulatory data, captured by means of ultrasound, principal component analysis (PCA) on the time-varying raw image data was performed (Hoole & Pouplier, 2017). Separately for each speaker, the principal components (PCs) were computed based on all frames from each repetition of the continuum. Upon visual inspection, the PC best reflecting the /u/-/y/ contrast was chosen per speaker. The corresponding F2 was extracted from the recordings of the speech signal.

An example of the F2 and PC score trajectory of a /u/-/y/ production by one female speaker is shown in Figure 1. The derivatives of these trajectories were calculated to detect regions of greater movements.

Our findings will form part of our long-term goal of determining whether there is evidence that in fluent speech coarticulatory effects, both in timing and magnitude, can be expected to be affected by the differences in the acoustic-to-articulatory mapping between female and male speakers.



1: /u/-/y/ production by a female speaker: On the top the low-pass filtered PC1 score trajectory in black and the corresponding smoothed F2 trajectory in blue; on the bottom their derivatives. In this example, articulatory movement for the transition from /u/ to /y/ begins earlier than the F2 movement and ends later (in the region around time points 1.25s to 2s).

References

Brunner, J., Fuchs, S., & Perrier, P. (2009). On the relationship between palate shape and articulatory behavior. *The Journal of the Acoustical Society of America, 125*(6), 3936–3949. https://doi.org/10.1121/1.3125313

Fitch, W. T., & Giedd, J. (1999). Morphology and development of the human vocal tract: A study using magnetic resonance imaging. *The Journal of the Acoustical Society of America, 106*(3), 1511–1522. https://doi.org/10.1121/1.427148

Hoole, P., & Pouplier, M. (2017). Öhman returns: New horizons in the collection and analysis of imaging data in speech production research. *Computer Speech and Language, 45*, 253–277. https://doi.org/10.1016/j.csl.2017.03.002

Lulich, S. M., Morton, J. R., Arsikere, H., Sommers, M. S., Leung, G. K. F., & Alwan, A. (2012). Subglottal resonances of adult male and female native speakers of American English. *The Journal of the Acoustical Society of America, 132*(4), 2592–2602. https://doi.org/10.1121/1.4748582

Simpson, A. P. (2002). Gender-specific articulatory-acoustic relations in vowel sequences. *Journal of Phonetics, 30*(3), 417–435. https://doi.org/10.1006/jpho.2002.0171

Stevens, K. N. (1989). On the quantal nature of speech. *Journal of Phonetics, 17*, 3–45. https://doi.org/10.4064/-41-2-257-262

Stevens, K. N., & Keyser, S. J. (2010). Quantal theory, enhancement and overlap. *Journal of Phonetics, 38*(1), 10–19. https://doi.org/10.1016/j.wocn.2008.10.004

THE EFFECT OF MANDIBLE SIZE ON TEMPORAL PROPERTIES OF SPEECH

Daniel Friedrichs, Volker Dellwo
University of Zurich, Zurich, Switzerland

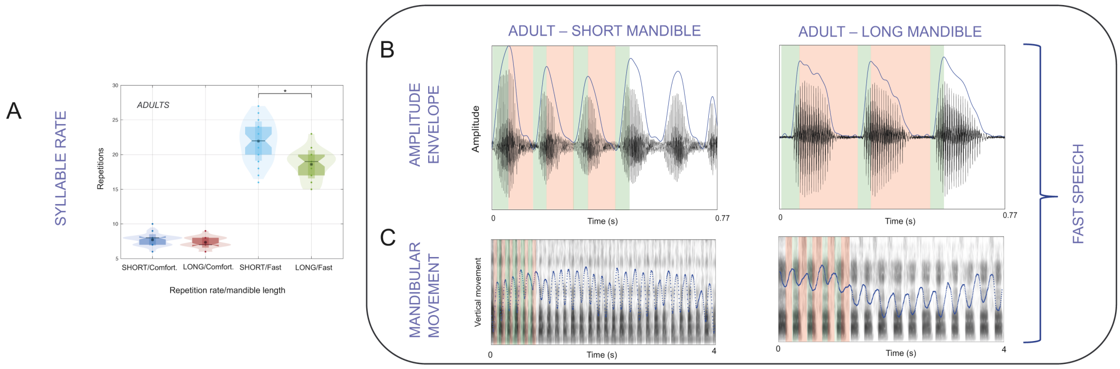
Prior research has shown that speech rate is influenced by various factors such as phonological complexity, social interaction, and a speaker's mental state. However, little is known about the effect of anatomical characteristics on the temporal properties of speech. This study assessed whether the length of the mandible affects the syllabic and supra-segmental timing of speech.

Two groups of speakers (mean age: 23.1, s.d.: 2.2; all students at the University of Zurich; no speech disorders) with either relatively short (N=7, five female, two male; mean length: 17cm, s.d.: 0.5cm) or long mandibles (N=7, two female, five male; mean length: 18.5cm, s.d.: 0.4cm) took part in an experiment, in which they were asked to produce five seconds of consonant-vowel (CV) repetitions at a comfortable and a maximum speech rate. The mandible length was calculated as the total distance between condylion, gonion, and menton (i.e., ramus and body of the lower jaw). CV combinations consisted of labial consonants and the back vowel /a/ (/ma/, /ba/, and /fa/) to ensure a high degree of vertical mandibular motion and mouth opening.

A repetition rate was calculated for a time window of four seconds starting from the onset of production. Furthermore, articulatory kinematics were investigated using electromagnetic articulography (EMA). Specifically, movements of the upper and lower incisors were recorded by a Carstens AG-501 system during the production of the CV repetitions. In addition, we examined the characteristics of the speech amplitude envelope, which was obtained by low pass filtering the full wave rectified signal at 30 Hz.

The results (see Figure 1) showed that speakers with shorter mandibles produced significantly more syllables in the maximally fast speech condition, while no differences between groups were found for the comfortable speech rate. Amplitude envelope characteristics revealed that the amplitude fall-time (i.e., the time from peak to minimum) was considerably longer in speakers with longer mandibles in fast speech, for which no differences were observed in the amplitude rise-time. Articulatory kinematics underline this as speakers from both groups required similar time to lower their mandibles (mouth opening motion), but those with longer mandibles needed considerably more time to raise their mandibles during mouth closure.

These findings support the hypothesis that anatomical properties influence speech timing. Further analyses are required to fully understand the different articulatory dynamics between speakers with varying jaw sizes and musculature. However, as the musculature involved in raising the mandible is grossly overpowered at a comfortable but not maximum speech rate, it seems likely that factors such as the mandible's resonant frequency have an effect on mandibular motion in fast speech. Such knowledge about anatomical properties encoded in the speech signal is relevant in explaining speaker-specific timing characteristics but may also play a role in understanding how segmental timing evolved in the world's languages.



(A) Boxplots showing the distribution of syllable repetitions for speakers with short and long mandibles at a comfortable and maximally fast speech rate. (B) Amplitude envelopes of CV segments produced by a speaker with a short (left) and long mandible in the fast speech condition. Amplitude rise and fall-time are highlighted in green and red color, respectively. (C) Graphical reconstruction of vertical mandible movements of the same speakers obtained by electromagnetic articulography (EMA) in the fast speech condition. To visualize differences in duration, up and down movements are highlighted in red and green, respectively.

BAYESIAN INFERENCE OF STATE FEEDBACK CONTROL MODEL PARAMETERS FOR PITCH PERTURBATION RESPONSES

Jessica L. Gaines[1,2], Kwang S. Kim[1], Ben Parrell[3], Vikram Ramanarayanan[1,4], Richard Ivry[2], Hardik Kothare[4], Srikantan Nagarajan[1], John Houde[1]
[1] *University of California, San Francisco, United States of America*
[2] *University of California, Berkeley, United States of America*
[3] *University of Wisconsin-Madison, Madison, Wisconsin, United States of Americas*
[4] *Modality.AI, San Francisco, California, United States of America*
[5] *University of California, Berkeley, Department of Psychology*

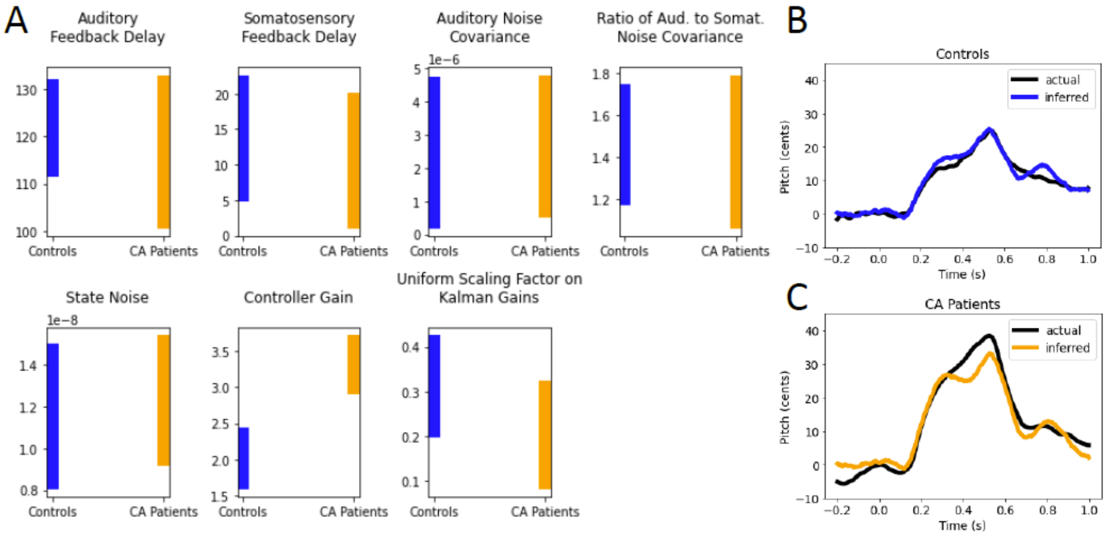
Introduction Computational modeling can be used to explore the neural mechanisms of speech motor control. Model parameters can be adjusted to test the effect of each parameter on the simulated response to a mid-trial pitch perturbation, which can then be compared with actual behavioral data. By finding the parameter values that best fit the model's output to the pitch perturbation response of a group of subjects, we can formulate mechanistic hypotheses about how that group of subjects differs, in terms of these parameters, from other subject groups (Gaines et al., 2020).

The present study compares behavioral data from previously published pitch perturbation studies to a state-feedback control (SFC) model of speech. In this model, motor commands are generated based on a neural estimate of vocal tract state. The estimated state and an efference copy of the commands are used to predict vocal tract state for the next time step, and thus the expected sensory feedback. The predictions are then compared with actual feedback to calculate error signals. These errors are scaled by a Kalman gain and used to update the predicted state, which is then used to generate the next set of commands. The values of tunable parameters like feedback noise covariance, feedback delays, state noise covariance, controller gain, and Kalman gain scaling affect the model's output of simulated vocal pitch over time in response to a mid-trial pitch perturbation (Houde et al., 2014).

Methods The Python package sbi (Tejaro-Cantero et al., 2020), a toolbox for simulation-based inference, was used to fit an SFC model of vocal pitch (Houde et al., 2014) to pitch perturbation response data from cerebellar ataxia patients and healthy controls (Houde et al., 2019). Using the sbi package, 50,000 simulations were run with different parameter sets randomly selected from a uniform prior distribution. The pitch time course outputs of these simulations were used to train a neural network that maps an output feature vector to the corresponding parameter set. From this neural network, a posterior distribution on each parameter was inferred and sampled to estimate the parameter set most likely to produce the observed data, along with a 95% confidence interval on each parameter value.

Results When we compare a 95% confidence interval for each fitted parameter between a healthy control group and a group of patients with cerebellar ataxia (Houde et al., 2019), we find that only controller gain is significantly different between controls and cerebellar ataxia patients (Fig. 1A). We can therefore hypothesize that cerebellar ataxia may be characterized by increased controller gain.

Future Work These procedures can be applied to other datasets to generate hypotheses about how additional groups of subjects may differ with regard to the parameters discussed here.



(A) A 95% confidence interval on each parameter fit for the pitch perturbation response of the control group (blue) and the group of cerebellar ataxia patients (orange). Only the controller gain parameter value is significantly different between the two groups. (B) A comparison between the actual (black) pitch perturbation response of the control group and the simulator output (blue) for the inferred parameters. (C) A comparison between the actual (black) pitch perturbation response of the patient group and the simulator output (orange) for the inferred parameters.

References

Gaines, J.L., Kothare, H., Raharjo, I., Ranasinghe, K., Ivry, R., Parrell, B., Agnew, Z., Nagarajan, S.S., & Houde, J.F. (2020). State feedback control model can account for differences in abnormal pitch perturbation responses in Alzheimer's disease and cerebellar ataxia [Conference presentation]. *12th International Seminar on Speech Production*, Providence, RI, United States. `https://issp2020.yale.edu/S02/gaines_02_14_153_poster.pdf`

Houde, J.F., Niziolek, C.A., Kort, N., Agnew, Z., & Nagarajan, S.S. (2014). Simulating a state feedback model of speaking. *10th International Seminar on Speech Production*, Cologne, Germany. `https://www.researchgate.net/publication/284486657_Simulating_a_state_feedback_model_of_speakisp`

Tejaro-Cantero, A., Boelts, J., Deistler, M., Lueckmann, J.-M., Durkan, C., Gonçalves, P.J., Greenberg, D.S., & Macke, J.H. (2020). sbi: A toolkit for simulation-based inference. *Journal of Open Source Software 5*(52), 2505. `https://doi.org/10.21105/joss.02505`

Houde, J.F., Gill, J.S., Agnew, Z., Kothare, H., Hickok, G., Parrell, B., Ivry, R.B., & Nagarajan, S.S. (2019). Abnormally increased vocal responses to pitch feedback perturbations in patients with cerebellar degeneration. *Journal of the Acoustical Society of America 145*(5), EL372. `https://doi.org/10.1121/1.5100910`

TOWARDS AN HD-SEMG MASK TO MEASURE OROFACIAL MUSCLE ACTIVITY DURING SPEECH PRODUCTION

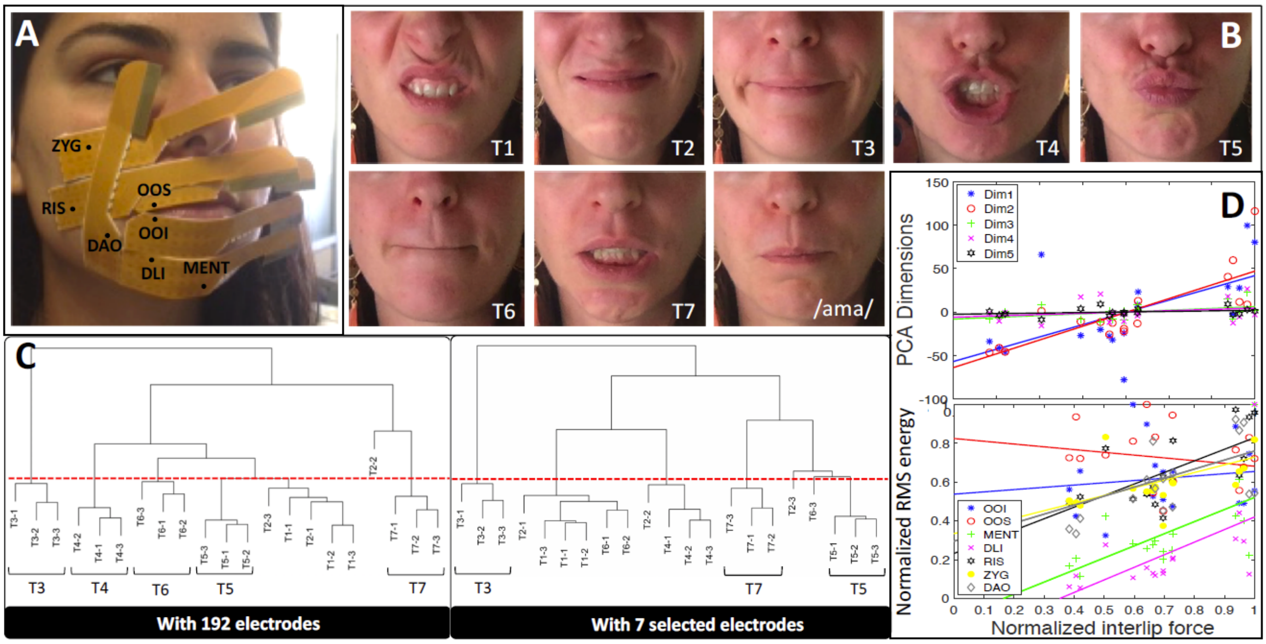
Maëva Garnier[1,2,3], Christophe Karam[1,2,3], Ladislas Nalborczyk[1,2], Alberto Botter[3], Sofiane Boudaoud[4]

[1]*University of Grenoble Alpes, CNRS, Grenoble INP, GIPSA-lab, Grenoble, France*
[2]*University of Aix Marseille, CNRS, LPC, Marseille, France*
[3]*Politecnico di Torino, Turin, Italy*
[4]*Université de Technologie de Compiègne, CNRS, Compiègne, France*

Objective The characterization of facial surface electromyographic (sEMG) activity has multiple interests for the understanding, modeling, recognition or synthesis of speech movements (Maeda & Honda, 1994; Huang et al., 2004; Stepp, 2012). The purpose of this work is to develop a face mask housing multiple sEMG electrodes to record HD-sEMG face activity (Lapatki et al., 2006; Wand et al., 2013), which would be easy to put on in experimental conditions and would adapt to varying face anatomies. We focused here on the feasibility of such an approach and on the characteristics of the detection system (i.e. number, location and density of the electrodes), compared to a simpler setup using isolated pairs of sEMG electrodes (Nalborczyk et al., 2020; Eskes et al., 2017; Schumann et al., 2021; Inzelberg et al., 2018; McClean et al., 1984; Janke & Diener, 2017).

Methodology A preliminary setup of 192 electrodes was developed, covering the whole cheek, chin and lip edge on one side of the face (see Figure 1A), made of 6 thin and flexible grids of 32 electrodes (LISiN), with an inter-electrode distance of 5mm. The setup was tested on a young female subject, who produced 3 repetitions of 7 basic orofacial movements (See Figure 1B) sustained during 30s, and 30 repetitions of the word /ama/ at comfortable level and speech rate. For that last speech condition, the interlip contact force was simultaneously recorded, using a pressure sensor (EPL-D11-25P from Measurement specialties). To evaluate the information gain brought by the electrode grids, seven out of the 192 electrodes were selected, corresponding to the typical location where isolated pairs of electrodes would have been placed (see Figure 1A).

Results The setup proved to be compatible with the production of "natural" orofacial movements, thanks to a layout of the grids that followed the "naso-labial fold" on the cheek and the "mental crease" on the chin. Only intense rounding and spreading of the lips were somehow restrained by the non-elasticity of the electrode grids. A simple PCA showed that more than 95% of the variance of the 192 channels could be explained by only 5 dimensions. However, these could not be easily interpreted in terms of group of adjacent electrodes or in terms of muscular synergies. Nevertheless, only 21.1% of electrode pairs provided highly correlated signals ($|R|>0.75$) while 55.5% of the electrode pairs were poorly correlated ($|R|<0.5$), suggesting a limited redundancy of the information provided by the electrode grids. This additional gain of information brought by the electrode grids resulted in a better discriminating power of basic orofacial movements. Thus, a standard Hierarchical Agglomerative Clustering (HAC) algorithm based on the average RMS energy of the 192 channels enabled the unsupervised classification of five movements (T3, T4, T5, T6, T7) out of the seven recorded here, whereas it was able to distinguish only 3 movement (T3, T5 and T7) when relying on seven well-selected electrodes (see Figure 1C). It also enabled a better prediction of the speech articulation effort, as measured by the interlip force during the words /ama/ (Adjusted $R\hat{2}$= 0.61 from the 5 PCA dimensions, against R2= 0.37 from the 7 well-selected electrodes) (see Figure 1D).



*A – Experimental setup with 6 grids of 32 sEMG electrodes. Seven out of the 192 electrodes are highlighted, corresponding to typical locations where isolated electrodes would have been placed to record the activity of specific facial muscles (OOS, OOI, DLI, DAO, MENT, ZYG, RIS). **B** – Representation of the seven basic orofacial movements investigated, in addition to productions of the word /ama/. **C** – Comparison of the unsupervised clustering (HAC) of these seven orofacial movements, based on the 192 electrodes or only on the 7 well-selected ones. **D** – Comparison of the prediction of the interlip force during the word /ama/ by the 5 principal components of the 192 RMS energies, or directly by the RMS energy of the 7 wellselected electrodes.*

References

Eskes M, van Alphen MJ, Balm AJ, Smeele LE, Brandsma D, van der Heijden F. (2017). Predicting 3D lip shapes using facial surface EMG. *PLoS One, 12*(4):e0175025.

Huang C, Chen C, Chung H. (2004). The Review of Applications and Measurements in Facial.

Inzelberg L, Rand D, Steinberg S, David-Pur M, Hanein Y. (2018). A wearable high-resolution facial electromyography for long term recordings in freely behaving humans. *Sci Rep., 8*(1):1–9.

Janke M, Diener L. (2017). EMG-to-speech: Direct generation of speech from facial electromyographic signals. *IEEEACM Trans Audio Speech Lang Process, 25*(12):2375–85.

Lapatki BG, Oostenveld R, Van Dijk JP, Jonas IE, Zwarts MJ, Stegeman DF. (2006). Topographical characteristics of motor units of the lower facial musculature revealed by means of high-density surface EMG. *J Neurophysiol., 95*(1):342–54.

Maeda S, Honda K. (1994). From EMG to formant patterns of vowels: the implication of vowel spaces.*Phonetica, 51*(13): 17–29.

McClean M, Goldsmith H, Cerf A. (1984). Lower-lip EMG and displacement during bilabial disfluencies in adult stutterers. *J Speech Lang Hear Res. 27*(3):342–9.

Nalborczyk L, Grandchamp R, Koster EH, Perrone-Bertolotti M, Loevenbruck H. (2020).

Can we decode phonetic features in inner speech using surface electromyography? *PloS One, 15*(5):e0233282.

Schumann NP, Bongers K, Scholle HC, Guntinas-Lichius O. (2021). Atlas of voluntary facial muscle activation: Visualization of surface electromyographic activities of facial muscles during mimic exercises. *PloS One, 16*(7):e0254932.

Stepp CE. (2012). Surface electromyography for speech and swallowing systems: measurement, analysis, and interpretation.

Wand M, Schulte C, Janke M, Schultz T. (2013). Grid-based Electromyographic Silent Speech Interface. *BIOSIGNALS*, p. 89–96. Available from: `http://csl.anthropomatik.kit.edu/downloads/BS13_WandSchulteJankeSchultz_GridBasedEMGSSI.pdf`

THE COORDINATION BETWEEN MOUTH OPENING-CLOSING RHYTHM AND INFORMATION IN SPEECH

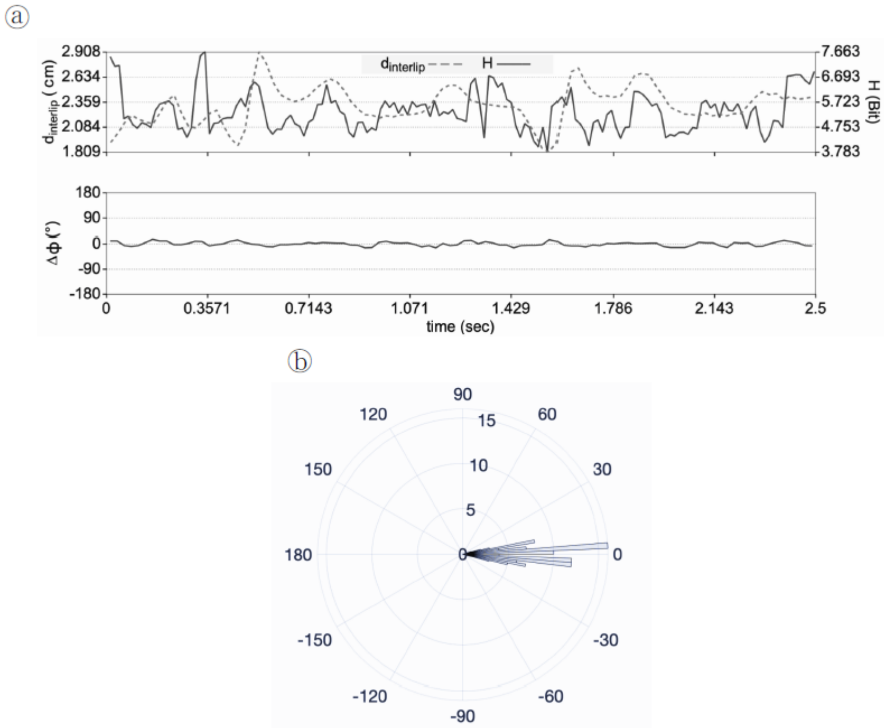
Lei He[1,2], Volker Dellow[1]
[1]*Department of Computational Linguistics, University of Zurich, Zurich, Switzerland*
[2]*Department of Phoniatrics and Speech Pathology, Clinic for Otorhinolaryngology, Head and Neck Surgery, University Hospital Zurich (USZ), Zurich, Switzerland*

The mouth opening-closing movements are believed to be the basis of speech rhythm which has its root in evolution according to the frame/content theory (MacNeilage 1998). The mouth opening-closing cycles tend to correlate with the sonority in speech, quantifiable in terms of the temporal envelopes, which are found to attract cortical entrainment at the slow theta- and delta-rates, playing an important role in comprehension in terms of sensory chunking, arguably (Morrill et al. 2012, Giraud and Poeppel 2012, Bucher et al. 2019, Poeppel and Assaneo 2020). These research findings point to the possibility that speech rhythm, in particular generated from the mouth opening-closing movements, functions as an information packaging mechanism. As a proof of concept, this study investigated the coordination between mouth opening-closing time series and the information fluctuation in speech (operationalized as spectral entropy at each windowed frame).

The mngu0 corpus (Richmond et al. 2011) was adopted. It contained 1354 recorded sentences (both audio and articulographic sensor histories; 1118 were selected) from one male speaker of British English. The mouth opening-closing time series was computed as the Euclidean distance between the two lips in the mid-sagittal plane. The spectral entropy was computed from each 16-msec frame (without overlap). The relative phase ($\Delta\ \Phi$) between the mouth opening-closing function (dinterlip) and the spectral entropy time series (H) were computed for each sentence (see Fig. 1A for illustration), and the distribution of $\Delta\ \Phi$ was revealed via the polar histogram (see Fig. 1B for illustration). From this particular example, it is evident that the mouth opening-closing time series and the spectral entropy time series are in phase to a great extent (MEAN = -0.33°, SE = 0.75°, MIN = -14.36°, MAX = 15.94°). To examine mutual similarities, the Bhattacharyya coefficient was calculated for all pairwise comparisons among all 1,118 $\Delta\ \Phi$ distributions (bin size = 2.5°); the median Bhattacharyya coefficient was 0.89 (1 - identical distributions; 0 - non-overlapping distributions).

The results point to the possibility that the opening-closing rhythm of the mouth functions as a packaging device of information in speech. In particular, the organized changes of the source-filter activities formed a train of spectral snapshots characterizable using the Shannon entropy. These spectral snapshots function as the essential building blocks of messages to be decoded by the receiver. As part of this source-filter mechanism, the mouth opening-closing rhythm packages these building blocks via in-phase coordinations to achieve comprehension.

This work was supported by the Swiss National Science Foundation (Grant No. PZ00P1 _193328 to LH), and the Forschungskredit of the University of Zurich (Grant No. FK-20-078 to LH).



A. An illustration of the calculation of the relative phase between mouth opening-closing time series ($d_{interlip}$) and the spectral entropy time series (H) using the first file in the corpus (mngu0_1_0001). B. The polar histogram showing the distribution of H for mngu0_1_0001.

References

Boucher, VJ; Gilbert, AC; Jemel, B (2019) The role of low-frequency neural oscillations in speech processing: Revising delta entrainment. *Journal of Cognitive Neuroscience 31*: 1205–1215.

Giraud, A-L; Poeppel, D (2012) Cortical oscillations and speech processing: Emerging computational principles and operations. *Nature Neuroscience 15*: 511–517.

MacNeilage, PF (1998) The frame/content theory of evolution of speech production. *Behavioral and Brain Sciences 21*: 499–546.

Morrill, RJ; Paukner, A; Ferrari, PF; Ghazanfar, AA (2012) Monkey lipsmacking develops like the human speech rhythm. *Developmental Science 15*: 557–568.

Poeppel, D; Assaneo, MF (2020) Speech rhythm and their neural foundations. Nature Reviews Neuroscience 21: 322–334.

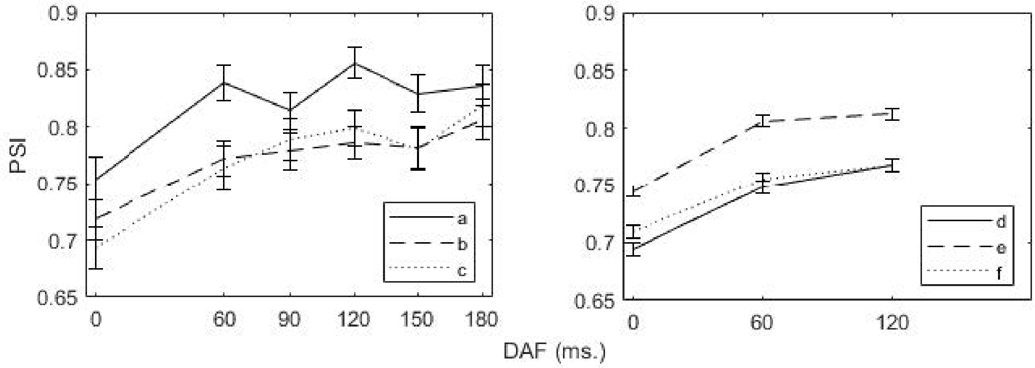
Richmond, K; Hoole, P; King, S (2011) Announcing the electromagnetic articulography (Day 1) subset of the mngu0 articulatory corpus. In *Proc. INTERSPEECH 2011* (Florence, Italy, 2011), pages 1505–1508.

UNVEILING THE COORDINATIVE ROLE OF PROSODY IN SPEECH PRODUCTION THROUGH MULTISCALE ANALYSIS OF THE MODULATORY ACTIVITY UNDERLYING SPEECH PRODUCTION

Leonardo Lancia, Jinyu Li, Cécile Fougeron
CNRS, Grenoble, France
Université Sorbonne Nouvelle, Paris, France

In the theoretical framework of coordination dynamics, emerging rhythmic patterns of behaviour may regulate the functioning of the sensorimotor system by stabilizing the interactions between the many degrees of freedom directly or indirectly controlled by speakers. In speech, the degree of perceivable rhythmicity is often related to the salience of prominent syllables or of supra-syllabic groupings (Dauer, 1987) and it has been related to the degree of coordination between syllabic and supra-syllabic patterns of activity (as extracted by band-pass filtering the acoustic amplitude signal to capture the relevant time scales) (Leong & Goswami, 2015; Lancia et al., 2019). The strength of this coordination varies significantly between so called stress-based languages whose phonology tends to refer to feet or to phonological words (as German or English) and syllable-based languages whose phonology tend to refer to syllables or to phrasal constructs (as French or Italian) (Lancia et al., 2019). Cross linguistic variation suggests that, by affecting the emerging rhythmic activity during speech production, the prosodic structure plays a role in organizing the moment-to-moment behaviour of the sensorimotor system. Further evidence comes from experiments showing that speakers' responses to perturbations of their auditory feedback are stronger in stressed or accented syllables (Li & Lancia, 2020; Oschkinat & Hoole, 2022).

A limit of studies on modulatory rhythmic activity is that these are mainly corpus-based, so that the effects of factors potentially affecting rhythm perception (e.g.: speech rate, segmental content) can mainly be controlled in post hoc analyses. We therefore conducted a Delayed Auditory Feedback (DAF) experiment aimed at controlling in a systematic fashion the rhythmic features of produced utterances. Indeed, to reduce the mismatch between the expected consequences of their motor commands and the perceived outcomes, speakers exposed to DAF lengthen the syllabic nuclei. However, the lengthening is stronger in prominent syllables Li & Lancia, 2020), which makes supra-syllabic activity more salient. In two different experiments, a total of 16 speakers of French (a syllable-based language) had to repeat several times three sentences with different orders and in which concurrently to DAF we manipulated either the likelihood that each sentence could produce a speech error, or the average complexity of the uttered syllables. Our hypothesis raised the prediction, confirmed by the data in the Figure (see supplemental material), that French speakers exposed to DAF would strengthen the coordination between syllabic and supra-syllabic activity (as speakers of a stress-based language). With such a stronger coordination the resulting complex rhythmic pattern gets simpler, nicely explaining the behaviour observed at the sub-syllabic level of activity. Indeed, the analysis of the amount of spectral change over time (computed as the total frame-to-frame squared distance of the first 12 MFCCs excluding the first) shows that, when feedback delays grow and local changes in speech rate become larger, the entropic content of the observed acoustic configurations decreases, suggesting a simplification of the underlying dynamics (Lancia et al., 2020). Altogether, these results show that structural relations between syllables, as encoded in the prosodic structure, contribute organizing speech dynamics across time scales of speech production in a way that provides both the stability and the flexibility that are required to rhythmic patterns of activity.



Average Phase Synchronization index (PSI; Rosenblum et al., 2000) between syllabic and suprasyllabic activity and standard error over auditory feedback delay (DAF in ms) grouped by. sentence[1] *Left panel sentences (uttered by 6 speakers, N= 1345 once speech errors removed): "Valmont voulait voir le vol" (a), "le norman hors norme mord" (b) , "maman et mami manient ma main" (c). Right panel sentences (uttered by 10 speakers, N= 8640 once speech errors removed): "Vivien vit le vin" (d), "Jaqueline jere e jour" (e), "Bradley brise le bras" (f). In both cases significance of the relation between DAF and PSI has been assessed via Mixed models with maximal random effects structure and as predictors the DAF level (coded as a categorical variable), the sentence, the sentence position (from first to third) and their double and triple interactions.*

References

Dauer R., M. (1987). Phonetic and phonological components of language rhythm. In Proceedings of the 11th International Congress of Phonetic Sciences (pp. 447-450).

Lancia, L., Krasovitsky, G., & Stuntebeck, F. (2019). Coordinative patterns underlying cross-linguistic rhythmic differences. Journal of Phonetics, 72, 66-80.

Lancia, L., Li, J., & Goldstein, L. (2020, December). Complexity patterns underlying speech production activity. Poster presented at ISSP 2020.

Leong, V., & Goswami, U. (2015). Acoustic-emergent phonology in the amplitude envelope of child-directed speech. PloS one, 10(12), e0144411.

Li, J., & Lancia, L. (2020). Speech production in response to multiple perturbations of auditory feedback.

Rosenblum, M., Tass, P., Kurths, J., Volkmann, J., Schnitzler, A., & Freund, H. J. (2000). Detection of phase locking from noisy data: application to magnetoencephalography. In Chaos In Brain? (pp. 34-51).

[1]Syllabic and supra-syllabic activities were estimated by band pass filtering the acoustic amplitude modulation signal with cut of frequencies adapted to the expected syllabic and accentual rates (as estimated from sentences duration in each recording). The PSI quantifies the variability of the generalized phase differences between the amplitude modulation signals on a time window of duration equal to 25 ms, sliding through the signal in steps of 5 ms. The PSI quantifies the variability of the generalized phase differences between syllabic and suprasyllabic amplitude modulations on a time window of duration equal to 25 ms, sliding through the signal in steps of 5 ms.

COMPARING TONGUE MOVEMENT VS. SHAPE REPRESENTATIONS FROM ULTRASOUND IMAGING OF /ɑ/ AND /ɹ/ ARTICULATORY STRATEGIES

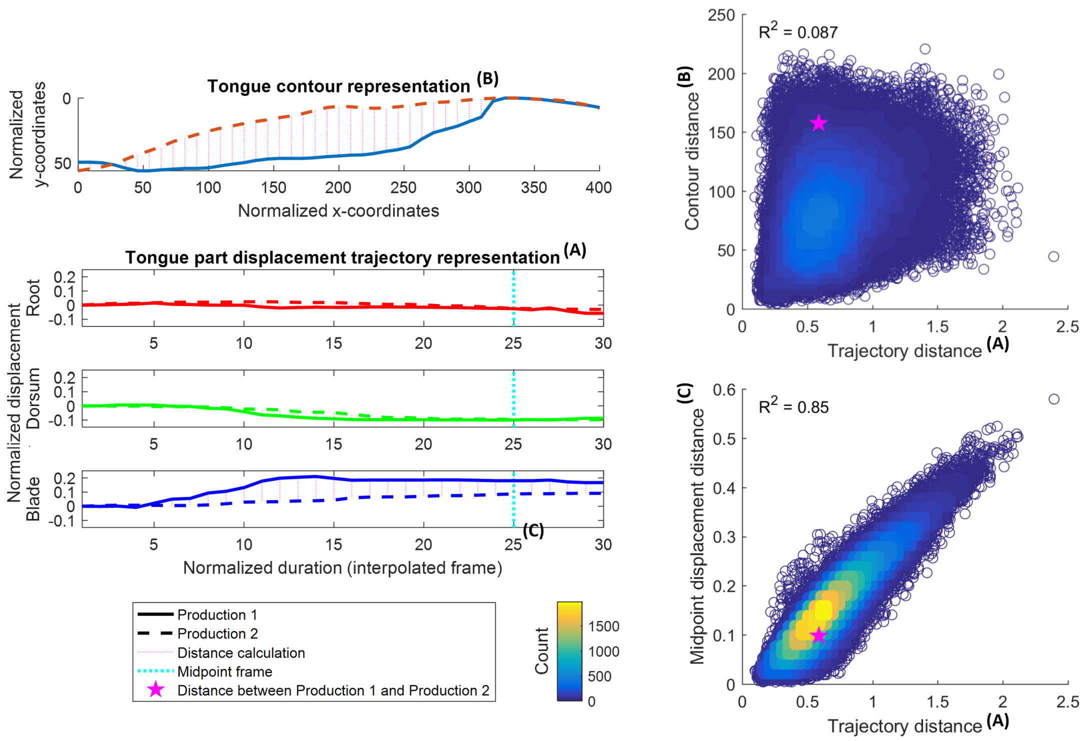
Sarah R. Li, Sarah Dugan, Sarah M. Schwab, Colin Annand, Jack Masterson, Suzanne Boyce, Michael A. Riley, T. Douglas Mast
University of Cincinnati, Cincinnati, United States

Introduction In speech sound production, multiple motor-control solutions are available to achieve a transition between articulated phonemes. Many studies have focused on static positionings of articulators (e.g., tongue shapes) as endpoints of these transitions. Studies of American English /ɹ/ have found that various tongue shapes yield perceptually equivalent outcomes (Tiede et al., 2004; Westbury et al., 1998). However, tongue shapes can vary based on adjacent vowels (Mielke et al., 2010), so that rhotic syllables can also be characterized by relative movements of articulators between phonemes (Dugan et al., 2019). The extent to which these relative movements are characterized by the tongue shapes comprising their endpoints is unknown. We used ultrasound imaging to explore relationships between representations of /ɑɹ/ articulation as tongue part displacement trajectories (time-dependent movements between /ɑ/ and /ɹ/) vs. tongue shapes (surface contours at /ɹ/).

Methods *Data representations.* Mid-sagittal ultrasound image sequences were recorded from 27 typical adult speakers of American English for 399 total productions of /ɑɹ/. TonguePART was used to automatically track the tongue surface contour and compute normalized displacement trajectories of the root, dorsum, and blade, representing time-dependent vocal-tract constrictions (Li et al., 2022). Tongue shapes during /ɹ/ production were characterized by normalized horizontal and vertical coordinates of the tongue contour at the midpoint-/ɹ/ frame. In addition to tongue-part trajectories and tongue shapes (Figure 1, left), two simplified representations were computed: (1) scores from principal component analysis (PCA) of trajectories and (2) relative displacements of root, dorsum, and blade from midpoint-/ɑ/ to midpoint-/ɹ/ frames.

Analyses To compare the articulatory information provided by each data representation, we evaluated Euclidean distances between each possible pair of productions. Distances for each data representation were cross-correlated (similar to assessing fits of distances in multi-dimensional scaling (Wang et al., 2013) and resulting R^2 values (Figure 1, right) were considered to measure similarity between the information about /ɹ/ strategies present within the respective representations. To compare identification of potential /ɹ/ articulation strategies, k-means clustering (k=2) was performed for each representation.

Results & Discussion R^2 values were high between the PCA-score (using three principal components) and trajectory representations (0.94) and between the relative-displacement and trajectory representations (0.85), showing that most of the variance in distances among trajectories was captured by these lower-dimensional representations. Between the contour and trajectory representations, R$\hat{2}$ was low (0.09), indicating that tongue shapes and trajectories captured different relationships among productions. Similar two-cluster solutions resulted for the trajectory, displacement, and PCA-score representations (Jaccard coefficients > 0.84). Two-cluster solutions for the trajectory and contour representations were dissimilar (Jaccard coefficients < 0.5), with contour clusters mainly separated by direction of concavity and trajectory clusters by blade displacement. While two-cluster solutions were optimal (based on gap evaluation) for the contour and trajectory representations, both suggested a continuum of strategies rather than distinct categories. Because displacement trajectories and tongue shapes capture different articulatory information, articulatory analyses may benefit from incorporating both representations.



Left panels display the contour (top) and displacement trajectory (bottom) representations for two selected productions. The relative-displacement representation comprises displacements of each tongue part at the acoustic midpoint frame of /ɹ/ (vertical cyan dashed lines). Euclidean distances between all recorded productions are plotted in the right panels, comparing contour vs. trajectory (top) and relative-displacement vs. trajectory (bottom) representations. Because of the large number of comparisons, a heatmap is used to indicate the count of pairs overlapping in these scatterplots.

References

Tiede, M. K., Boyce, S. E., Holland, C. K., & Choe, K. A. (2004). A new taxonomy of American English /r/ using MRI and ultrasound. *J. Acoust. Soc. Am., 115*(5), 2633–2634.

Westbury, J. R., Hashi, M., & Lindstrom, M. J. (1998). Differences among speakers in lingual articulation for American English /ɹ/. *Speech Commun., 26*(3), 203–226. https://doi.org/10.1016/S0167-6393(98)00058-2

Mielke, J., Baker, A., & Archangeli, D. (2010). Variability and homogeneity in American English /r/ allophony and /s/ retraction. In *Variation, Detail, and Representation.* (pp. 699–719). Mouton de Gruyter.

Dugan, S., Li, S. R., Masterson, J., Woeste, H., Mahalingam, N., Spencer, C., Mast, T. D., Riley, M. A., & Boyce, S. E. (2019). Tongue part movement trajectories for /r/ using ultrasound. *Perspect. ASHA Spec. Interest Groups, 4*(6), 1644–1652. https://doi.org/10.1044/2019_PERS-19-00064

Li, S. R., Dugan, S., Masterson, J., Hudepohl, H., Annand, C., Spencer, C., Seward, R., Riley, M. A., Boyce, S., & Mast, T. D. (2022). Classification of accurate and misarticulated /ɑɹ/ for ultrasound biofeedback using tongue part displacement trajectories. *Clin. Linguist. Phon.*, 1–27. https://doi.org/10.1080/02699206.2022.2039777

Wang, J., Green, J. R., Samal, A., & Yunusova, Y. (2013). Articulatory distinctiveness of vowels and consonants: a data-driven approach. *J. Speech Lang. Hear. Res., 56*(5), 1539–1551. https://doi.org/10.1044/1092-4388(2013/12-0030)

PEAK ACCELERATION DETERMINES SEGMENT BOUNDARY

Malin Svensson Lundmark[1,2]
[1]*Lund University, Lund, Sweden*
[2]*University of Southern Denmark, Odense, Denmark*

This study is about a one-to-one connection between articulation and acoustics, specifically, one that involves rapid articulatory movements resulting in acoustic changes we know as phonological segment boundaries (Fant & Lindblom, 1961; Zsiga, 1994). Rapid movements are associated with very high acceleration, peak acceleration, and the timing of peak acceleration determines the timing of a phonological segment boundary (Lundmark, 2020). When an articulator changes position, there is an instant speed change. For example, in a nasal bilabial stop, the lips slow down just before closing and accelerate just after opening. Likewise, tongue tip speed changes occur just before and after contact with the palate. The velocity changes in either direction of a movement are because of added energy resulting in very high acceleration (i.e., peak deceleration and peak acceleration).

Articulatory intervals based on peak acceleration and peak deceleration are hereafter referred to as posture intervals (Lundmark, 2020). A posture interval includes the steady-state and would in theory be equal to the duration of the constriction, i.e. the segment. Thus, we see a causal relationship between a posture interval of any given active articulator and the resulting segment.

As a first step of testing this proposed articulatory-acoustic relationship, correlation tests are performed. The two variables are 1) the posture interval of a consonantal articulator, and 2) the resulting segment duration. The expected result is a very strong relationship if, and only when, the posture interval involves the active articulator (here: either the lips or the tongue tip).

18 Southern Swedish speakers were recorded with an AG501 at the Lund University Humanities Lab. The full dataset includes 3000 disyllabic target words. The aggregated dataset consists of 875 tokens: CVC word onset with /a/ (long/short), and /n/, /l/ or /m/. Kinematic measures are on lip aperture and tongue tip peak deceleration and peak acceleration. The kinematic signal has been filtered and smoothed with low-pass filter, using the R function loess (span = 0.1). Acoustic measurements consist of consonant segment duration.

Results show that there is a very strong relationship ($>r=.90$) between the lip posture interval and the segment duration of [m], and between the tongue tip posture and [n]. However, this strong correlation is only present when the posture interval is measured on the crucial active articulator, and only for the nasal stops. A next step is to further test the causal relationship by investigating other manner and places of articulation. It is hoped that the approach introduced in this study will lead to a more comprehensive understanding of the acoustic-articulatory interface in speech.

References

Fant, G. & Lindblom, B. (1961). Studies of minimal speech sound units. *Speech Transmission Laboratory: Quarterly Progress Report, 2*, 1–11.

Svensson Lundmark, M. (2020). Articulation in time. *Some word-initial segments in Swedish.* Lund University.

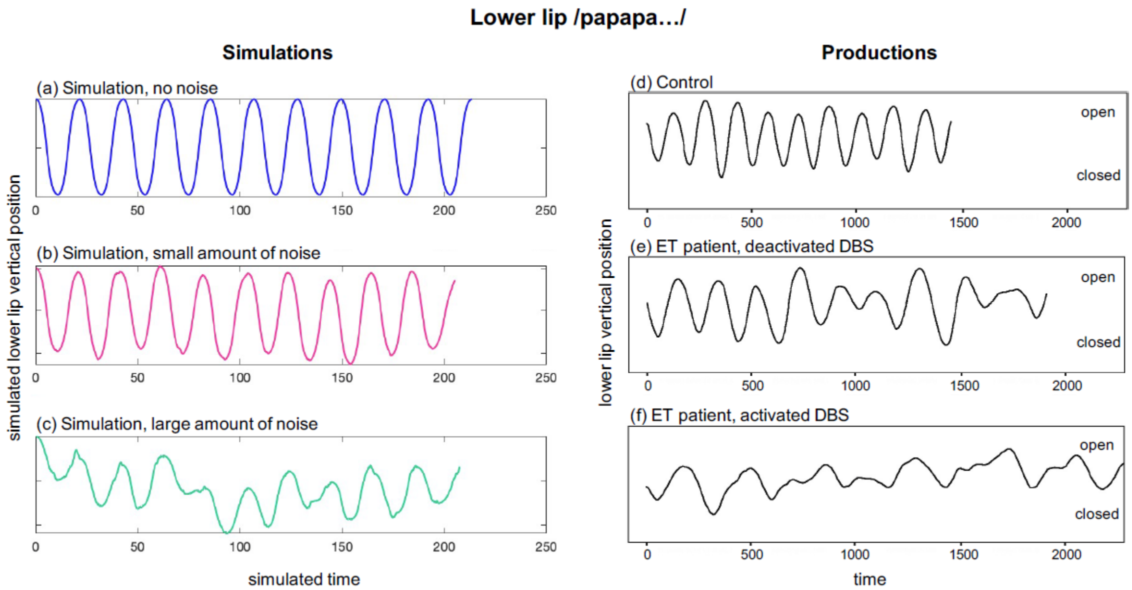
Zsiga, E. (1994). Acoustic evidence for gestural overlap in consonant sequences, *Journal of Phonetics, 22*, 121–140.

ANALYSIS AND MODELLING OF IMPAIRED SPEECH MOVEMENTS: CHALLENGES AND FUTURE DIRECTIONS

Doris Mücke[1], Antje Mefferd[3], Tabea Thies[1,2], Simon Roessig[1], Anne Hermes[4]
[1] *University of Cologne, Cologne, Germany*
[2] *University Hospital Cologne, Cologne, Germany*
[3] *Vanderbilt University Medical Center, Nashville, Tennessee, United States of America*
[4] *CNRS/Sorbonne Nouvelle, Paris, France*

A common goal of speech kinematic studies on dysarthric speech is the identification of the speech motor impairments that negatively impact speech function (e.g., precision, rate, intelligibility). However, to date, such research efforts are methodologically challenging. One major obstacle is the complex mapping between poorly defined (variable) articulatory movement patterns (i.e., the phonetic surface) and well defined (discrete) linguistic units (i.e., the phonological forms). Articulatory phonology (AP), based on Task dynamics (TD), is a framework that has succeeded in accounting for the mapping of continuous, variable phonetic signals to discrete linguistic forms (in many languages) (Saltzman & Munhall, 1989; Browman & Goldstein, 1992). However, it has been predominantly applied to the kinematic analysis of healthy speakers. Although AP/TD approaches can handle a high amount of context-dependent articulatory variability, their flexibility has been recently described as being insufficient to account for perturbated or highly noisy speech movement contours commonly seen in impaired speakers (Parrell & Lammert, 2019). Even in fast syllable repetition tasks, which are considered kinematically less complex than running speech tasks, AP/TD cannot fully account for highly deviant speech motor behaviors of impaired speakers. Therefore, the overall goal of this paper was to discuss the applicability and potential extension of a dynamical systems approach for kinematic studies of speakers with dysarthria. First, we built a case for such extensions by discussing current shortcomings of 1) dimension-reduction approaches, 2) spatial target definitions based on velocity profiles, and 3) the use of organ groups in lieu of individual articulators to account for synergies in a speech production model (Namasivayam et al., 2020).

As the second part of this paper, we explored the potential of introducing stochastic noise to AP/TD simulations. Based on pilot data, we sought to illustrate how the introduction of stochastic noise to the AP/TD model may provide a promising approach to modelling typical and deviant articulatory behavior, at least for simple oscillatory movements such as those occurring in syllable repetitions. Stochastic components can be integrated at various levels in the models, among these the gestural parametrization (e.g., stiffness) and the differential equation describing the gesture itself (as a stochastic differential equation). Figure 1a shows the output of an AP/TD simulation for /papapa…/ using damped anharmonic oscillators (Sorensen & Gafos, 2016). Figures 1b and 1c show simulation examples that apply a small amount and a large amount of noise, respectively. As can be seen, low noise levels resemble the contours of the typical speakers (Figure 1d) whereas high noise levels can resemble the deviant contours of the impaired speech movements (Figure 1e); however, perhaps not as well for increased severity levels as for mild to moderate severity levels (Figure 1f). In sum, the movement trajectories of speakers with speech motor impairments can be highly deviant from typical speakers to an extent that their characteristics may even be incompatible with current theories of speech motor control. Introducing stochastic noise to dynamical systems in speech motor control such as AP/TD holds promise to provide an account of modelling deviant articulatory movement patterns.



TD simulation (/papapa.../, simulated lower lip movements) with damped anharmonic oscillators is provided in (a), by applying a small amount of noise on the gestural parametrization level (stiffness and damping) and the differential equation in (b), and a large amount of noise in (c). The output is a concatenation of solutions that follow point-attractor dynamics. Lower lip trajectories during a fast syllable repetition task of /papapa.../, produced by a healthy speaker (d) and a speaker with Essential Tremor with deactivated (e) and activated (f) thalamic deep brain stimulation (DBS).

References

Browman, C. P., & Goldstein, L. (1992). Articulatory phonology: An overview. *Phonetica, 49*(3-4), 155-180. https://doi.org/10.1159/000261913

Namasivayam, A. K., Coleman D., O'Dwyer A. & van Lieshout, P. (2020). Speech sound disorders in children: an articulatory phonology perspective. *Frontiers in Psychology, 10*:2998, https://doi.org/10.3389/fpsyg.2019.02998

Parrell, B., & Lammert, A. C. (2019). Bridging dynamical systems and optimal trajectory approaches to speech motor control with dynamic movement. *Frontiers in Psychology, 10*:2251. https://doi.org/10.3389/fpsyg.2019.02251

Saltzman, E., and Munhall, K. G. (1989). A dynamical approach to gestural patterning in speech production. *Ecological Psychology, 1*, 333–382. https://doi.org/10.1207/s15326969eco0104_2

Sorensen, T. and A. Gafos. (2016). The gesture as an autonomous nonlinear dynamical system. *Ecological Psychology 28*(4), 188-215. https://doi.org/10.1080/10407413.2016.1230368

SUPRA-LARYNGEAL ARTICULATION UNDER VOCAL EFFORT VARIATION

Lena Pagel[1], Simon Roessig[1,2], Doris Mücke[1]
[1]*University of Cologne, Cologne, Germany*
[2]*Cornell University, Ithaca, United States of America*

Speakers vary the amount of effort they put into speech production according to communicative demands (Lindblom, 1990). Vocal effort manifests itself in various articulatory modulations and varies on the global level of the utterance (e.g. in loud speech compared to habitual speech) and on the local level of syllables (e.g. to mark prosodic prominence) (Cho, 2004; Mefferd & Green, 2010; Roessig & Mücke, 2019). This study investigates the interaction of the two levels of vocal effort: If individuals speak loudly with a globally high vocal effort, can they still encode prominence relations as local effort variations?

An experiment with 20 German speakers was carried out using 3D Electromagnetic Articulography (AG501) to capture lip and tongue kinematics. Speakers were engaged in an interactive question-answer task. They produced the syllables /bi/ and /mi/ in target words in two utterance positions (initial or medial). Two focus conditions related to local variation of vocal effort were elicited through a virtual avatar's questions: In one condition, both the initial and the medial word were moderately prominent (broad focus); in the other condition, the medial word was emphasised (corrective focus) and the initial word was attenuated (background). To test for global variation of vocal effort, all utterances were produced in habitual and loud speech. We investigated articulatory extremum positions of the lips (as the Euclidean distance of the lips) and tongue body (in the vertical and horizontal movement dimension) during the target vowel /i/.

The preliminary results show that lip opening is modulated as a function of local vocal effort variation: The aperture increases when prominence increases, i.e. from broad to corrective focus (corresponding to a sonority expansion), and decreases when prominence decreases, i.e. from broad focus to background (corresponding to a sonority reduction). In loud speech, the lip aperture is overall greater than in habitual speech, revealing a global increase in vocal effort, which yet does not hinder the encoding of prominence relations. In tongue kinematics as well, prominence is encoded through similar strategies in both speaking styles: When prominence increases, the high front vowel /i/ is produced with a fronted target (corresponding to a hyperarticulation), when prominence decreases, the target is lowered and retracted (corresponding to a hypoarticulation). This can be observed in both speaking styles, even though the overall tongue position is fronted and lowered in loud speech as compared to habitual speech.

In summary, our findings show that prominence relations are systematically encoded in supra-laryngeal articulation as a function of local vocal effort variation. At the same time as one entity is strengthened, another entity is weakened, thereby potentially increasing the perceptual prominence of the strengthened entity. Crucially, this can not only be observed in habitual but also in loud speech, where vocal effort is globally increased. Our results underline the flexibility of the prosodic system in adapting to communicative demands. These adaptions are best conceived of as the maintenance of parameter relations in a dynamical model of speech production rather than fixed, hard-coded parameter sets.

References

Cho, T. (2004). Prosodically conditioned strengthening and vowel-to-vowel coarticulation in English. *Journal of Phonetics, 32*, 141–176. https://doi.org/10.1016/S0095-4470(03)00043-3.

Lindblom, B. (1990). Explaining Phonetic Variation: A Sketch of the H&H Theory. In W. J. Hardcastle & A. Marchal (Eds.): *Speech Production and Speech Modelling* (pp. 403–439)

Mefferd, A. S., & Green, J. R. (2010). Articulatory-to-acoustic relations in response to speaking rate and loudness manipulations. *Journal of Speech, Language, and Hearing Research, 53*(5). https://10.1044/1092-4388(2010/09-0083)

Roessig, S., & Mücke, D. (2019). Modeling Dimensions of Prosodic Prominence. *Front. Commun., 4*, https://doi.org/10.3389/fcomm.2019.00044

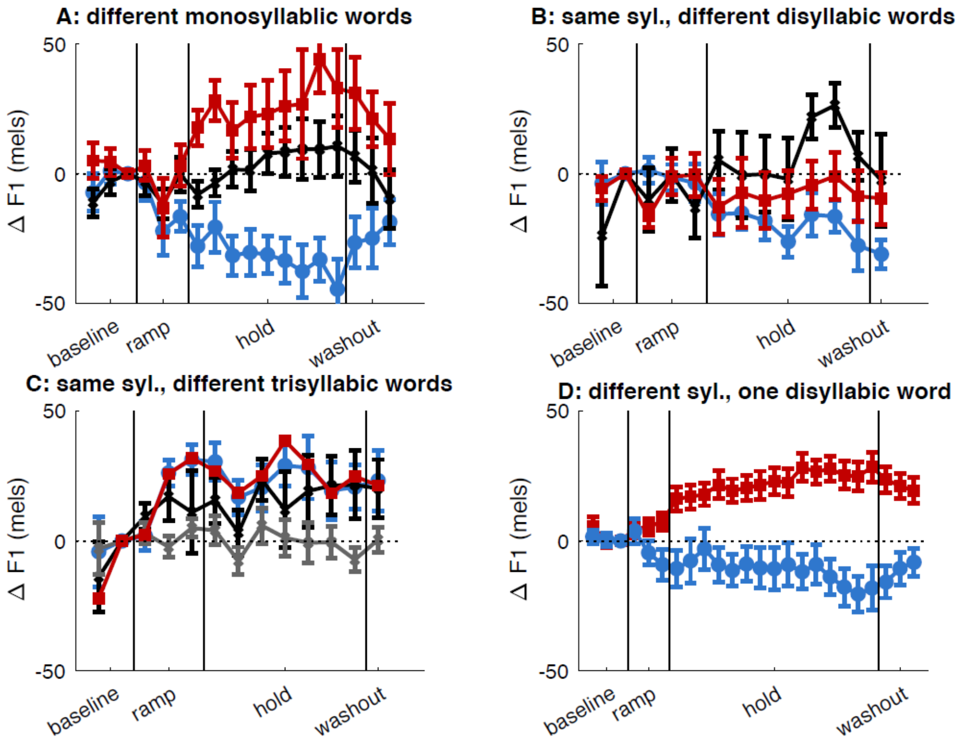
ASSESSING THE SCOPE OF SPEECH MOTOR PLANNING WITH SENSORIMOTOR ADAPTATION

Benjamin Parrell*, Caroline A. Niziolek*
University of Wisconsin-Madison, Madison, United States of America
**equal contribution*

Speech production is classically conceptualized as a series of hierarchical processing stages ending with the selection of low-level linguistic units which are then read out into articulatory movements by the motor planning system. However, such models are insufficient to explain growing evidence that speech motor planning encompasses more than phonemes, syllables, or articulatory gestures. The current work addresses the scope of speech motor planning by building on recent studies in upper limb motor control that have shown motor planning is critical to differentiate movements for learning. For example, participants are able to simultaneously learn to adapt their reaching movements to two opposing force fields if those fields are linked to planned follow-through movements to different targets (Sheahan et al., 2016), but not if they are linked to an arbitrary visual cue (Woolley et al., 2007). These results suggest that motor plans for complex actions form a cohesive unit that can scaffold learning. In speech, participants can simultaneously adapt their production of a single vowel in different ways in response to auditory perturbations applied to different monosyllabic words (Rochet-Capellan & Ostry, 2011), suggesting the existence of multiple motor plans for the same vowel, differentiated by surrounding context. However, the precise scope of motor planning in speech remains unclear, as current evidence cannot distinguish whether adaptive behavior is tied to individual syllables or larger word- or utterance-level units.

Here, we report the results of four experiments designed to examine the scope of planning in speech using differential sensorimotor adaptation. In all studies, the first vowel formant (F1) was shifted in opposing directions (up/down) on different stimuli. First, we replicated the results of (Rochet-Capellan & Ostry, 2011), confirming that speakers are able to adapt to opposing F1 perturbations applied to different monosyllabic words sharing the same vowel (difference of 77±20 mels, Fig. 1A, n = 15). Learning was present but attenuated when perturbations were applied to the same syllable in different disyllabic words (*sever* vs *seven*, 19±11 mels, Fig 1B, n = 7, anticipated final n = 20), and essentially eliminated when perturbations were applied to the same syllable in different trisyllabic words (*pedicure* vs *pedigree*, 5±5 mels, Fig 1C, n = 8, anticipated final n = 20). Conversely, robust learning was observed when opposing perturbations were applied to different syllables within the same disyllabic word (*bedhead*, 46 ± 9 mels, Fig 1D, n = 20).

These results are largely consistent with theories that posit speech motor planning relies on the selection and sequencing of syllables. However, the fact that learning was observed when the same syllable was perturbed in different disyllabic words suggests that larger, word-level plans may also be recruited in some cases. The lack of learning in trisyllabic words may suggest that such word-level plans only exist for more frequent words, consistent with suggestions of the DIVA model (Guenther, 2016). It should be noted that these conclusions must be drawn tentatively given the ongoing data collection for two of these experiments.



For all experiments, stimuli with a downward perturbation applied to F1 (-125 mels) are shown in red, stimuli with an upward perturbation applied to F1 (+125 mels) are shown in blue, and control stimuli with no perturbation are shown in black and gray. ***A:*** *Opposing formant perturbations applied to different monosyllabic words results in robust motor learning to counter the applied perturbation, consistent with Rochet-Capellan & Ostry (2011).* ***B:*** *When the same perturbations were applied to the same initial syllable in different disyllabic words ("sever" and "seven"), learning was substantially attenuated.* ***C:*** *learning was essentially eliminated with perturbations were applied to the same initial syllable in different trisyllabic words ("pedigree" and "pedicure").* ***D:*** *Conversely, robust learning was observed when opposing perturbations were differentially applied to the first and second syllables in the disyllabic word "bedhead".*

References

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

Rochet-Capellan, A. and D. J. Ostry. 2011. Simultaneous acquisition of multiple auditory-motor transformations in speech. *J Neurosci, vol. 31*, no. 7, pp. 2657–62, `https://doi.org/10.1523/JNEUROSCI.6020-10.2011`

Sheahan, H. R., D. W. Franklin, and D. M. Wolpert (2016). Motor Planning, Not Execution, Separates Motor Memories. *Neuron, vol. 92*, no. 4, pp. 773–779, Nov. 2016, `https://doi.org/10.1016/j.neuron.2016.10.017`

Woolley, D. G., J. R. Tresilian, R. G. Carson, and S. Riek (2007). Dual adaptation to two opposing visuomotor rotations when each is associated with different regions of workspace. *Exp. Brain Res., vol. 179*, no. 2, pp. 155–165. `https://doi.org/10.1007/s00221-006-0778-y`

LATERALIZATION IN ONSET AND CODA ENGLISH LATERAL CONSONANTS: A MULTISLICE RTMRI ANALYSIS

Anisia Popescu[1,2], Mairym Llorens Monteserin[2], Louis Goldstein[2], Shrikanth Narayanan[2]
[1]*University of Potsdam, Potsdam, Germany*
[2]*University of Southern California, Los Angeles, California, United States of America*

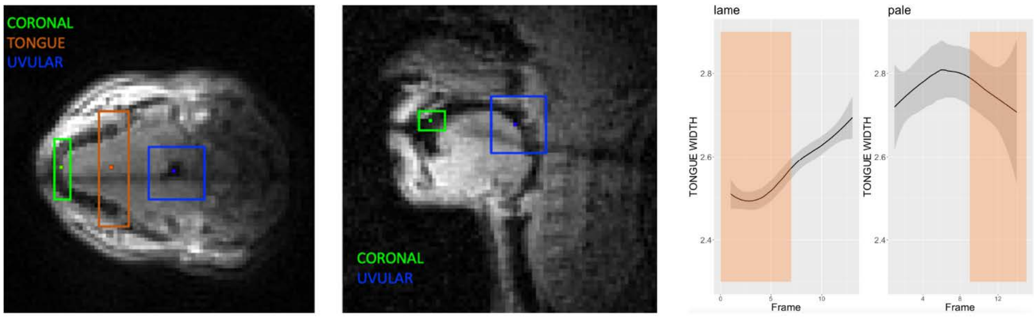
Lateral consonants, as the name indicates, are characterized by the presence of lateral channels that allow airflow paths along the sides of the tongue (Stevens, 1998). A possible cause for the creation of side channels is the narrowing of the front part of the tongue, by retracting the sides of the tongue from the molars. Another possibility is the tilting of the tongue to one side of the oral cavity (Katz et al., 2017; Ying et al., 2021). English laterals are produced with two antagonistic gestures: a tongue tip advancement creating a coronal constriction and a tongue dorsum retraction creating a uvular constriction (Sproat & Fujimura, 1992; Browman & Goldstein, 1995). Because of the tongue's incompressibility (Fujimura & Kakita, 1979) the opposite direction constrictions lengthen the tongue, possibly causing the tongue to narrow (as hypothesized by (Browman & Goldstein, 1995). The present study investigates whether the tongue narrowing correlates over time with coronal and uvular constrictions in English laterals and with the frequency of F3, using multi-slice rtMRI analysis.

Three participants were recorded producing CVC words involving a lateral consonant (onset or coda position), a variety of vocalic contexts and a labial consonant, in either phrasal, phonological or no boundary prosodic contexts.

Two rtMRI slices were used for the analysis – a commonly used mid-sagittal slice and an additional oblique slice, which was defined by a line along the mid-sagittal plane between two anatomically references anchor points: the inferior surface of the upper lip and the intersection between the spinal cord and the C2-C3 intervertebral disk (Llorens et al., 2017). Tongue width, coronal and uvular constriction degrees (CD) were estimated using a region of interest (ROI) analysis (Blaylock, 2021) on both slices. Frames of interest (i.e., lateral-vowel and vowel-lateral sequences) were determined based on acoustic landmarks extracted manually in Praat. Individual measures of coronal and uvular CD, as well as their product (CORONAL x UVULAR), together with syllable position, prosodic boundary and vowel position were used as independent variables in a multiple regression analysis to predict the value of tongue width. A second multiple regression linked F3 value to tongue width, syllable position, prosodic boundary and vowel context.

Preliminary results from one participant show tongue width is negatively correlated with the combined measure of coronal and uvular constrictions ($p < 0.\ 0001$). The tongue is narrower in coda vs. onset laterals ($p < 0.0001$) and in back vs. front vowels ($p < 0.0001$). Acoustic-articulatory relations show F3 values are negatively correlated with tongue width ($p < 0.0001$). Coda laterals, which show higher degree of narrowing exhibit higher F3.

Results point to a systematic relation between tongue width and tongue tip and tongue dorsum gestures of the lateral. The coordination of both coronal and uvular gestures results in narrowing of the tongue. Acoustically, higher degrees of narrowing result in higher values of F3.



Regions of interest (ROI) for the oblique and mid-sagittal MRI slices (left); Tongue width values over time in words lame and pale (right) indicating tongue narrowing during the laterals (in orange)

References

Blaylock, R.; 2021, *VocalTract ROI Toolbox*. `https://github.com/reedblaylock/VocalTract-ROI-Toolbox`

Browman, C. P., & Goldstein, L. (1995). Gestural syllable position effects in American English. *Producing Speech: Contemporary Issues*, 19–33.

Fujimura, 0. & Kakita, Y. 1979. Remarks on quantitative description of the lingual articulation, In *Frontiers of speech communication research* (S. Ohman & B. Lindblom, editors). London: Academic Press.

Katz, W., F.; Mehta, S.; Wood, M.; & Wang, J.; 2017. Using electromagnetic articulography with a tongue lateral sensor to discriminate manner of articulation, *JASA, 141*(1).

Llorens Monteserin, M.; Byrd, D., Goldstein, L. & Narayanan, S.; 2017. Indexing tongue profile narrowing for English using 3D volumetric MR imaging. *174th Meeting of the ASA.*

Sproat, R., & Fujimura, O. (1993). Allophonic variation in English/l/and its implications for phonetic implementation. *Journal of Phonetics, 21*(3), 291–311.

Stevens, K., N. 1998. *Acoustic phonetics*, The MIT Press, Cambridge, Massachusetts.

Ying, J.; Shaw, J., A.; Carignan, C., Proctor, M.; Derrick, D., & Best, C., T.; 2021. Evidence for active control of tongue lateralization in Australian English /l/, *Journal of Phonetics, 86.*

AGE-RELATED CHANGES ON TONGUE BODY MOVEMENTS

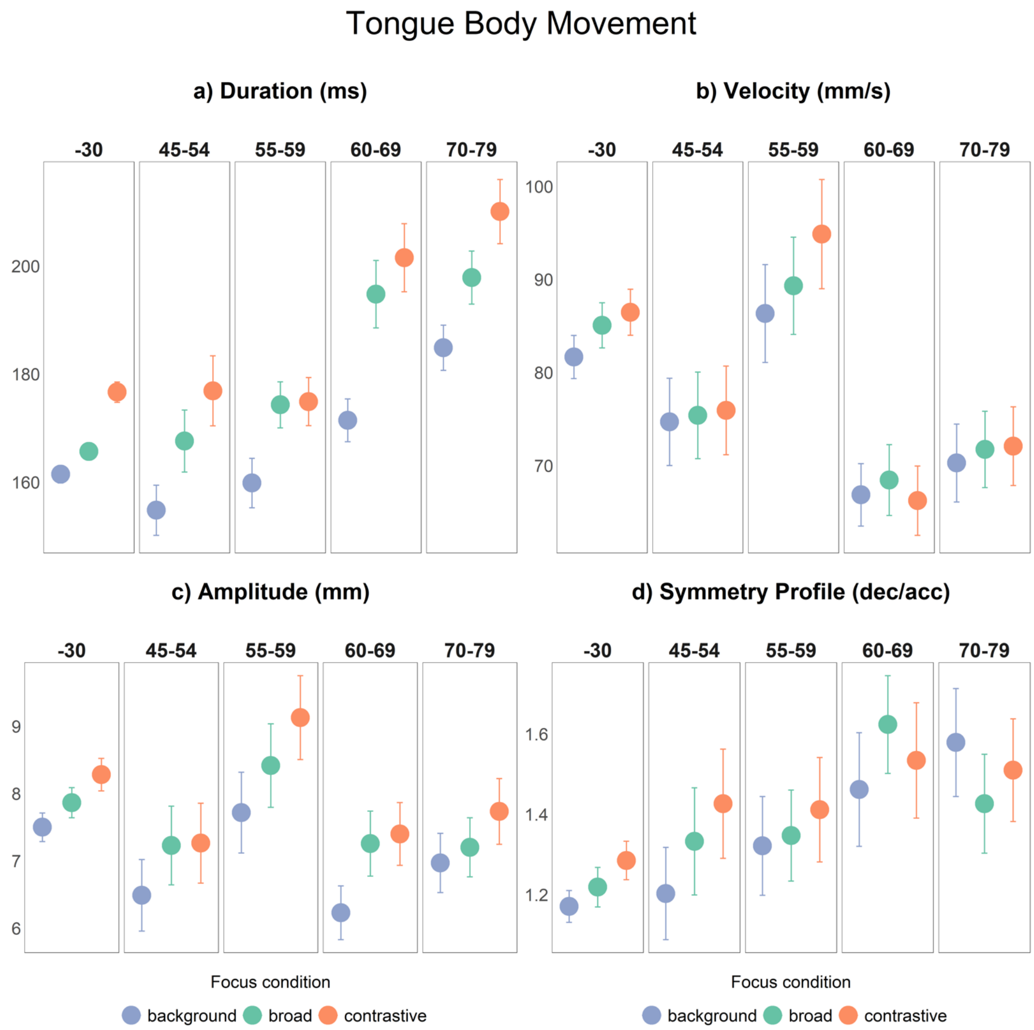
Janine Schreen[1,3], Tabea Thies[1,3], Anne Hermes[2], Doris Mücke[2]
[1] *University of Cologne, Cologne, Germany*
[2] *CNRS/Sorbonne-Nouvelle, Paris, France*
[3] *University Hospital Cologne, Cologne, Germany*

Introduction Acoustic studies have shown, that effects of aging can impact the level of speech motor control including a slower speech tempo (Amerman & Parnell, 1992) and a reduced vowel space (Thies et al., 2022), especially with an increased instability of speech patterns above the age of 60 years (Karlsson & Hartelius, 2021; D'Alessandro & Fougeron, 2021). Furthermore, a gradual, non-linear decrease in the degree of coarticulation has been reported for French with a plateau from the mid-50s to 70 years and an abrupt drop beyond 70 years (D'Alessandro & Fougeron, 2021). Not much is known about aging effects on speech kinematics in the articulatory dimension. In an exploratory study with eight German speakers in total (D'Alessandro & Fougeron, 2021) compared older and younger speakers and reported longer and spatially reduced vocalic tongue body movements with asymmetrical velocity profiles in older speakers. The present study investigates at what age changes in vocalic tongue body movements appear and how kinematic parameters are modulated in different age groups. In this study, we provide further evidence that changes in speech motor control occur gradually across the lifespan in a non-linear way.

Method Speech data of 44 German speakers aged between 19 and 79 years were collected by using electromagnetic articulography (AG501). Target words (girl names with C1V1.C2V2-structure) were embedded in a carrier sentence and appeared in three different focus conditions. V1 contains one of the five peripheral vowels: /i, e, a, o, u/. Vocalic tongue body movement durations, amplitudes, maximum speed and symmetry profiles of the backmost tongue body sensor for the production of V1 are reported. Speakers were divided into five age groups with differing group spans to especially capture age-related effects from 55-59, 60-69, and over 70 years.

Results Our results show that above the age of 60 years, articulatory movements considerably slow down (Fig.1a), which is also reflected in longer deceleration phases (Fig.1d), entailing asymmetric movement profiles. Furthermore, we found an increase of maximum speed from 55-59 years, before an abrupt decrease of speed over 60 years (Fig.1b). Simultaneously, amplitudes of tongue body movements are larger in the group 55-59 years and become considerably smaller with increasing age (Fig.1c), while movement durations remain similar to the younger speaker groups. Above the age of 60 years, movement duration increases while at the same time hyper-articulation (movement amplitude) is reduced. The increase of speed and the hyper-articulation of tongue body movements at the age of 55-59 can be related to target modifications of the speakers' speech system. To maintain the speed, speakers probably use more physical effort while speaking to compensate for the aging process. This could further lead to more variability within this middle-aged group, as reported in previous studies (D'Alessandro & Fougeron, 2021). The observed reduction in speed and prolonged deceleration phases for speakers over 60 years could be explained by a decrease in sensory feedback (probably due to changes in damping of the speakers' speech system).

Conclusion We conclude that tongue kinematics change above the age of 55. After a transition phase from 55-60 years, speech patterns level out and show characteristics of longer, slower and smaller movements.



Means and standard errors for articulatory parameters of vocalic tongue body movements across different age groups and focus conditions: a) duration in ms, b) velocity mm/s, c) amplitude/displacement in mm and d) symmetry (ratio of deceleration to acceleration phase).

References

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

D'Alessandro, D., & Fougeron, C. (2021). Changes in Anticipatory VtoV Coarticulation in French during Adulthood. *Languages, 6*(4), 181.

Karlsson, F., & Hartelius, L. (2021). On the primary influences of age on articulation and phonation in maximum performance tasks. *Languages, 6*(4), 174.

Thies, T., Hermes, A., & Mücke, D. (2022). Compensation in time and space: Prominence marking in aging and disease. *Languages, 7*(1), 21.

LEXICAL TONE BUT NOT ARBITRARY F0 IS CO-PLANNED WITH SEGMENTAL GESTURES

Emily Tesch*, Robin Karlin*, Ding-lan Tang, Caroline A. Niziolek°, Benjamin Parrell°
University of Wisconsin-Madison, Madison, United States of America
**equal contribution*
°equal contribution

The relationship between segmental and lexical tone planning in speech has long been debated. Some evidence suggests that tone is planned after segments, based on speech errors (Chen, 1999), facilitation effects (O'Seaghdha et al., 2010), and the coordination of articulatory gestures (Geissler et al., 2021); however, other evidence in the same experimental paradigms suggests that tone is co-planned with segments (Wan & Jaeger, 1998; Zeng et al., in prep; Gao, 2008; Zhang et al., 2019). In this study, we directly probe how the planning of tone relates to segmental planning using simultaneous adaptation to opposing sensory perturbations, contrasting lexical tone in Mandarin with arbitrary f0 in English.

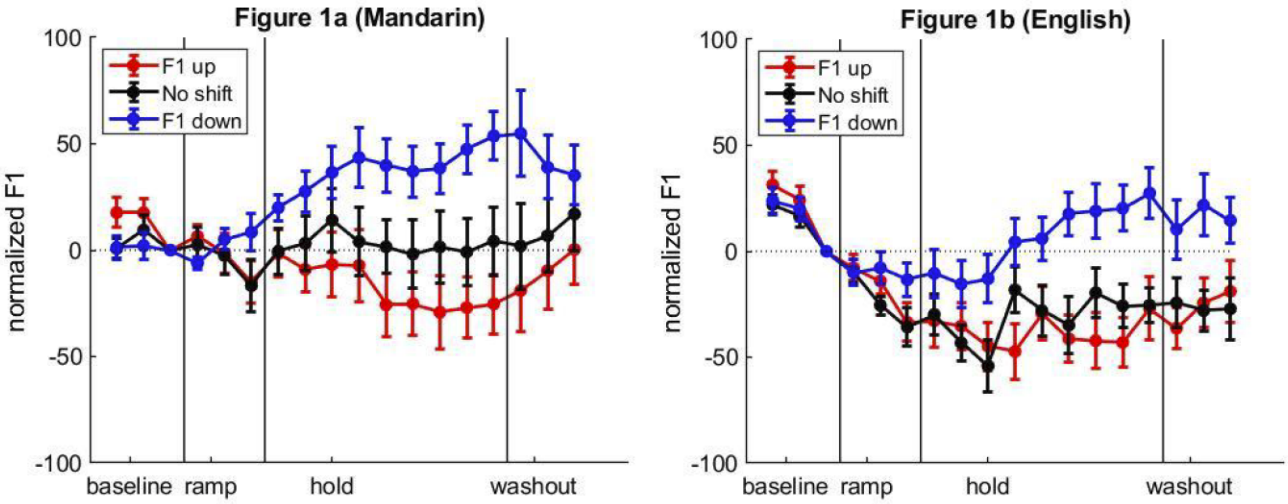
Sensory perturbations can be used to assess the scope of motor planning: people differentially adapt to simultaneous, opposing perturbations of kinematically identical movements, as long as those perturbations are tied to unique movement plans. For example, when participants received one perturbation on a forward reach followed by a rightward reach, and an opposing perturbation on a forward reach followed by a leftward reach, they differentially adapted the (identical) forward reach to counteract the perturbations (Sheahan et al., 2016). Critically, they also adapted when they planned to execute both reaches but stopped before performing the second reach, indicating that planning, rather than execution, provides the critical context for differential learning. Conversely, participants did not adapt if they saw a left/right cue but did not plan the second reach, indicating that cues unrelated to motor planning, even if fully predictive, do not provide sufficient context for motor adaptation.

Similar results have been reported in speech studies, where speakers adapt to opposing formant perturbations applied to the same vowel in different monosyllabic words, suggesting that words or syllables are planned as cohesive units (Rochet-Capellan & Ostry, 2011). Here, we assess whether differences in lexical tone provide a similar context for differential learning as segmental context in order to directly test whether tone and segmental articulation form a cohesive motor plan.

Methods Experiment 1: 9 native speakers of Mandarin Chinese (20 anticipated) said three words differing only in lexical tone: 飞 /fei1/ "fly"; 肥 /fei2/ "fat", and 费 /fei4/ "cost". Experiment 2: 9 native speakers of American English (20 anticipated) said the word "head", matching pitch to a pure tone that was high, mid, or low compared to their typical f0. In both experiments, each f0 context received one of three vowel formant perturbations (F1 up, F1 down, no perturbation), counterbalanced across participants.

Results Mandarin speakers consistently differentially adapt each word: speakers lower F1 in words that received upward perturbation, raise F1 in words that received downward perturbation, and do not change words that were not perturbed (Fig 1a). In contrast, there is more variability across English-speaking participants, who consistently fail to produce a robust three-way F1 contrast between the three f0 values of "head" (Fig 1b).

The reduced ability of English speakers to differentially adapt their productions of "head" as cued by arbitrary f0 levels suggests that they do not plan f0 and segmental movements together, even though they are executed simultaneously. In Mandarin, however, the differential adaptation observed here strongly suggests that tone is co-planned with segmental articulation.



(a) changes in the production of F1 in the Mandarin vowel /ei/ in three words differentiated only by tone; (b) changes in the production of F1 in the English vowel /ɛ/ in the word "head" with three f0 contexts.

References

Chen, J.-Y. "The representation and processing of tone in Mandarin Chinese: Evidence from slips of the tongue," *Applied psycholinguistics, vol. 20*, no. 2, pp. 289–301, 1999.

Gao, M. "Mandarin tones: An articulatory phonology account," PhD Thesis, Yale University, 2008.

Geissler, C., J. Shaw, F. Hu, and M. Tiede, "Eccentric CV timing across speakers of diaspora Tibetan with and without lexical tone contrasts," 2021.

O'Seaghdha, P. G., J.-Y. Chen, and T.-M. Chen, "Proximate units in word production: Phonological encoding begins with syllables in Mandarin Chinese but with segments in English," *Cognition, vol. 115*, no. 2, pp. 282–302, 2010.

Rochet-Capellan, A. and D. J. Ostry, "Simultaneous acquisition of multiple auditory–motor transformations in speech," *Journal of Neuroscience, vol. 31*, no. 7, pp. 2657–2662, 2011.

Sheahan, H.R., D. W. Franklin, and D. M. Wolpert, "Motor planning, not execution, separates motor memories," *Neuron, vol. 92*, no. 4, pp. 773–779, 2016.

Wan, I.-P. and J. Jaeger, "Speech errors and the representation of tone in Mandarin Chinese," *Phonology, vol. 15*, no. 3, pp. 417–461, 1998.

Zeng, Y., Zhang, J, and Fiorentino, R., "Early encoding of lexical tone in Chinese spoken word production: An online picture-word interference study," in prep.

Zhang, M., C. Geissler, and J. Shaw, "Gestural representations of tone in Mandarin: evidence from timing alternations," in *Proceedings of the 19th International Congress of Phonetic Sciences*, 2019, pp. 1803–1807.

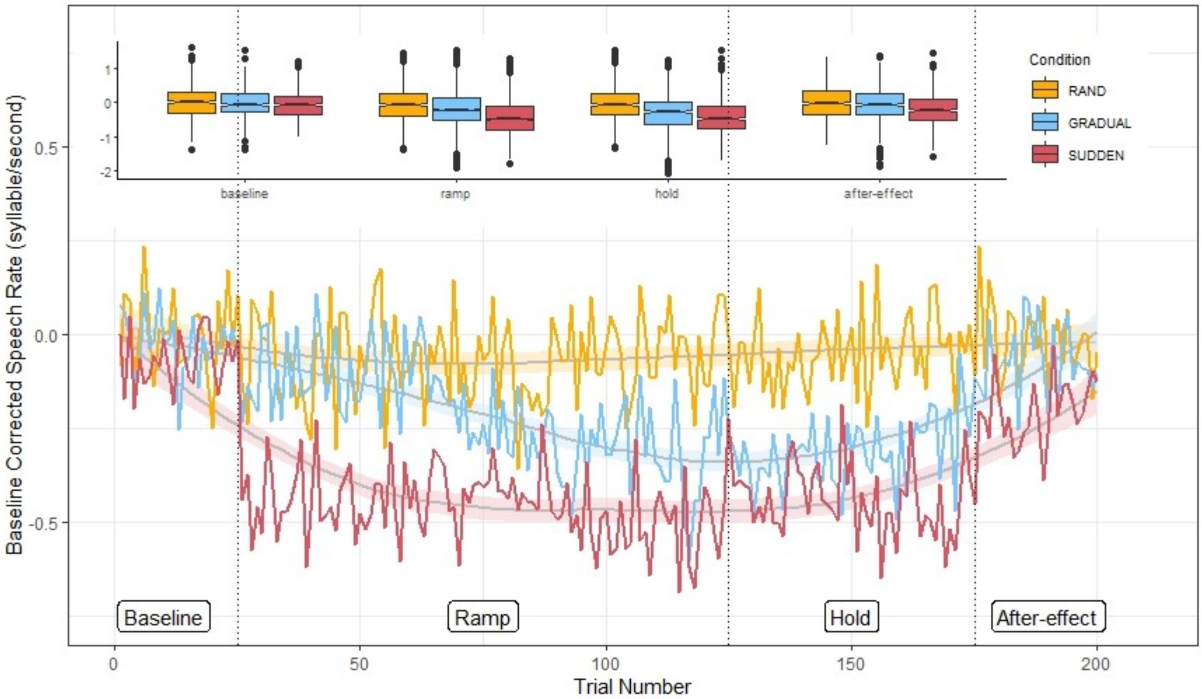
ADAPTATION TO DELAYED AUDITORY FEEDBACK

Monique Tardif[1,2], Jason Bohland[1,2]
[1] *University of Pittsburgh, Pittsburgh, Pennsylvania, United States of America*
[2] *Carnegie Mellon University, Pittsburgh, Pennsylvania, United States of America*

Models of speech motor control typically include internal forward models that predict the auditory consequences of planned speech, which is then compared to auditory feedback. Auditory-motor adaptation experiments are used to investigate the short-term plasticity of spectral (what) predictions from the forward model by gradually altering the feedback signals heard by participants over many trials. However, these studies have not established the extent of plasticity of temporal (when) forward model predictions. Delayed auditory feedback (DAF) is commonly used to modify when feedback arrives and causes noticeable rate reductions during speech (Yates, 1961). Karlin, Naber, and Parrell (2021), showed that individuals are sensitive to focal changes in feedback timing during vowel productions. **The goal of this study is to quantify adaptation to gradual changes in latency of delayed auditory feedback (DAF).** We did so by applying paradigms previously used for measuring adaptation to gradual spectral changes. Katz and Lackner (1977) showed that participants adapted to a DAF task similar to the one used presently, but to our knowledge that study has not been replicated.

In this study, participants repeated 3-syllable nonword sequences in three experimental conditions (200 trials each), with order counterbalanced across participants. The dependent variable was speech rate, which is typically reduced under DAF. Each condition was divided into phases: baseline (25 trials), ramp (100 trials), hold (50 trials), and after-effect (25 trials). In the GRADUAL condition, delay was linearly increased from 25 to 100 ms during the ramp and held at 100 ms throughout the hold. In the SUDDEN condition, delay was set to 100 ms throughout the ramp and hold. In the RANDOM condition, the same delays experienced throughout the GRADUAL condition were presented in random order. **We hypothesized that participants would partially adapt to delays in the GRADUAL condition, leading to reduced changes in speech rate compared to the SUDDEN and RANDOM conditions.**

Speech rate was estimated for each trial using the VocalToolkit plugin for Praat and normalized to the mean rate in the baseline phase. A linear mixed-effects model estimated fixed effects of condition (SUDDEN, GRADUAL, RANDOM) and phase (baseline, ramp, hold, after-effect) and random effects of participant and condition order on baseline-corrected speech rate. We found a main effect (relative to baseline) of ramp (β =-.18, SE = .02, p $<$ .001), and hold (β =-.28, SE = .02, p $<$ .001) (see Figure 1). Importantly, there were significant differences between GRADUAL and SUDDEN conditions across the ramp (β =-.11, SE = .03, p $<$ .001) and hold (β =-.22, SE = .04, p $<$ .001) compared to the baseline phase. The effects of the GRADUAL vs. RANDOM conditions were also different for the ramp (β =-.20, SE = .03, p $<$ .001) and hold (β =-.09, SE = .04, p = .01). Overall, these results demonstrate that speakers were able to maintain a faster speaking rate when delays were introduced gradually over time rather than suddenly or randomly. This evidence supports the idea that forward temporal predictions are adaptable during speech.



Mean speech rate (syllables per second) across participants under RANDOM, SUDDEN, or GRADUAL onset of delayed auditory feedback (DAF). Coloured lines represent moving averages (across 3 trials) after normalization to the baseline phase of the experiment (labelled in the main graph and indicated on the x-axis of the inset plot). The grey lines represent smoothed loess function estimates of the results. The colours represent the three experimental conditions: yellow –RANDOM presentation of DAF (between 25-100ms). Blue –GRADUAL increase of DAF (from 25-100ms during the ramp phase), and red - SUDDEN onset of 100ms on the 26th trial. The inset shows mean speech rate boxplots averaged within each of the experimental phases for each condition, coded in the same colours as the main plot.

References

Karlin, R., Naber, C., & Parrell, B. (2021). Auditory Feedback Is Used for Adaptation and Compensation in Speech Timing. *Journal of Speech, Language, and Hearing Research, 64*(9), 3361-3381.

Katz, D. I., & Lackner, J. R. (1977). Adaptation to delayed auditory feedback. *Perception & Psychophysics, 22*(5), 476-486.

Yates, A. J. (1963). Delayed auditory feedback. *Psychological Bulletin, 60*(3), 213.

COMPARISON OF ACOUSTIC PARAMETERS OF INHALATIONS VS. EXHALATIONS WITH 3D-PRINTED VOCAL TRACT MODELS

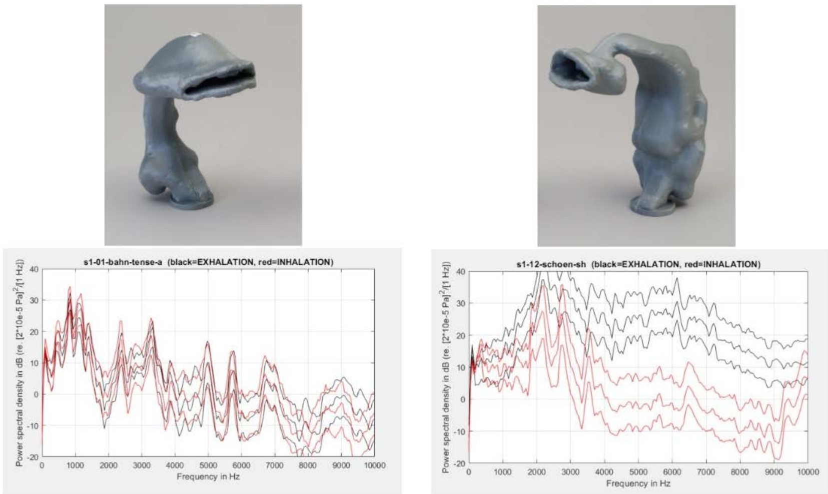
Raphael Werner[1], Susanne Fuchs[2], Jürgen Trouvain[1], Steffen Kürbis[3], Bernd Möbius[1], Peter Birkholz[3]
[1]*Saarland University, Saarbrücken, Germany*
[2]*Leibniz Centre General Linguistics (ZAS), Berlin, Germany*
[3]*TU Dresden, Dresden, Germany*

Our research focuses on the acoustic characteristics of inhalation noises in speech and the underlying articulatory mechanisms. Previously, we have shown (Werner et al., 2021) two main similarities of breathing noise to selected speech sounds: a) enhanced amplitudes at frequencies corresponding to low vowel formants and b) spectral characteristics of voiceless obstruents with a back place of articulation. This comparison is limited in that speech sounds are generally egressive, while inhalation noise is produced with an ingressive airstream. However, the airstream direction might be crucial for the production of noise, downstream of a constriction. To better understand the impact of airstream direction on acoustic properties, we used vocal tract models, producing four vowels and four fricatives. This allows us to analyze the spectra of the radiated sounds concerning a change of airflow direction while keeping the oral configuration constant.

For this, 3D-printed vocal tract models, excluding the nasal cavity, were used that were based on MRI data of a male and female German speaker's vocal tract producing /iː, aː, uː, ɘ, x, ç, ʃ, s/ (Birkholz et al., 2020) (see Fig. 1). To imitate in- and exhalations they were supplied with static airflow through the glottis at three fluid power levels in two airflow directions. Overall, we thus had 96 recordings (8 vocal tract configurations $\times$ 2 directions $\times$ 2 model speakers $\times$ 3 power levels). Although analyzing inhalations using LPC-based formant tracking looked promising (Werner et al., 2021), here we used the averaged power spectral density of the sounds produced over 10 s, as the assumption for a voiced source is not met here. To characterize and compare the sound spectra we calculated coefficients of the Discrete Cosine Transforms (DCT) 0-3. DCT0 corresponds to its mean amplitude, DCT1 to its slope, DCT2 to its curvature, and DCT3 to the amplitudes of the higher frequencies (Jannedy & Weirich, 2017). For each DCT coefficient, we fitted a linear mixed effects model with direction and vocal tract setting, and their interactions, as predictors. The models also included random intercepts for speaker and power level. We used lme4 (Bates et al., 2015) for model fitting and emmeans (Lenth, 2021) for pairwise post-hoc comparisons between in- and exhalations for each configuration. Since we found no significant interaction for DCT3, we used an additive model there.

None of the models returned a main effect for direction. For DCT0, the post-hoc comparisons showed a significant direction contrast for /iː, ç, ʃ, s/ with significantly higher DCT0 values in exhalation for all four. These sounds are produced with high tongue positions that lead to a concentrated airstream hitting the incisors. This obstacle source amplifies the signal in exhalation, but is much weaker in inhalations. For DCT1, inhalation and exhalation were significantly different in /ʃ, s/, and for DCT2 only in /ʃ/.

These results suggest that reversing the airstream direction with a noisy source has segment-specific effects on the spectrum's mean amplitude, slope, curvature, and higher frequencies. Rather than a general effect of direction, differences are found for sibilants, especially /ʃ/, and for mean intensity in settings involving high tongue positions.



Top: two of the 3D-printed vocal tracts corresponding to a male speaker producing the sounds /aː/ (left) and /ʃ/ (right); bottom: power spectral densities (0–10 kHz) of the radiated sounds generated with a static airflow through the glottis at different fluid power levels for the model representing /aː/ (left) and /ʃ/ (right). Exhalations are shown in black, inhalations in red.

References

Bates, D., Mächler, M., Bokler, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software, 67*(1), pp. 1–48.

Birkholz, P., Kürbis, S., Stone, S., Häsner, P., Blandin, R., & Fleischer, M. (2020). Printable 3D vocal tract shapes from MRI data and their acoustic and aerodynamic properties. *Scientific Data, 7*(1), pp. 1–16.

Jannedy, S., & Weirich, M. (2017). Spectral moments vs discrete cosine transformation coefficients: Evaluation of acoustic measures distinguishing two merging German fricatives. *The Journal of the Acoustical Society of America, 142*(1), pp. 395–405.

Lenth, R.V. (2021). *emmeans: Estimated Marginal means, aka Least-Squares Means*, r package version 1.6.1.

Werner, R., Fuchs, S., Trouvain, J., & Möbius, B. (2021). Inhalations in Speech: Acoustic and Phys-iological Characteristics. *Interspeech 2021*, pp. 3186–3190.

RATE-RELATED CHANGES IN MOVEMENT TRAJECTORY CHARACTERISTICS AND THEIR EFFECT ON SPATIOTEMPORAL VARIABILITY

Alan Wisler[1], Annalise Fletcher[1], Antje Mefferd[2]
[1] *Utah State University, Logan, Utah, United States of America*
[2] *Vanderbilt University, Nashville, Tennessee, United States of America*

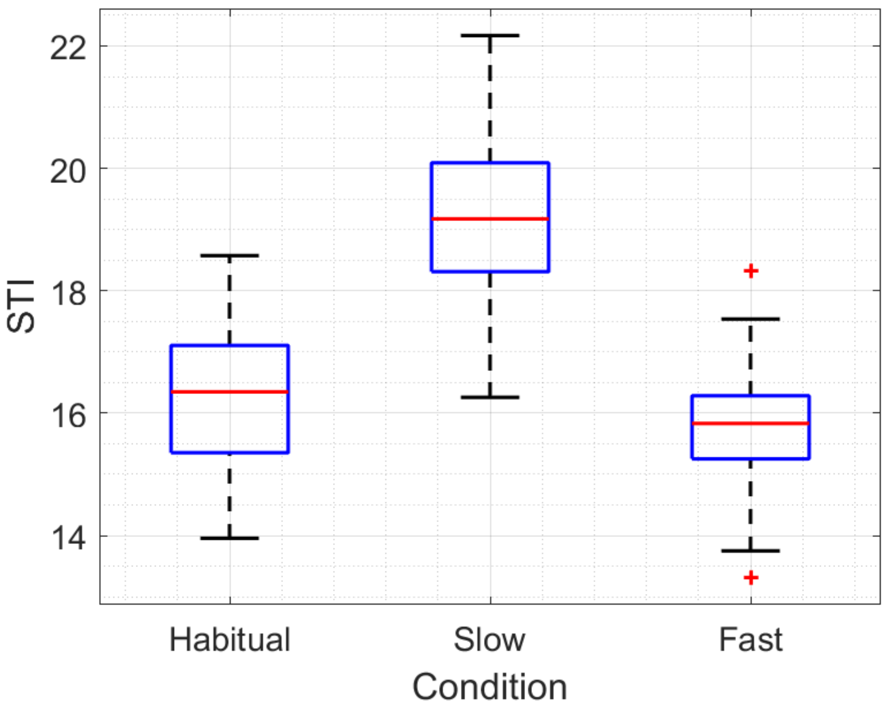
Background Due to their theoretical and clinical relevance, numerous kinematic studies have examined how changes in speech rate affect articulatory control. Two important findings are that 1) trial-to-trial variability of spatially and temporally normalized movement trajectories increases significantly during slow speech relative to typical and fast speech in typical adult talkers and 2) pattern recognition software can reliably differentiate the spatiotemporal patterns of slow rates from those of other speech rates (Smith et al., 1995). A distinct spatiotemporal pattern for slow speech is not surprising considering the relatively long periods of positional plateaus that commonly occur during slow speech but not during habitual and fast speech. However, the question arises whether the elevated spatiotemporal index (STI) values associated with slow speech may be predominantly driven by trajectory characteristics rather than rate-related differences in motor control. Therefore, the current study pursued two goals: 1) to identify differences between the signal characteristics of slow, habitual and fast speech and 2) to determine the effect of rate-specific trajectory properties on STI values.

Methods To address the first goal, we examined sets of 10 lower lip movement trajectories of the phrase "Buy Bobby a puppy" produced by a typical young adult under habitual, fast and slow speech. For each set of productions, we selected the trajectory that exhibited the highest average correlation with the other nine trajectories as the rate-specific template. Fourier analysis was applied to each template to decompose the signal into 15 sine and cosine waves. To address the second goal, we applied the findings of the Fourier analysis to a simulation experiment (see also Lucero, 2005) to determine changes in STI values in response to a controlled behavioral system (i.e., a fixed level of inherent variation of the trajectory pattern). Using this procedure, 100 sets of ten synthetic trajectories were generated for each rate. To evaluate the effect of the signal properties, the magnitude of change in STI values from habitual to slow and fast speech was calculated using Cohen's d.

Findings Inspection of the Fourier coefficients showed that slow speech yielded greater magnitudes in higher order coefficients than habitual and fast speech. However, the Fourier coefficients did not reveal any differences in signal characteristics between habitual and fast speech. Incorporating these templates into the simulation showed a noticeable increase in the STI for slow speech (Cohen's $d = 2.586$) and a small decrease in the STI for fast speech (Cohen's $d = 0.548$) relative to habitual speech.

Conclusion Findings suggest that signal properties differed between slow speech relative to habitual and fast speech. The higher order coefficients that were observed for slow speech represent the relative frequency following temporal normalization and are likely a result of higher frequency transition periods and lower frequency periods of positional plateaus. These rate-related differences in the signal characteristics resulted in increased STI values. Thus, STI changes in response to slow speech may not be solely driven by

changes in speech motor control. Future work is warranted to determine if findings of our modeling study can be replicated using empirical data of various speaker groups (e.g., typical talkers, talkers with motor speech disorders).



Range of spatiotemporal index (STI) values for slow, habitual, and fast speech simulations.

References

Lucero, J. C. (2005). Comparison of measures of variability of speech movement trajectories using synthetic records. *Journal of Speech, Language, and Hearing Research, 48*(2), 336–344. `https://doi.org/10.1044/1092-4388(2005/023)`

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. `https://doi.org/10.1007/BF00231983`

A MATTER OF TIME: AN ONLINE EXPERIMENT INVESTIGATING THE IMPACT OF POST-PRACTICE REST AND SLEEP ON SPEECH-MOTOR LEARNING

Anne L. van Zelst, F. Sayako Earle
University of Delaware, Newark, United States of America

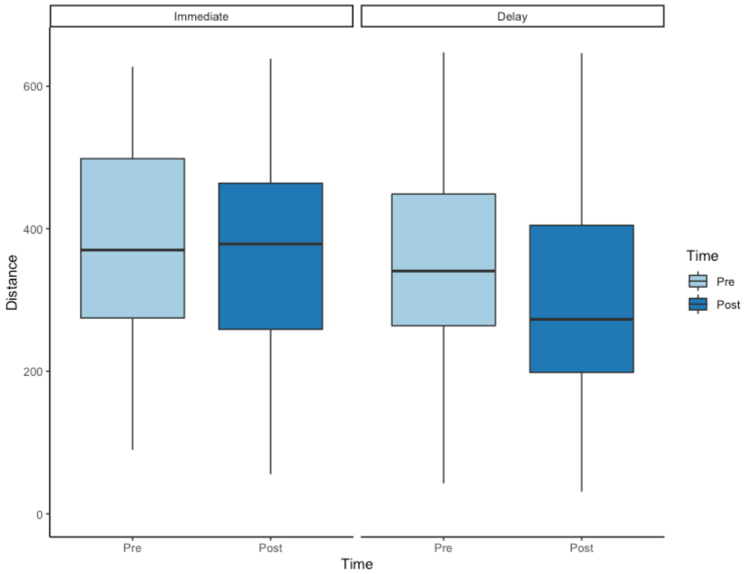
Timing speech-motor practice relative to wakeful rest or sleep may promote memory consolidation of novel speech-motor representations (Dudai, 2004). For speech-motor learners, how time is spent immediately following speech-motor practice may modulate how performance is enhanced in the absence of further training (Dudai, 2004; Robertson, 2009; Walker &Stickgold, 2004). Here, using a web-administered experiment delivered via Gorilla Experiment Builder (Anwyl-Irvine et al., 2020), we evaluated whether a post-practice period of rest or one containing nocturnal sleep could facilitate increased movement accuracy. Within a pseudo-randomized, between-group research design, we examined the effects of time (REST, SLEEP, IMMEDIATE) on the speech-motor production of a trained vowel contrast.

To date, a total of 46 native speakers of American English aged 18 to 25-years, with typical speech, language, hearing, and cognitive skills have been recruited (recruitment target: n= 48). Depending on their experimental group, learners logged on between 8:00-9:00AM or 8:00-9:00PM and trained in the production of two non-native Danish vowels, [y] and [ø]. Participants first completed baseline assessments and sleep habit questionnaires, followed by the speech sound production training. They then had a 12-hour delay with (SLEEP) or without nocturnal sleep (REST) or proceeded immediately (IMMEDIATE) to a post-training production assessment of Danish vowels in trained ([V]) and untrained ([hVd]) contexts. F1-F2-F3 formants were measured using Praat (Boersma & Weenink, 2022) . Then, movement accuracy was measured by the Euclidean distances between the F1-F2-F3 values in participant productions recorded during the pre and post assessments, against that of the Danish speaker model.

Group differences were analyzed via linear mixed-effects analyses on the outcome measures with an interaction between Time and Group, subsuming main effects, and context (trained vs. untrained) as factors, and the maximal random effects structure as supported by the data. After applying this approach to the preliminary data, we observed an emerging trend where Euclidean distance decreases between pre-training and post-training for the two 12-hour delay groups (REST and SLEEP), but not for the IMMEDIATE group (see Figure 1). When the two 12-hour delay groups were compared, there was also a significant three-way interaction between Group, Time, and Training, suggesting that the magnitude of improvement over Time and Training differs across the REST and SLEEP groups.

To ensure the ecological validity of our outcome measures, we are also obtaining perceptual ratings of the productions by native Danish speakers. Six native Danish speakers have been recruited to date (recruitment target: n=40) as participants in a web-administered perceptual identification task of the American English speakers' productions. This data will be collected and analyzed by the time of the conference.

These initial results advance the state of the current literature on the time course of speech-motor learning. Specifically, to optimize speech-motor learning, there may be practical applications in educational and clinical settings for timing speech-motor practice relative to a period of rest or sleep.



Euclidean Distance Pre and Post Training

References

Anwyl-Irvine, A. L., Massonnié, J., Flitton, A., Kirkham, N., & Evershed, J. K. (2020). Gorilla in our midst: An online behavioral experiment builder. *Behavior research methods, 52*(1), 388-407.

Boersma, Paul & Weenink, David (2022). Praat: doing phonetics by computer [Computer program]. Version 6.2.09.

Dudai, Y. (2004). The neurobiology of consolidations, or, how stable is the engram?. Annu. Rev. Psychol., 55, 51-86.

Robertson, E. M. (2009). From creation to consolidation: a novel framework for memory processing. *PLoS biology, 7*(1), e1000019.

van Zelst, A. L., & Earle, F. S. (2021). A case for the role of memory consolidation in speech-motor learning. *Psychonomic Bulletin & Review, 28*(1), 81-95.

Walker, M. P., & Stickgold, R. (2004). Sleep-dependent learning and memory consolidation. *Neuron, 44*(1), 121-133.

[Funded by a University of Delaware Summer Doctoral Fellowship].

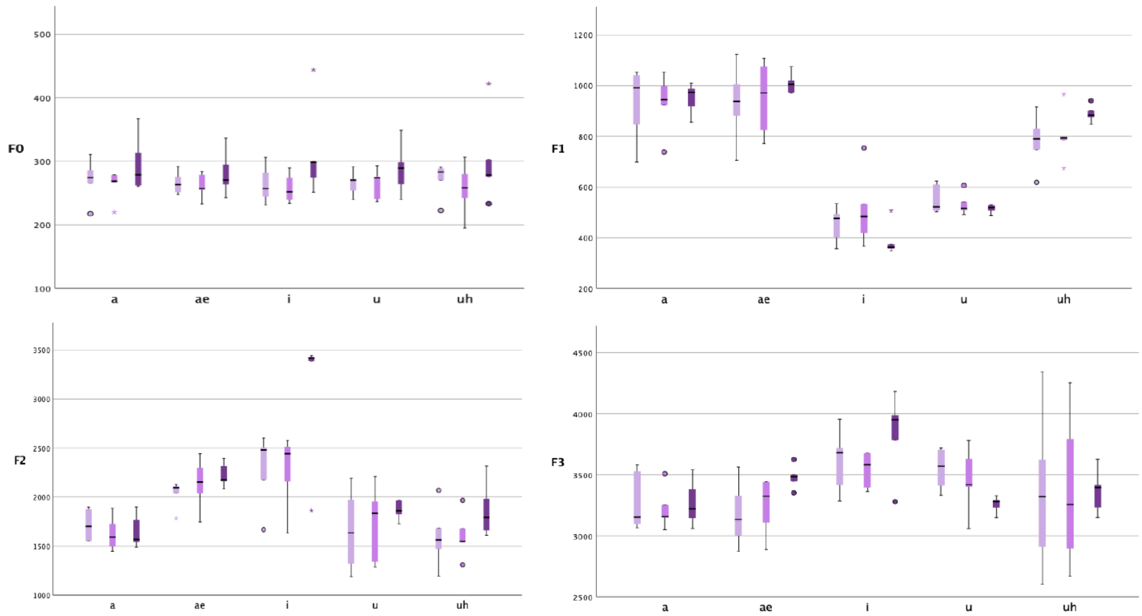
CHILD FORMANT MEASUREMENTS FROM REMOTELY-COLLECTED WAV AND M4A FILES ARE SIMILAR

Karen V. Chenausky[1,2,3], Helen Tager-Flusberg[3], Michaela Flaherty[1],Jordan Green[1]
[1] *MGH Institute of Health Professions, Boston, Massachusetts, USA*
[2] *Harvard Medical School, Boston, Massachusetts, USA*
[3] *Boston University, Boston, Massachusetts, USA*

One of the challenges of remote speech data collection is that apps using lossy compression audio file formats are more common than lossless file formats. While previous work (Sanker et al., 2021) suggests that there is no significant effect of file format on mean F0, F1, or F2 from adult speech, no similar comparisons have been performed for children's speech, especially children who may have speech disorders. To determine whether F0 and formant frequency measurements for (possibly disordered) children's speech differed according to file format, we compared measurements made across WAV and M4A file types.

Five low- and minimally verbal autistic children repeated the syllables [ba], [bæ], [bi], [bu], and [bʌ] at least three times each. These were recorded during remote assessment sessions by children's parents using Lexis Audio Recorder, a free app available for iOS and Android that has no ads and does not collect or share user information. Parents shared the audio files using DropBox. We used Wavesurfer to measure F0, F1, F2, and F3 values at one spectral slice in mid-vowel under three conditions: (1) automatically on the original file in WAV format ("WAV"), (2) automatically on a version of the file that was converted to M4A format and back to WAV format ("M4A"), and (3) manually on the original file in WAV format ("manual"). We then compared measured values from the three conditions using a repeated-measures ANOVA and post-hoc paired t-tests. There were no significant differences in F0, F1, F2, or F3 frequency values according to file format except for F3 for [æ], F3 for [u], and F2 for [ʌ]. Post-hoc tests indicated that in no case were the automatic measurements made from WAV and M4A files significantly different. In all cases, it was the manually-measured values that differed from the automatic measurements. Manually-measured F3 for [æ] was significantly higher than that for automatically-measured WAV F3 (Δ = 303 Hz, p = .019, Cohen's d = 1.4) and automatically-measured M4A F3 (Δ = 241 Hz, p = .013, Cohen's d = 1.5). Manually-measured F3 for [u] was significantly lower than that for automatically-measured WAV F3 (Δ = -293 Hz, p = .015, Cohen's d = 1.8). Manually-measured F2 for [ʌ] was significantly higher than that for automatically-measured WAV F2 (Δ = 275 Hz, p = .028, Cohen's d = 1.5) and automatically-measured M4A F2 (Δ = 262 Hz, p = .028, Cohen's d = 1.5).

These results suggest that remote recordings of children's speech are feasible using the Lexis app and that automatic F0 and formant measurements made using Wavesurfer are unaffected by file format (WAV vs. M4A). Further work will include determining whether acoustic measures of vowel distinctiveness derived from automatic F0 and formant measurements are related to other longitudinal variables of interest in the speech of minimally verbal autistic children, such as expressive language scores.



Box-and-whisker plots of mean F0, F1, F2, and F3 values measured automatically on WAV files (ligh purple colour), automatically on M4A files (medium purple color), and manually on WAV files (dark purple color). Circles represent values more than 1.5 times the interquartile range from the first or third quartiles. Stars represent values more than 3 times the interquartile range from the first or third quartiles.

References

Sanker, C., Babinksi, S., Burns. R., Evans. M., Johns J., Kim, J., Smith, S., Weber, N. & Bowern, C. (2021) (Don't) try this at home! The effects of recording devices and software on phonetic analysis. *Language 97*(4): e360-e382.

INTELLIGIBILITY AND CLINICAL COMMUNICATION SCALES IN CHILDREN

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

Background & Aims Children with neurologic conditions often experience considerable communicative restrictions which are reflected, for example, in reduced intelligibility (Hustad et al., 2013). To describe their communication skills, sensitive and objective measures of intelligibility (usually expressed as the percentage of correctly transcribed words/syllables) constitute the gold standard. Furthermore, various subjective scales have been established as standard measures, e.g., the Communication Function Classification System (CFCS; Hidecker et al., 2011), and the Viking Speech Scale (VSS; Pennington et al., 2013). In addition, ratings of primary caregivers are often considered as well, particularly with regard to everyday communication (Natzke et al., 2020; Sakash et al., 2020). All these subjective measures are efficient in terms of time and resources, yet they are also rather global in nature and it has not yet been clarified conclusively how they relate to objective intelligibility measures. Moreover, studies have not yet systematically considered whether and to what extent these measures are subject to developmental influences. In this study, we aim to investigate the relationship between an objective intelligibility measure and the above-mentioned subjective clinical measures in children with neurological conditions. In a second step, we aim to analyze how much these measures are affected by developmental influences and what this implies for their clinical interpretation.

Methods Thirty children (11f; 5;1 - 9;10 years;months) with neurological conditions, and 144 typically developing children (72f; 3;0 – 9;11 years;months) participated. Speech samples were collected using the materials of the Bogenhausen Dysarthria Scales for Childhood Dysarthria (BoDyS-KiD; Haas et al., 2021), a German assessment tool for childhood dysarthria. To assess intelligibility, naïve listeners transcribed sentences repeated by the children. CFCS and VSS were rated by the examiners (i.e., the first and last author). Ratings of primary caregivers, that is of the children's parents, were obtained by means of a self-developed questionnaire that asks about the child's ability to communicate verbally with familiar/unfamiliar children/adults.

Results & Discussion Overall, the group of children with neurological conditions was rather heterogeneous with regard to all communication measures, with scores ranging from very mild to most severe. We found significant correlations between the objective intelligibility measure and CFCS, VSS as well as parents´ ratings. At the same time, it was apparent that intelligibility, as well as CFCS, were still developing considerably in the typically developing children between the ages of 3 and 10. This indicates that if - especially younger children - do not achieve high scores in intelligibility and CFCS, this does not necessarily reflect a speech motor disorder, but may still be age-appropriate (Haas et al., accepted). With regard to the parent questionnaire, we did not see any pronounced improvement across age. This could be because parents apply different standards to different age groups. Thus, even with older children, a lower parent rating does not necessarily indicate a disorder. Consequently, subjective measures represent important clinical metrics. However, there is strong evidence that they should be age-normalized in the future.

References

Haas, E., Ziegler, W., & Schölderle, T. (2021). Developmental courses in childhood dysarthria: Longitudinal analyses of auditory-perceptual parameters. *Journal of Speech, Language, and Hearing Research, 64*(5), 1421–1435.

Haas, E., Ziegler, W., & Schölderle, T. (accepted). Intelligibility, speech rate, and communication efficiency in children with neurologic conditions. A longitudinal study of childhood dysarthria. *American Journal of Speech-Language Pathology.*

Hidecker, M. J. C., Paneth, N., Rosenbaum, P. L., Kent, R. D., Lillie, J., Eulenberg, J. B., Chester, K., Johnson, B., Michalsen, L., Evatt, M., & Taylor, K. (2011). Developing and validating the Communication Function Classification System for individuals with cerebral palsy. *Developmental Medicine & Child Neurology, 53*(8), 704–710.

Hustad, K. C., Schueler, B., Schultz, L., & DuHadway, C. (2012). Intelligibility of 4-Year-Old Children with and without Cerebral Palsy. *Journal of Speech, Language, and Hearing Research, 55*(4), 1177–1189.

Natzke, P., Sakash, A., Mahr, T., & Hustad, K. C. (2020). Measuring Speech Production Development in Children With Cerebral Palsy Between 6 and 8 Years of Age: Relationships Among Measures. *Language, Speech, and Hearing Services in Schools, 51*(3), 881-896.

Pennington, L., Virella, D., Mjøen, T., da Graça Andrada, M., Murray, J., Colver, A. F., Himmelmann, K., Rackauskaite, G., Greitane, A., Prasauskiene, A., Andersen, G. L., & La Cruz, J. de (2013). Development of The Viking Speech Scale to classify the speech of children with cerebral palsy. *Research in Developmental Disabilities, 34*(10), 3202–3210.

Sakash, A., Mahr, T., & Hustad, K. C. (2020). Validity of Parent Ratings of Speech Intelligibility for Children with Cerebral Palsy. *Developmental Neurorehabilitation, 24*(2), 98-106.

ATYPICAL SPEECH ACOUSTICS AND JAW KINEMATICS DURING AFFECT PRODUCTION IN CHILDREN WITH AUTISM SPECTRUM DISORDER ASSESSED BY AN INTERACTIVE MULTIMODAL CONVERSATIONAL PLATFORM

Hardik Kothare[1], Vikram Ramanarayanan[1,2], Oliver Roesler[1], Michael Neumann[1], Jackson Liscombe[1], William Burke[1], Andrew Cornish[1], Doug Habberstad[1], Brandon Kopald[2], Alison Bai[2], Yelena Markiv[2], Levi Cole[2], Sara Markuson[2], Yasmine Bensidi-Slimane[2], Alaa Sakallah[2], Katherine Brogan[2], Linnea Lampinen[2], Sara Skiba[2], David Suendermann-Oeft[1], David Pautler[1], Carly Demopoulos[2]

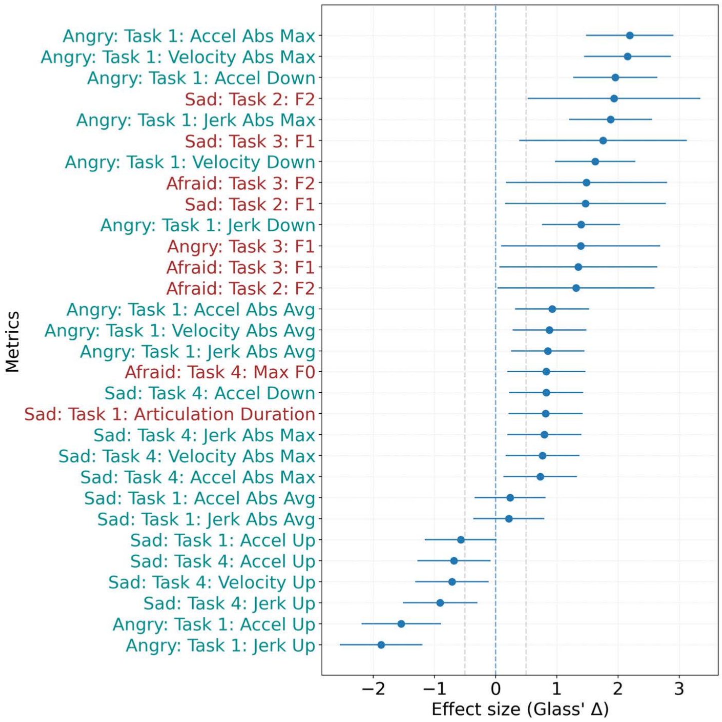
[1]*Modality.AI, Inc., San Francisco, California, United States of America*

[2]*University of California, San Francisco, California, United States of America*

A characteristic feature of Autism Spectrum Disorder (ASD) is atypical production of vocal and facial affect during emotional speech (Hubbard et al., 2017; Loveland et al., 1994). Prior work has shown that this atypical affect production correlates with accuracy in recognition of vocal and facial affect and can be captured using an interactive multimodal conversational platform (Kothare et al., 2021). In this work, we investigate which speech acoustic and facial kinematic metrics that are automatically extracted by the conversational platform show significant differences between participants with ASD and controls. In an affect production task designed in a conversational format, 44 ASD participants (16 female, mean age = 11.74 ± 2.56) and 17 controls (8 female, mean age = 12.80 ± 2.59) were prompted to produce one of four emotions (happy, sad, angry and afraid) through the following four tasks: repeating the monosyllable "oh" after (1) a video stimulus, (2) an audio stimulus, (3) a situation narration followed by a picture stimulus and (4) repeating the sentence "I'll be right back" after a video stimulus. The automatically-extracted speech acoustic and facial kinematic metrics were normalised by gender to account for gender-specific differences.

Non-parametric Kruskal-Wallis tests performed to investigate differences in metrics (see figure) between ASD and controls revealed that the ASD cohort exhibited greater velocity, acceleration and jerk of the jaw. These higher-order derivatives of the vertical movement of the jaw have larger values in ASD only for two of the four emotions— angry and afraid— and only in two of the four tasks where participants were asked to repeat monosyllabic or sentential speech after a video stimulus. These facial kinematic metrics also had greater variance in the ASD cohort as evaluated by Fligner-Killeen tests. This suggests that the ASD cohort exhibits exaggerated jaw movement while mimicking speech with negative emotions from a video stimulus but not when affect production is naturally elicited via a picture stimulus or the repetition of an audio-only stimulus. Additionally, spectral metrics of speech such as formant frequency values of the monosyllabic vowel /o/ elicited by a picture stimulus or the audio repetition of sad, afraid and angry emotions and the maximum fundamental frequency during afraid sentential repetition was larger in ASD than in controls. Articulation duration during the sad monosyllabic production of 'oh' was also greater in ASD. All the above differences showed a statistically significant difference at an alpha threshold of 0.05 and were controlled for false discovery rate.

These findings point towards exaggerated and variable speech motor control in ASD during repetition of emotional speech when the production is cued via a video. Additionally, acoustic properties of emotional speech in ASD are also atypical. Notably, these differences are specific to certain emotions providing a novel insight into the atypical production of vocal and facial affect during emotional speech in ASD.



Effect sizes of acoustic metrics (brown) and jaw kinematics (teal) that show statistically significant differences between pASD and controls at an alpha threshold of 0.05. Task 1: monosyllable "oh" video stimulus, Task 2: monosyllable "oh" audio stimulus, Task 3: monosyllable "oh" picture stimulus and Task 4: sentence video stimulus.

References

Hubbard, D. J., Faso, D. J., Assmann, P. F., & Sasson, N. J. (2017). Production and perception of emotional prosody by adults with autism spectrum disorder. *Autism Research, 10*(12), 1991-2001.

Loveland, K. A., Tunali-Kotoski, B., Pearson, D. A., Brelsford, K. A., Ortegon, J., & Chen, R. (1994). Imitation and expression of facial affect in autism. *Development and Psychopathology, 6*(3), 433-444.

Kothare, H., Ramanarayanan, V., Roesler, O., Neumann, M., Liscombe, J., Burke, W., Cornish, A., Habberstad, D., Sakallah, A., Markuson, S., Kansara, S., Faerman, A., Bensidi-Slimane, Y., Fry, L., Portera, S., Suendermann-Oeft, D., Pautler, D., Demopoulos, C. (2021) Investigating the Interplay Between Affective, Phonatory and Motoric Subsystems in Autism Spectrum Disorder Using a Multimodal Dialogue Agent. *Proc. Interspeech 2021*, 1967-1971

THE EFFECTIVENESS OF RAPID SYLLABLE TRANSITION TREATMENT IN IMPROVING COMMUNICATION IN CHILDREN WITH CEREBRAL PALSY: A RANDOMIZED CONTROLLED STUDY

Johana Korkalainen[1], Patricia McCabe[1], Andy Smidt[1], Catherine Morgan[1,2]
[1] *The University of Sydney, Sydney, Australia*
[2] *Cerebral Palsy Research Institute, Australia*

Background Cerebral palsy is a movement disorder that impacts communication. Most children with CP have speech disorders, particularly dysarthria or a combination of dysarthria and other speech impairments (Australian Cerebral Palsy Register 2018; Mei et al., 2020). Rapid Syllable Transition Treatment (ReST; McCabe et al., 2020) differs from the current motor speech interventions for children with CP in that it targets practising nonwords accurately and fluently rather than improving the individual speech subsystems (Hayden, 2006). A pilot study using ReST provided a proof of concept (Korkalainen et al., in press), so it was trialed in a larger study.

Aim The aim was to determine whether six weeks of ReST improved accuracy, speech intelligibility and communication participation in children with CP more than usual care.

Methods A single blind randomised control trial recruited 14 children aged 8-14 years old with moderate CP. Children were randomised using a concealed allocation to the ReST treatment group (n=8) or the usual-care group (n=6). The usual-care group remained with usual intervention for 6 weeks, the length of the ReST, after which they too received ReST therapy. ReST was provided for six weeks of therapy on a frequency of three one-hour sessions per week.

The primary outcomes included the percentage of phonemes correct in words (PPC_W), intelligibility in words (WIPI), and Intelligibility in Context (ICS). These were measured at pre and post intervention and at maintenance, two weeks post intervention. An additional time point for the usual-care group was post usual care. Student SLP clinicians, who were blind to group allocation, the identity of the participant, and the timing of the assessment and probes, scored and transcribed de-identified audio-recordings.The within-subjects variables were based on time (treated items at base line and one week post treatment). The between-subjects variable was treatment (ReST or usual care). This gave the results for each variable and the interaction between them (time x group). Repeated measures analysis of variance (ANOVA) with 95% confidence was conducted.

Results Significant group difference in favour of ReST therapy over usual care was shown in all the primary outcomes, WIPI (F=5.25, p=.04) and PPC in word (F=1.24, p= .29) and ICS (F=5.25, p= .04). Significant time by group difference was shown in PPC at word level (F=5.88, p= .03) and ICS (F=4.94, p= .001).

Conclusion ReST is more effective in improving speech intelligibility, speech accuracy and communicative participation in children with CP than usual care. Based on these results, ReST can offer an alternative motor speech intervention for children with CP. However, due to the small sample size, and inability to blind parents and treating clinicians, the results should be interpreted cautiously.

DIAGNOSTIC FEATURES OF CHILDHOOD APRAXIA OF SPEECH IN ESTONIAN, FINNISH AND LITHUANIAN: A SURVEY OF SLPS

M. Lahtein[1], M.-L. Mailend[1,2], M. Padrik[1], S. Daniute[2], D. Kariene[3], A.-L. Martikainen[4], M. Vanhala-Haukijärvi[5]
[1] *University of Tartu, Tartu, Estonia*
[2] *Moss Rehabilitation Research Institute, Elkins Park, Pennsylvania, United States of America*
[3] *Vilnius University, Vilnius, Lithuania*
[4] *Vytautas Magnus University, Kaunas, Lithuania*
[5] *University of Eastern Finland, Kuopio, Finland*
[6] *University of Turku, Turku, Findland*

Research on diagnostic features of childhood apraxia of speech (CAS) in English is making headway, however, studies in other, languages are scarce (Wong et al., 2020). Due to this gap in cross-linguistic research, it is unclear how applicable English-based diagnostic features are in languages other than English. This study investigated the diagnostic features that SLPs use for CAS diagnosis in three non-Germanic languages: Estonian and Finnish, both of which belong to the Finno-Ugric language family, and Lithuanian – a Balto-Slavic language.

A cross-linguistic survey was conducted among SLPs experienced with CAS from Estonia, Finland, and Lithuania. The participants were asked to rate the significance of 63 features for diagnosing CAS in their native language. The list included features that are commonly reported for English-speaking children with CAS, however, we also included language-specific features that may be associated with CAS diagnosis. For example, phonemic quantity is a central aspect in Estonian and Finnish, but not in Lithuanian. Given the duration errors prevalent in CAS (Grigos et al., 2015), we hypothesized that Estonian and Finnish SLPs are more likely to associate phonemic quantity errors with CAS compared to Lithuanian SLPs. Errors related to vowels are also common in children with CAS (ASHA, 2007). Thus, we expected diphthong reduction errors to be more common in Estonian (36 diphthongs) and Finnish (16 diphthongs) compared to Lithuanian (9 diphthongs). Finally, we predicted that palatalization errors are strongly associated with CAS in Lithuanian, where most consonants have a palatalized and non-palatalized variants that distinguish meaning. This could be challenging for children with CAS considering the precise coordination of tongue movements that is needed for palatalization (Grigos et al., 2015).

A total of 197 SLPs completed the survey (Estonia: n=59, Finland: n=69, Lithuania: n=69). An overview of the features rated as most significant for CAS diagnosis in the three languages is presented in Table 1. The table lists the features that were rated as a “very significant feature” by at least 50% of the SLPs from each country. Interestingly, prosodic errors were not rated as very significant in any of the languages, although they are considered a core feature of CAS in English (ASHA, 2007). The language-specific hypotheses were evaluated with ordinal logistic regression analysis. In line with our predictions, Lithuanian SLPs were more likely to regard palatalization errors as a significant feature for CAS diagnosis compared to Finnish SLPs ($t=2.80$; $p<.05$). Similarly, Estonian SLPs were more likely to regard diphthong simplification errors as a significant feature of CAS compared to Finnish SLPs ($t=2.10$; $p<.05$). Other comparisons did not reach statistical significance. Contrary to our predictions, no significant differences between languages were observed for quantity errors. In summary, the most highly rated features across the three languages include features that are commonly reported in English-speaking children with CAS. However, language-specific differences were also identified. These findings highlight the need for empirical studies of CAS in different languages.

	Estonian	Finnish	Lithuanian
1.	articulatory groping, restarts and starting difficulties 89,8%	difficulties with diadochokinetic tasks 73,9%	articulatory groping, restarting difficulties 56,5%
2.	difficulties with diadochokinetic tasks 84,7%	speech is unintelligible even for familiar listeners 69,6%	difficulties with volitional non-speech oral motor tasks, oral apraxia 55,1%
3.	ability to use sound/word spontaneously but not volitionally 64,4%	Increased errors with increased utterance length 59,4%	increased errors with increased utterance length 53,6%
4.	inconsistent errors in repeated productions of the same word, where a word is sometimes produced correctly and sometimes not 62,7%	increased errors as the complexity of the syllable shape increases 59,4%	difficulty with fine motor movements and poor coordination 53,6%
5.	increased errors as the complexity of the syllable shape increases 61,0%	inconsistent errors in repeated productions of the same word, where errors are made on different phonemes in each production 59,4%	speech is unintelligible even for familiar listeners 53,6%
6.	inconsistent errors in repeated productions of the same word, where errors are made on different phonemes in each production 61,0%	limited phonemic inventory of consonants and vowels 53,6%	increased errors as the complexity of the syllable shape increases 52,5%
7.	high amount of within-class manner or place substitution errors 52,5%	ability to use sound/word spontaneously but not volitionally 50,7%	

Features that were rated very significant for CAS diagnosis by the majority (at least 50%) of the Estonian, Finnish and Lithuanian SLPs.

References

American Speech-Language-Hearing Association. (2007). Childhood apraxia of speech [Technical Report]. https://www.asha.org/policy/TR2007-00278/

Grigos, M. I., Moss, A., & Lu, Y. (2015). Oral articulatory control in childhood apraxia of speech. *Journal of Speech, Language, and Hearing Research, 58*(4), 1103-1118.

Wong, E. C., Lee, K. Y., & Tong, M. C. (2020). The applicability of the clinical features of English childhood apraxia of speech to Cantonese: A modified Delphi survey. *American Journal of Speech-Language Pathology, 29*(2), 652-663.

WHEN DO SEQUENTIAL MOTION RATE TASKS GET FASTER THAN ALTERNATING MOTION RATE TASKS DURING DEVELOPMENT? ORAL-DIADOCHOKINETIC RATES OF NEUROTYPICAL FRENCH-SPEAKING CHILDREN, ADOLESCENTS AND YOUNG ADULTS

Monica Lancheros, Marina Laganaro
University of Geneva, Geneva, Switzerland

Oral diadochokinetic (DDK) tasks are commonly used to assess motor speech skills in both pediatric and adult populations. They require the repetition of a syllable (alternating motion rate -AMR- tasks) or of a cluster of varying syllables (sequential motion rate -SMR- tasks) at a maximum rate in a single breath. Several studies on neurotypical adults have shown that SMR tasks are produced faster than AMR tasks (e.g. Jang et al., 2020; Mousavi et al., 2020) . However, a screening of the results summarized in the review paper by Kent, Kim & Chen (2021) seems to indicate that such pattern is not present at all ages. In the present study, AMR and SMR tasks (/ba/, /de/, /go/ and /badego/, respectively) were produced by three different age groups: 7 to 9-year-old children, 14 to 16-year-old adolescents and 20 to 30-year-old young adults. The number of syllables produced over an interval of 4 seconds of continuous repetition (i.e. syllabic rate) was computed for each DDK type per participant.

Preliminary results on 15 participants per age group show no difference in the syllabic rates of SMR and AMR tasks in children (syllabic rates (SD): /ba/=4.42(1); /de/=4.58(0.81); /go/=4.15(0.69); /badego/=4.43(1.09)). Concerning the group of adolescents, the syllabic rate of the SMR task is higher from that of the AMR syllables /go/ and /de/ whereas no difference is found for the syllable /ba/ (ba/=5.16(0.76); /de/=5.10(0.50); /go/=4.79(0.48); /badego/=5.42(0.67)). As for the adults, they show higher rates for the SMR task as compared to each of the three AMR tasks (/ba/=6.08(0.74); /de/=5.92(0.81); /go/=5.59(0.61); /badego/=6.95(0.92)). Adults being faster in repeating a cluster of varying syllables relative to the repetitive production of a single syllable is consistent with findings of previous studies (e.g. Neel & Palmer, 2012) and suggests that producing an oromotor task closely related to speech (i.e. SMR task) reaches higher rates than those sharing less properties (i.e. AMR tasks; Lancheros, Pernon & Laganaro, 2022). Concerning adolescents, their results reflect adult-like performances since the same pattern was found between SMR and AMR tasks reported for adults, except for the syllable /ba/ which did not reach the threshold of significance by a small margin when compared to the SMR task. 7 to 9-year-old children show similar performances on both DDK tasks.

Our preliminary results suggest that it is only during middle adolescence that SMR and AMR tasks start showing the same patterns of performance usually reported in adults, which brings important insights on the moment at which closely-related speech tasks (e.g SMR) differ from speech-like tasks sharing less properties and principles with speech (e.g. AMR). In the following months we aim at including more participants per age group in order to confirm those findings with a more representative sample size.

References

Jang, J., Choe, Y. G., & Ha, S. (2020). Characteristics of diadochokinesis in typically developing children and adults. *Audiology and Speech Research, 17*(1), 73-80.

Kent, R. D., Kim, Y., & Chen, L. M. (2022). Oral and Laryngeal Diadochokinesis Across the Life Span: A Scoping Review of Methods, Reference Data, and Clinical

Applications. *Journal of Speech, Language, and Hearing Research, 65*(2), 574-623.

Lancheros, M., Pernon, M., & Laganaro, M. (2022). Is there a continuum between speech and other oromotor tasks? evidence from motor speech disorders. *Aphasiology*, 1-20.

Mousavi, S. Z., Mehri, A., Nabavi, D., Faraji, M., & Maroufizadeh, S. (2020). Comparing the diadochokinetic rate in Farsi-speaking young and older adults. *Iranian Rehabilitation Journal, 18*(1), 57-64.

Neel, A. T., & Palmer, P. M. (2012). Is tongue strength an important influence on rate of articulation in diadochokinetic and reading tasks? *Journal of Speech, Language, and Hearing Research, 55*(1), 235–246.

CONSONANT PRODUCTION IN CHILDREN WITH COCHLEAR IMPLANTS AND EXPOSED TO CANADIAN FRENCH CUED SPEECH: AN ACOUSTIC AND ARTICULATORY STUDY

Laura Machart[1,2], Anne Vilain[2], Hélène Lœvenbruck[1], Lucie Ménard[3]
[1] *University Grenoble Alpes, CNRS, LPNC, Grenoble, France*
[2] *University Grenoble Alpes, CNRS, Grenoble INPP*, GIPSA-Lab, Grenoble, France*
[3] *UQÁM, Montréal, Canada*

Although cochlear implant improves deaf children's speech intelligibility (Turgeon et al., 2017; Grandon et al., 2020), the auditory information it provides remains degraded (Colin et al., 2017). This auditory limitation can impact on oral language development and lead to persistent language disorders (Geers et al., 2015). Several acoustic studies have highlighted specific impairments in the speech production of children with CI (Grandon, 2016; Reidy et al., 2017) but only a few studies have examined how children with CI use their articulators to produce speech (Turgeon et al., 2017). To supplement degraded acoustic information, cued speech manual gestures can be used in conjunction with auditory and visual speech (Hage & Leybaert, 2006). French Cued Speech has been shown to improve auditory processing of sentences in children with hearing aids (Périer et al., 1990, Leybaert et al., 2010) and to help children build more stable phonological representations (Charlier & Leybaert, 2000), even in children with CI (Hage & Leybaert, 2006; Leybaert et al., 2010). It has also been suggested that exposure to French Cued Speech could improve production (Hage & Leybaert, 2006; Machart et al., 2019).

The aim of the present study is to examine the influence of Canadian French Cued Speech proficiency on the acoustic quality and accuracy of tongue gestures in children with CI. Based on speech production data obtained in French-speaking children with CI (Machart et al., 2019; in revision), we chose to focus on the two following stops /t/, /k/ and the two fricatives /s/, /ʃ/, for which substitution errors are common. Articulatory data were obtained using an ultrasound system (described in Ménard et al., 2014). The consonants were recorded during the production of simple words (Picture-naming task), each including one of the target consonants followed by the vowel /a/. Two groups of children were examined: 10 typical children with normal hearing (NH) and 9 children with cochlear implants who showed different cued speech proficiency levels (CI- and CI+). Acoustic data were phonetically annotated and analysed using PRAAT (Boersma & Weenink). Accuracy scores have been computed. Acoustic measurements such as formant transition for stops and spectral moments for fricatives have been obtained. Articulatory data have been extracted using SLURP (Laporte & Ménard, 2018). Tongue contour were assessed on horizontal positions of specific flesh points (Ménard et al., 2013; Ohkubo & Scobbie, 2019) as well as on MCI (Dawson et al., 2015).

The results suggest that cued speech proficiency sustains the development of speech production in children with cochlear implants, and improves their articulatory gestures, particularly for the place contrast in plosives as well as fricatives. Moreover, it appears that high cued speech proficiency may allow children with CI to produce acoustic and articulatory contrasts in the same way as children with normal hearing.

References

Boersma, P. & Weenink, D., 2019. Praat: doing phonetics by computer [Computer program]. Version 6.1.08.

Charlier, B. L., & Leybaert, J., 2000. The rhyming skills of deaf children educated with phonetically augmented speechreading. *Quarterly Journal of Experimental Psychology Section A: Human Experimental Psychology, 53*(2), 349–375.

Colin, S., Ecalle, J., Truy, E., Lina-Granade, G., Magnan, A., 2017. Effect of age at cochlear implantation and at exposure to Cued Speech on literacy skills in deaf children. *Res. Dev. Disabil. 71*, 61–69.

Dawson, K. M., Tiede, M. K., & Whalen, D. H., 2016. Methods for quantifying tongue shape and complexity using ultrasound imaging. *Clin. Ling. Phon., 30*(3–5), 328–344.

Geers, A., Nicholas, J., Tobey, E., Davidson, L., 2015. Persistent language delay versus late language emergence in children with early cochlear implantation (pp. 155-170). *Journal of Speech, Language, and Hearing Research, Volume 59.*

Grandon, B., 2016. *Développement typique et atypique de la production de parole : caractéristiques segmentales et intelligibilité de la parole d'enfants porteurs d'un implant cochléaire et d'enfants normo-entendants de 5 à 11 ans.* (Thèse de Linguistique). Université Grenoble Alpes.

Grandon, B., Martinez, M. J., Samson, A. & Vilain, A., (2020). Long-term effects of cochlear implantation on the intelligibility of speech in French-speaking children. *Journal of Child Language.*

Hage, C. & Leybaert, J., 2006. The Effect of Cued Speech on the Development of Spoken Language. In P. E. Spencer & M. Marschark, *Advances in the Spoken Language Development of Deaf and Hard-of-Hearing Children.* Oxford University Press, pp. 193–211.

Laporte, C., Ménard L., 2018. Multi-hypothesis tracking of the tongue surface in ultrasound video recordings of normal and impaired speech. *Medical Image Analysis. 44*: 98-114.

Leybaert, J., Colin, C., Hage, C., LaSasso, C.J., 2010. Cued Speech for Enhancing Speech Perception and First Language Development of Children With Cochlear Implants. *Trends Amplif. 14*, 96–112.

Machart, L., Meloni, G., Vilain, A., Lœvenbruck, H. & Schott-Brua, V. (2019, mai). Le développement phonologique de l'enfant implanté : l'apport de la Langue française Parlée Complétée. *VIIIème Journées de Phonétique Clinique*, Mons, 75-77.

Ménard, L., Leclerc, A. & Tiede, M. (2014) : Articulatory and acoustic correlates of contrastive focus in congenitally blind adults and sighted adults, *Journal of Speech, Language, and Hearing Research, 57*, 793-804. doi: 10.1044/2014.

Ménard, L., Perrier, P., & Aubin, J. (2013). The role of auditory feedback in speech development: A study of compensation strategies for a lip-tube perturbation. Proceedings of *Meetings on Acoustics ICA2013, 19*(1), 060181.

Ohkubo, M., & Scobbie, J. M. (2019). Tongue Shape Dynamics in Swallowing Using Sagittal Ultrasound. *Dysphagia, 34*(1), 112–118. https://doi.org/10.1007/s00455-018-9921-8

Périer, O., Charlier, B., Hage, C., Alegría, J., 1990. Evaluation of the Effects of Prolonged Cued Speech Practice Upon the Reception of Spoken Language. *Cued Speech Journal, IV*, 47-59.

Reidy, P. F., Kristensen, K., Winn, M. B., Litovsky, R. Y., & Edwards, J. R. (2017). The

Acoustics of Word-Initial Fricatives and Their Effect on Word-Level Intelligibility in Children With Bilateral Cochlear Implants. *Ear & Hearing, 38*(1), 42–56. https://doi.org/10.1097/AUD.0000000000000349

Turgeon, C., Trudeau-Fisette, P., Fitzpatrick, E., & Ménard, L. (2017). Vowel intelligibility in children with cochlear implants: An acoustic and articulatory study. *Int. Journn. Ped. Otor., 101,* 87–96.

REMOTE MARKERLESS FACIAL MOTION TRACKING OF MINIMALLY VERBAL CHILDREN WITH AUTISM SPECTRUM DISORDER

Marc Maffei[1], Karen V Chenausky[1], Helen Tager-Flusberg[2], Jordan Green[1,3]
[1]*MGH Institute of Health Professions, Boston, Massachusetts, United States of America*
[2]*Boston University, Boston, Massachusetts, United States of America*
[3]*Harvard University, Boston, Massachusetts, United States of America*

Introduction One in four children diagnosed with autism spectrum disorder (ASD) will remain minimally verbal (MV) at school age (Tager-Flusberg et al., 2013). It has long been suspected that motor speech disorders are present in at least some MV children with ASD (Prizant, 1996; Shriberg et al., 2011) and may contribute to language impairments (Chenausky et al., 2019). However, this hypothesis has rarely been investigated instrumentally. Facial motion tracking yields objective, replicable kinematic data that have been used in pediatric populations to quantify typical development (Nip & Green, 2013) and examine oromotor functioning in disorders including cerebral palsy (Chen et al., 2010). Only two known studies have utilized facial motion tracking with individuals with ASD, one using reflective markers (Gladfelter & Goffman, 2018) and one using markerless technology (Parish-Morris et al., 2018). Markerless tracking has improved significantly over the last decade and offers significant advantages over traditional marker-based technologies – specifically, lower cost, decreased discomfort for subjects, and overcoming COVID-19-related limitations on in-person data collection. However, the implementation of this technology for speech research has been minimal likely due to a lack of research focused on validation and reliability. The purpose of this preliminary study is to assess the feasibility of commercially available markerless tracking software for the evaluation of speech characteristics among MV children with ASD.

Methods Over 100 MV children with ASD completed a remote syllable repetition task over Zoom. Video quality of 20 videos (projected 100 by August 2022) was assessed using a checklist of issues related to technology, environment, and subject placement/behavior. Tracking was achieved using Dynamixyz Performer, a commercially available facial motion capture and 3D processing tool. A custom MATLAB program (Green et al., 2013) was used for post-processing and analysis of lip and jaw movements during productions of /bɑ/. For this abstract we analyzed data from two subjects (age 4:8 and 5:11) (projected 10+ by August 2022) and compared them to existing normative data(10).

Results Jaw (J) movement was subtracted from lower lip (LL) movement. LL and J coupling was examined using zero-lag cross correlation coefficients of displacement signals and revealed an immature coupling pattern consistent with that of typically developing 1-year-olds (Green et al., 2000). The relative contributions of LL and J to combined LL+J movement were calculated using linear regression analysis and revealed a similarly immature movement profile (i.e., LL+J movement was driven almost entirely by the jaw)(10). The most common video quality issues during the syllable repetition task were excessive head movements (75% of videos) and face not in frame for entire task (65%).

Conclusions Preliminary evidence suggests that commercially available markerless facial motion tracking has potential to overcome specific limitations of traditional methods, although markerless approaches are time-consuming and susceptible to data quality issues. Future work should examine the validity of such applications by comparing the performance of markerless approaches to established methods.

Work supported by NIH grants F31DC020108 and P50DC018006.

References

Chen C, Chen H, Hong W, Yang F, Yang L, Wu C. Oromotor variability in children with mild spastic cerebral palsy: a kinematic study of speech motor control. *J NeuroEngineering Rehabil. 2010;7*(1):54. doi:10.1186/1743-0003-7-54

Chenausky K, Brignell A, Morgan A, Tager-Flusberg H. Motor speech impairment predicts expressive language in minimally verbal, but not low verbal, individuals with autism spectrum disorder. *Autism & Developmental Language Impairments. 2019;4*:239694151985633

Gladfelter A, Goffman L. Semantic richness and word learning in children with autism spectrum disorder. *Developmental Science. 2018;21*(2):e12543.

Green JR, Moore CA, Higashikawa M, Steeve RW. The Physiologic Development of Speech Motor Control: Lip and Jaw Coordination.(Brief Article). *Journal of Speech, Language, and Hearing Research. 2000;43*(1):239-255.

Green JR, Wang J, Wilson DL. SMASH: A Tool for Articulatory Data Processing and Analysis. Proceedings of Interspeech 2013.

Nip ISB, Green JR. Increases in Cognitive and Linguistic Processing Primarily Account for Increases in Speaking Rate With Age. Child Dev. 2013;84(4):1324-1337.

Parish-Morris J, Sariyanidi E, Zampella C, et al. Oral-Motor and Lexical Diversity During Naturalistic Conversations in Adults with Autism Spectrum Disorder. *Proceedings of the Conference of the Association for Computational Linguistics North American Chapter Meeting 2018.*

Prizant B. Brief report: Communication, language, social, and emotional development. *Journal of Autism and Developmental Disorders. 1996;26*(2):173-178.

Shriberg L, Paul R, Black L, Santen J. The Hypothesis of Apraxia of Speech in Children with Autism Spectrum Disorder. *Journal of Autism & Developmental Disorders. 2011;41*(4):405-426.

Tager-Flusberg H, Kasari C. Minimally Verbal School-Aged Children with Autism Spectrum Disorder: The Neglected End of the Spectrum. *Autism Research. 2013;6*(6):468-478.

NONWORD REPETITION IN CHILDREN WITH CHILDHOOD APRAXIA OF SPEECH (CAS) AND SPEECH MOTOR DELAY (SMD) – DOES IT REFLECT ORAL MOTOR OR LINGUISTIC DIFFICULTIES

Åsa Mogren[1], Emilia Carlsson[2]
[1]*Karolinska Institutet, Stockholm, Sweden*
[2]*University of Gothenburg, Gothenburg, Sweden*

Reproducing novel words i.e. nonword repetition (NWR), is a well-known and frequently used task in both research and clinical practice. NWR tasks have been used to measure a range of linguistic, speech and motor abilities (Coady and Evans, 2008) and NWR has shown to be a sensitive marker for a range of speech and language disorders (Estes et al., 2007, Coady and Evans, 2008, Conti-Ramsden et al., 2001, Bishop et al., 1996). The overall aim with this study was to investigate NWR in children with Childhood Apraxia of Speech (CAS) and Speech Motor Delay (SMD) in relation to linguistic and oral motor performance and also to compare those results with a comparison group of children with Typical Speech Development (TSD).

The participants included 57 children with SSD aged 6:0-16:7 years (mean age 8:4), 12 girls and 45 boys, 34 children were assessed as having CAS and 23 had SMD. The study also included 39 children with TSD aged 6:0-12:2 years (mean age, 8:8), 18 girls and 21 boys. Differential diagnostics of SSD was made using the operationalised 12 CAS features list by Iuzzini-Seigle & Murray (2017) and Shriberg's classification system (Shriberg et al., 2010). Speech was assessed in children with SSD by phonetic transcription of consonant and vowel production in a word naming test. Parents completed a questionnaire including anamnestic information of pre-linguistic skills, heredity, and reading and writing ability. A non-word repetition task consisting of 18 nonwords with two – four syllables was used. Percentage Consonants Correct (PCC) was calculated for both word naming and NWR. Auditory discrimination was assessed with nine of the included nonwords presented in two pairs. Oral motor performance was assessed using a screening test for orofacial function (NOT-S) (Bakke et al., 2007) consisting of an interview part and an examination part.

The results showed a difference between children with SSD (CAS and SMD) and children with TSD on all anamnestic and assessed variables. Children with CAS had a lower PCC on NWR and a lower result on auditory discrimination task compared to children with SMD and there were also more children with CAS that reported difficulties with reading and writing ability. No difference was found in oral motor performance assessed with NOT-S between children with CAS and SMD. A multiple regression analysis was conducted to further investigate the association between the variables. The results showed that the variable that affected the performance of NWR the most when age was controlled for was number of CAS features from the operationalised list. Number of CAS features and age explained 37.8 % of the variability in NWR. Auditory discrimination performance and age together explained 27.6% and oral motor performance and age explained 24.7% of the variability in NWR. Age alone explained 16% of the variability in NWR.

Results from this study indicates that children with CAS had more severe SSD and more difficulties related to auditory discrimination and reported difficulties with reading and writing than children with SMD. To conclude, assessing both linguistic and oral-motor skills in relation to NWR is important since results indicates that these skills are intertwined.

References

Bakke, M., Bergendal, B., McAllister, A., Sjogreen, L., & Asten, P. (2007). Development and evaluation of a comprehensive screening for orofacial dysfunction. *Swedish Dental Journal, 31*(2), 75-84.

Bishop, D. V. M., North, T. & Donlan, C. (1996). Nonword Repetition as a Behavioural Marker for Inherited Language Impairment: Evidence from a Twin Study. *Journal of Child Psychology and Psychiatry, 37*, 391-403.

Coady, J. & Evans, J (2008) Uses and interpretations of non-word repetition tasks in children with and without specific language impairments (SLI). *International Journal of Language and Communication disorders 43*(1) 1-10.

Conti-Ramsden, G., Botting, N. & Faragher, B. (2001). Psycholinguistic markers for specific language impairment (SLI). *Journal of child psychology and psychiatry, 42*, 741-748.

Estes, K. G., Evans, J. L. & Else-Quest, N. M. (2007). Differences in the nonword repetition performance of children with and without specific language impairment: A meta-analysis.

Iuzzini-Seigel, J., Murray, E. (2017). Speech Assessment in Children with Childhood Apraxia of Speech. *Perspectives of the ASHA Special Interest Groups, 2*, 47-60.

Shriberg, L. D., Fourakis, M., Hall, S. D., Karlsson, H. B., Lohmeier, H. L., McSweeny, J. L., Wilson, D. L. (2010). Extensions to the Speech Disorders Classification System (SDCS). *Clinical Linguistics & Phonetics, 24*(10), 795-824.

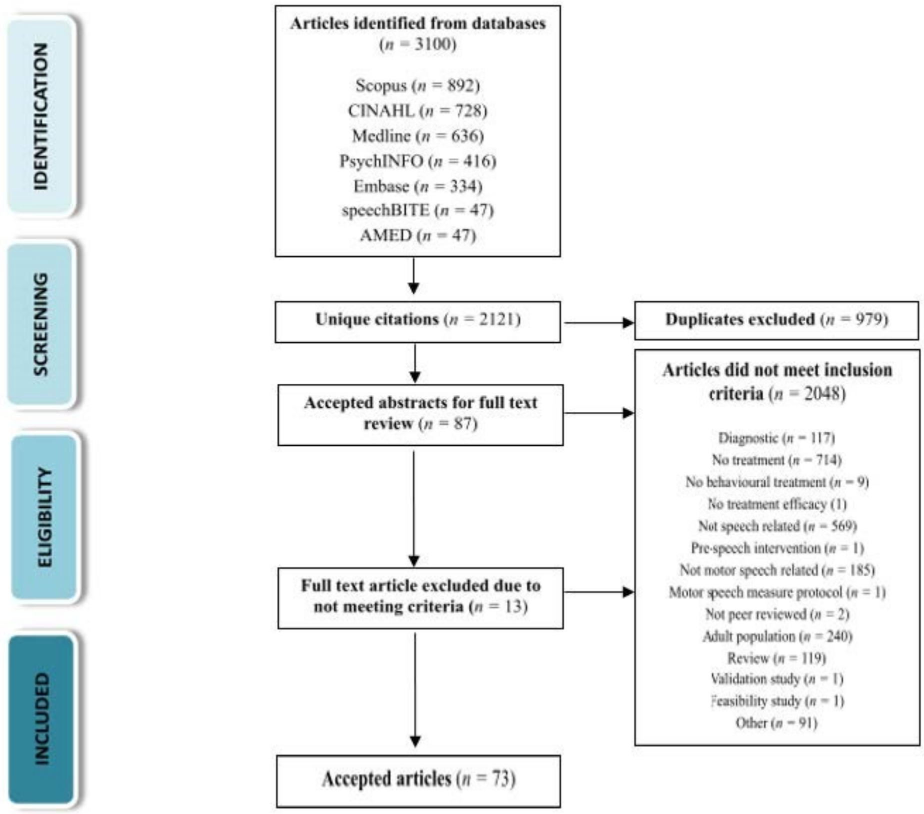
IMPLEMENTING MOTOR SPEECH OUTCOME MEASUREMENT IN PRESCHOOL SPEECH-LANGUAGE PROGRAMS

Aravind K. Namasivayam[1], Elissa Flagg[2] , Margit Pukonen[2] , Nilgoun Bahar[3], Bavika Atputhajeyam[1], Jatin Gill[1], Pascal van Lieshout[1]
[1]*University of Toronto, Toronto, Canada*
[2]*The Speech and Stuttering Institute, Toronto, Canada*
[3]*University of Oxford, Oxfordshire, United Kingdom*

The Ministry of Children, Community and Social Services (MCCSS) in Ontario (Canada) has engaged The Speech & Stuttering Institute (Toronto) in a project on outcome measures for children receiving motor speech therapy in the preschool speech and language (PSL) program. This project is conducted in collaboration with the Speech Research Centre Inc. and the University of Toronto. It is part of a larger scale focus on quality assurance in Ministry-funded programs serving special needs children in Ontario, Canada. The purpose of this study is to identify motor speech outcome measures that are feasible, sustainable, practical and relevant for the PSL context.

For the first phase of this project, a literature search (scoping review) was conducted with the aim of identifying candidate outcome measures (both standardized and non-standardized) with good psychometric properties that are sensitive to measure change following intervention. Our analyses were guided by the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA; See Figure 1) framework (Page et al., 2021) to ensure adherence to quality review protocols. Seven databases (e.g., Scopus, CINAHL, Medline) were searched for journal articles published between January 2015 and June 15th, 2021, to identify intervention studies in children with speech sound disorders. A total of 73 articles were included for final analysis from 3100 that were initially identified. Following this we consolidated data from Kearney et al., 2015 (similar review of articles published between 1985 to Dec 31st, 2014) and the current study (Jan 2015 to June 15th, 2021) to extract 57 pediatric motor speech outcome measures. These measures were mostly perceptual transcription and/or perceptual ratings with only a small number of physiological or acoustic measures.

In the next phase we are engaging clinicians as key stakeholders to help us understand current practices with respect to outcome measurement, and to provide input on what a feasible/practical plan for outcome measurement would look like in clinical settings. Our goal is to narrow down the measures needed to those that are easiest and most efficient to obtain, but still sensitive to changes in speech production skills. That way, outcome measurement can eventually be implemented system wide as a means of demonstrating the positive impact of PSL services on clients.



Screening and review process.

References

Kearney, E., Granata, F., Yunusova, Y., Van Lieshout, P., Hayden, D., & Namasivayam, A. (2015). Outcome measures in developmental speech sound disorders with a motor basis. *Current Developmental Disorders Reports, 2*(3), 253-272.

Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., ... & Moher, D. (2021). The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. *International Journal of Surgery, 88*, 105906

LONGITUDINAL NORMATIVE DATA ON DEVELOPMENTAL SPEECH ERRORS IN FRENCH-SPEAKING PRESCHOOLERS, THE AVERAGE PERCENTAGE OF OCCURRENCES OF PHONOLOGICAL PROCESSES

Léonor Piron[1], Andrea MacLeod[2], Morgane Warnier[1], Christelle Maillart[1]
[1] *University of Liège, Liège, Belgium*
[2] *University of Alberta, Edmonton, Canada*

Introduction Phonological process (PPs) are considered as developmental errors (Rvachew & Brosseau-Lapré, 2018). Assessment of PPs remains one of the most frequent in-depth analyses of phonology (Brosseau-Lapré et al., 2018; Macrae, 2016; Kirk & Vigeland, 2015). PPs vary from one language to another (Brosseau-Lapré et al., 2018; Hayes, 2009). Each SLP should therefore be as familiar as possible with the PPs of their language. However, longitudinal developmental data in the French-speaking area of PPs are scarce despite the fact that French is currently the 5th largest language in the world (Semo, 2018). In particular, it would be interesting to determine the average percentage of occurrences (POC) of the different PPs in French, from longitudinal data. Average POC will make it easier for French-speaking SLP to categorize a patient's PPs occurrence and will improve their qualitative phonological assessment (Macrae, 2016; Kirk & Vigeland, 2015). Such an analyze will also meet the cross-linguistic research movement on phonological development (Petinou & Armostis, 2017; Keren-Portnoy et al., 2003).

Goals Determining in French-speaking preschoolers' speech:
1. changes over time in the average POC of segmental and suprasegmental PPs
2. normative data on the average POC of PPs

Methods 29 French-speaking typically developing children were followed longitudinally on their phonological productions from 3 to 5 years old. We have collected speech samples through the shorten version of the Eulalies test (Meloni et al., 2017; Warnier et al., in press). Only typically developing children were selected after a thorough control of their audition, IQ, lexical level and history. Analyses were carried out at ages 3, 4 and 5. The analyzes were carried out using the Phon software Rose et al., 2006). The average reliability between transcribers for all evaluation times was excellent (90,93%). The POC of PPs was calculated following the methodology of similar studies and of recommended assessment methods (Franklin & McDaniel, 2016; Olswang et al., 1987):
1. For each PP and each child
2. calculation of the mean and standard deviation of the 29 children's POC of the PP

The changes over time were measured using Friedman ANOVAs for non-parametric repeated measures. A PP was considered frequent when its average POC was above or equal to 5%.

Results PPs globally decrease in occurrence over time. At age 3, devoicing, fronting, cluster reduction and final deletion are frequent PPs. At age 4, devoicing and cluster reduction are still frequent. At age 5, only cluster reduction remains a frequent PP. Means, Standard Deviation of POC and Friedman Anova results are shown in Table 1.

Discussion Our results on PPs are consistent with similar findings in French [2,15] and in other languages (Petinou & Armostis, 2017; Keren-Portnoy et al., 2003; Cohen & Anderson, 2011). Cluster reduction decreases more slowly than other PPs. A possible explanation is that French clusters are acquired slower than singletons (MacLeod et al., 2011). Devoicing is more frequent than fronting, contrary to the results of Brosseau-Lapré et al. (2018). In comparison with English (Rvachew & Brosseau-Lapré, 2018; Cohen & Anderson, 2011), devoicing is a frequent PP at ages 3 and 4. On the opposite, onset consonant reduction, gliding, stopping and syllable deletion are no frequent PPs at any age.

Conclusion Our longitudinal normative data developed with the average POC provide French-speaking SLP with further landmarks and norms on PPs.

Average POC of segmental and suprasegmental PPs and results of Friedman ANOVAs

		3 Years old		4 Years old		5 Years old		Friedman Anovas for non-parametric repeated measures		
		Mean	*SD*	*Mean*	*SD*	*Mean*	*SD*	*X2*	*Df*	*p*
Segmental	*Devoicing*	10,57	7,83	6,22	4,39	4,54	2,95	17,2	2	<0,001
	Fronting	8,185	6,27	3,420	4,31	1,06	2,26	38,8	2	<0,001
	Stopping	3,63	6,41	1,317	3,25	0,48	1,52	18	2	<0,001
	Backing	3,1	3,04	1,179	1,38	0,49	0,69	24,5	2	<0,001
	Gliding	2,16	3,35	0,11	0,6	0	0	21,8	2	<0,001
Suprasegmental	*Cluster Deletion*	34,48	25,64	22,88	19,38	13,48	8,97	10,7	2	0,005
	Final Deletion	9,42	9,65	4,19	5,32	1,63	3,33	22,3	2	<0,001
	Phoneme Deletion	6	5,55	2,42	3,24	0,56	0,704	28,2	2	<0,001
	Syllable Deletion	4,94	7,04	1,5	2,21	0,36	0,69	33,8	2	<0,001
	Onset Deletion	4,12	5,05	2,74	4,14	0,77	1,82	14,1	2	<0,001

References

Brosseau-Lapré, F., Rvachew, S., Macleod, A. A. N., Findlay, K., Bérubé, D., Bernhardt, B., & Findlay, K. (2018). Une vue d'ensemble : les données probantes sur le développement phonologique des enfants francophones canadiens. *Revue Canadienne d'orthophonie et d'audiologie, 42*(1), 1–19.

Cohen, W., & Anderson, C. (2011). Identification of phonological processes in preschool children's single-word productions. *International Journal of Language and Communication Disorders, 46*(4), 481–488. https://doi.org/10.1111/j.1460-6984.2011.00011.x

Franklin, A., & McDaniel, L. (2016). The Development of English as a Second Language With and Without Specific Language Impairment: Clinical Implications. *American Journal of Speech-Language Pathology, 25*(2), 172–182. https://doi.org/10.1044/2015_AJSLP-14-0172

Hayes, B. (2009). *Introductory phonology.* Wiley-Blackwell.

Keren-Portnoy T, Majorano M, Vihman MM. (2003) From phonetics to phonology: the emergence of first words in Italian. *J Child Lang, 36*, 235–267.

Kirk, C., & Vigeland, L. (2015). Content coverage of single-word tests used to assess common phonological error patterns. *Language, Speech, and Hearing Services in Schools, 46*, 14–29. https://doi.org/10.1044/2014_LSHSS-13-0054

MacLeod, A. A. N., Sutton, A., Trudeau, N., & Thordardottir, E. (2011). The acquisition of consonants in Québécois French: A cross-sectional study of pre-school aged children. *International Journal of Speech-Language Pathology, 13*(2), 93–109.

Macrae, T. (2016). Comprehensive Assessment of Speech Sound Production in Preschool Children. *Perspectives of the ASHA Special Interest Groups, 1*(1), 39–56. https://doi.org/10.1044/persp1.sig1.39

Meloni G., Loevenbruck H., Vilain A., Macleod A. A. N. (2017, July 17-21). *EULALIES, The France-Québec speech sound disorders project* [Poster presentation]. IASCL 14th international congress, Lyon, France.

Olswang LB, Stoel-Gammon C, Goggins E, Carpenter RL. (1987). *Assessing Pre-Linguistic and Early Speech Skills in Developmentally Young Children.* Seattle, University of Washington Press.

Petinou, K., & Armostis, S. (2017). Phonological Process Occurrence in Typically Developing Toddlers. *Folia Phoniatrica et Logopaedica, 68*(5), 199–204. https://doi.org/10.1159/000454950

Rose, Y., MacWhinney, B., Byrne, R., Hedlund, G., Maddocks, K., O'Brien, P., & Wareham, T. (2006). Introducing Phon: A Software Solution for the Study of Phonological Acquisition. *Proceedings of the Annual Boston University Conference on Language Development*, 489–500.

Rvachew, S., & Brosseau-Lapré, F. (2018). Developmental phonological disorders: foundations of clinical practice (Second edition). Plural Publishing, Inc.

Schelstraete, M. A., Maillart, C., & Jamart, A.-C. (2004). Les troubles phonologiques : cadre théorique, diagnostic et traitement. *Les Troubles Du Langage et Du Calcul Chez l'enfant*, 81–112.

Semo, M. (2018, october 11). Le français, cinquième langue la plus parlée dans le monde. Le Monde.

Warnier, M., Maillart, C., Rose, Y., & MacLeod, A. (In press). Exploring word production in three-year-old monolingual French-speaking children. *Clinical Linguistics and Phonetics.*

THE EFFECT OF LEXICAL STATUS AND VOCABULARY ON COARTICULATION.

Tom Starr-Marshall[1], Joanne Cleland[1], Claire Timmins[1], James M. Scobbie[2]
[1]*University of Strathclyde, Glasgow, United Kingdom*
[2]*Queen Margaret University, Edinburgh, United Kingdom*

Background This study aims to explore whether the size of unit of planning in children's speech is influenced by the lexical status of the words and receptive vocabulary. It also explores whether different measures of coarticulation, give different answers to these questions. Previous research suggests that lower expressive vocabulary points to unstable representations at a phonemic level. (Macrae and Sosa, 2015). When representations are unstable or under-specified, words are planned as larger units. Increasing specification of representations with maturation leads to planning with smaller units. This progression from larger to smaller units of planning can be measured by investigating coarticulation. A larger degree of coarticulation signifies a larger unit of planning, whereas less coarticulation signifies a smaller unit of planning (Noiray et al., 2018).

Research Questions

- What is the relationship between coarticulation and vocabulary?
- What effect does lexical status (i.e. real or nonsense word) have on coarticulation?
- How does lexical status and vocabulary interact with coarticulation?

Methods This study uses a subset of data from the Ultrasuite corpus of ultrasound tongue imaging and audio data comprising 30 typical developing Scottish English speaking children aged 5;8 years to 12;10 years (Eshky et al., 2019).

Forty-Five nonsense words were selected from this dataset. All were VCV in structure, containing most consonants of English in a symmetrical vowel context between the Scottish English corner vowels /a/, /o/, and /i/. Twenty-Three real words CVC in structure were also selected from this data set. These words were a mix of Diagnostic Evaluation of Articulation and phonology (Dodd et al., 2002) and analytically important words (Eshky, et al., 2019), and also used the Scottish English corner vowels /a/, /o/, and /i/. Segments for analysis were highlighted and annotated in Articulate Assistant Advanced (AAA) (Articulate Instruments Ltd, 2012) based on the acoustic signal. The onset of the consonant was defined as the end of the periodic signal of the preceding vowel; the end was defined as the end of the aperiodic noise of the consonant. Consonant midpoints were exported automatically. These were batch splined and trimmed to the hyoid and mandible shadows. Measures of Dorsum Excursion Index (Zharkova, 2013), Tongue Constraint Position Index (Zharkova, 2013), LOC_a-i (Zharkova et al., 2015), and Highest Point of the tongue (Noiray et al., 2018) were exported. These were compared with lexical status (real or nonsense word), age, and receptive vocabulary raw scores from the The British Picture Vocabulary Scale (BPVS) (Dunn et al.,1997) using a multilevel model.

Results Data analysis is ongoing. Initial results suggest a greater degree of coarticulation in real words than in nonsense words. Which suggests children plan real words using larger units. Further data will be presented.

References

Dodd, B. and Zhu, Hua and Crosbie, S. and Holm, A. and Ozanne, A. (2002) *Diagnostic evaluation of articulation and phonology (DEAP).* London: Psychology Corporation.

Dunn, L. M., Dunn, L. M., Whetton, C., & Burley, J. (1997). *British Picture Vocabulary Scale (BPVS-II)*, 2nd ed. NFER-Nelson Publishing Company.

Eshky, A., Ribeiro, M. S., Cleland, J., Richmond, K., Roxburgh, Z., Scobbie, J., & Wrench, A. (2018). UltraSuite: A repository of ultrasound and acoustic data from child speech therapy sessions. In *Proceedings of the Annual Conference of the International Speech Communication Association, INTERSPEECH* (Vol. 2018-September).

Macrae, T., & Sosa, A. V. (2015). Predictors of token-to-token inconsistency in preschool children with typical speech-language development. *Clinical Linguistics and Phonetics, 29*(12).

Noiray, A., Abakarova, D., Rubertus, E., Krüger, S., & Tiede, M. (2018). How do children organize their speech in the first years of life? Insight from ultrasound imaging. *Journal of Speech, Language, and Hearing Research, 61*(6).

Zharkova, N. (2013). Using ultrasound to quantify tongue shape and movement characteristics. *The Cleft Palate Craniofacial Journal, 50*, 76–81.

Zharkova, N., Gibbon, F. E., & Hardcastle, W. J. (2015). Quantifying lingual coarticulation using ultrasound imaging data collected with and without head stabilisation. *Clinical Linguistics and Phonetics, 29*(4).

TREATMENT OF CHILDHOOD APRAXIA OF SPEECH WITH SPEECH AND MUSIC THERAPY

Mirjam van Tellingen[1], Joost Hurkmans[1], Hayo Terband[2], Anne Marie van de Zande[3], Ben Maassen[4], Roel Jonkers[4]
[1]Rehabilitation Center "Revalidatie Friesland", Beetsterzwaag, The Netherlands
[2]University of Iowa, Iowa City, Iowa, United States of America
[3]Rehabilitation Center "Rijndam Revalidatie", Rotterdam, The Netherlands
[4]University of Groningen, Groningen, The Netherlands

Background A combination of speech therapy and music therapy, Speech-Music Therapy for Aphasia (SMTA; De Bruijn et al., 2005; Hurkmans et al., 2018), is used in the treatment of children with Childhood Apraxia of Speech (CAS; Van Tellingen et al., 2022). SMTA has been shown to impact speech production at the level of speech motor planning and programming in five adults with apraxia of speech (AoS; Hurkmans et al., 2015). Studies on musical elements in the treatment of children with speech sounds disorders are limited and while these studies mostly report positive outcomes, methodological quality is insufficient (Van Tellingen et al., accepted).

The use of music in the treatment of speech disorders is based on (1) similarities between and overlap in the processing of language and music (Patel, 2011), (2) overlap in prosodic features in music and speech (Hurkmans, 2016; Terband et al., 2019), (3) principles of motor learning (Maas et al., 2014; Wulf et al., 2018) and (4) mechanisms of music with regards to motivation and mood (Merrett et al., 2014).

SMTA combines speech therapy, which includes three levels (1) speech sounds (including syllables), (2) words and (3) sentences with music therapy, which follows a structured procedure in which musical support is phased out from singing to rhythmical chanting and speaking. The musical interventions are designed to musically support the speech exercises using melody, rhythm, meter, tempo and dynamics. The aim of the current study is to determine whether SMTA can be effective for children suffering CAS, based on the similarities of AoS and CAS, both being considered a disorder in the planning and programming of speech movements (American Speech-Language-Hearing Association, 2007a; Hurkmans, 2016).

Method The current single subject design study is part of a study that will include five children with (suspected) CAS in a pilot study. The current participant is a 5 year old boy with CAS.

Design: Single subject design with pretest, baseline, treatment, posttest and follow-up.

Treatment: 20 sessions of SMTA, following standard treatment protocol, consisting of two 30-minute therapy sessions per week and practice at home with recordings.

Outcome measures: The following assessments were used to measure changes in functioning at different linguistic levels, including activity and participation: Intelligibility in Context Scale – Dutch (ICS-Dutch; McLeod et al., 2012); Computer Articulation Instrument (CAI; Maassen et al., 2019); Modified Diadochokineses Test (MDT; Hurkmans et al., 2012); Phonological Analysis of Dutch (FAN; Beers, 1995); a personalized imitation task with trained and related untrained items.

Results Results for the 5 year old boy point show improved intelligibility, improved production of consonant clusters and increased consistency.

Conclusion This boy with CAS had improved intelligibility after treatment with SMTA. Results showed transfer from trained items to untrained speech tasks, including tasks at the level of speech motor planning and programming. The results of the first study in this pilot are promising and provide leads for the further study of SMTA in the treatment of CAS.

References

American Speech-Language-Hearing Association. (2007a). Childhood apraxia of speech [Technical report]. Available from www.asha.org/policy. *American Speech-Language-Hearing Association.*

Beers, M. (1995). *Fonologische Analyse Nederlands (FAN).* IFOTT.

De Bruijn, M., Zielman, T., & Hurkmans, J. J. S. (2005). Speech-Music Therapy for Aphasia (SMTA). Revalidatie Friesland.

Hurkmans, De Bruijn, M., Reitsma, T., & Koek, P. (2018). Speech-Music Therapy for Aphasia (E-learning). Revalidatie Friesland.

Hurkmans, J. J. S. (2016). *The treatment of apraxia of speech.* Rijksuniversiteit Groningen.

Hurkmans, J., Jonkers, R., Boonstra, A. M., Stewart, R. E., & Reinders-Messelink, H. A. (2012). Assessing the treatment effects in apraxia of speech: Introduction and evaluation of the Modified Diadochokinesis Test. *International Journal of Language & Communication Disorders, 47*(4), 427–436. https://doi.org/10.1111/j.1460-6984.2012.00155.x

Hurkmans, Jonkers, R., Bruijn, M. de, Boonstra, A. M., Hartman, P. P., Arendzen, H., & Reinders-Messelink, H. A. (2015). The effectiveness of Speech–Music Therapy for Aphasia (SMTA) in five speakers with Apraxia of Speech and aphasia. *Aphasiology, 29*(8), 939–964. https://doi.org/10.1080/02687038.2015.1006565

Maas, E., Gildersleeve-Neumann, C. E., Jakielski, K. J., & Stoeckel, R. (2014). Motor-Based Intervention Protocols in Treatment of Childhood Apraxia of Speech (CAS). *Current Developmental Disorders Reports, 1*(3), 197–206. https://doi.org/10.1007/s40474-014-0016-4

Maassen, B., Van Haaften, L., Diepeveen, S., Terband, H., Van den Engel-Hoek, L., Veenker, T., & De Swart, B. (2019). *Computer Articulatie Instrument.* Boom Uitgevers.

McLeod, S., Harrison, L. J., & McCormack, J. (2012). *Schaal voor Verstaanbaarheid in de Context* [Intelligibility in Context Scale: Dutch]. (J.C. van Doornik-van der Zee & H.R. Terband, Trans). Charles Sturt University. Retrieved from http://www.csu.edu.au/research/multilingual-speech/ics

Merrett, D. L., Peretz, I., & Wilson, S. J. (2014). Neurobiological, cognitive, and emotional mechanisms in melodic intonation therapy. *Front Hum Neurosci, 8*(JUNE), Article JUNE. https://doi.org/10.3389/fnhum.2014.00401

Patel, A. (2011). Why would Musical Training Benefit the Neural Encoding of Speech? The OPERA Hypothesis. *Frontiers in Psychology, 2.* https://www.frontiersin.org/article/10.3389/fpsyg.2011.00142

Terband, Namasivayam, Maas, van Brenk, Mailend Marja-Liisa, Diepeveen, van Lieshout, & Maassen. (2019). Assessment of Childhood Apraxia of Speech: A Review/Tutorial of Objective Measurement Techniques. *Journal of Speech, Language, and Hearing Research, 62*(8S), 2999–3032. https://doi.org/10.1044/2019_JSLHR-S-CSMC7-19-0214

Van Tellingen, M., Hurkmans, J., Terband, H., Jonkers, R., & Maassen, B. (Accepted).

Music and musical elements in the treatment of childhood speech sound disorders: A systematic review of the literature. *International Journal of Speech-Language Pathology.*

Van Tellingen, M., Hurkmans, J., Van de Zande, A. M., Terband, H., Meinsma - van der Tuin, M., Jonkers, R., & Maassen, B. (2022). *Speech and music therapy in the treatment of Childhood Apraxia of Speech: A case study* [Poster]. European Academy of Childhood Disability, Barcelona.

Wulf, G., Lewthwaite, R., Cardozo, P., & Chiviacowsky, S. (2018). Triple play: Additive contributions of enhanced expectancies, autonomy support, and external attentional focus to motor learning. *Quarterly Journal of Experimental Psychology, 71*(4), 824–831. https://doi.org/10.1080/17470218.2016.1276204

THE USE OF OBJECTIVE ARTICULATORY KINEMATIC MEASURES TO SUPPORT CLINICAL DECISION MAKING IN THE DIAGNOSIS OF MOTOR SPEECH DISORDERS: A PILOT STUDY

Roslyn Ward[1], Richard Palmer[1], Neville Hennessey[1], Linda Orton[1], Paul Davey, Geoff Strauss[1], Petra Helmholz[1], Deborah Hayden[2], Aravind K. Namasivayam[3]
[1]Curtin University, Bentley, Australia, [2]PROMPT Institute, Santa Fe, New Mexico, United States of America
[3]University of Toronto, Canada

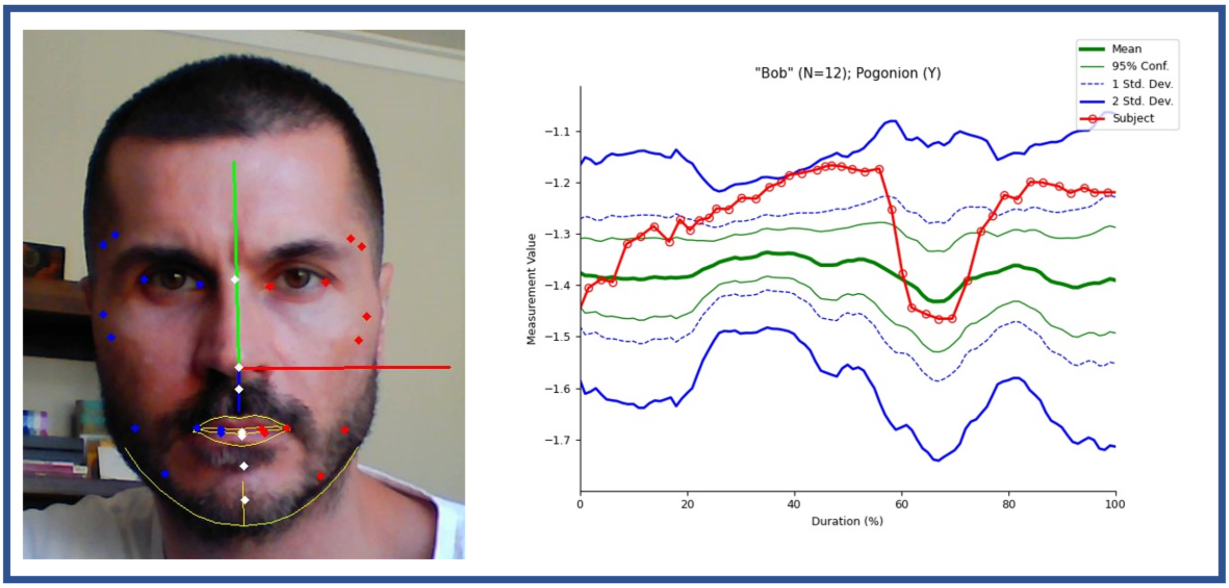
Introduction Differential diagnosis of motor-speech disorders (MSD) in children is difficult (Iuzzini-Seigel et al., 2022). Accurate, automated assessment instruments could assist clinicians' identification of typical or atypical speech movements, thereby supporting diagnostic decision making. This study reports on the accuracy of speech-language pathologists (S-LPs) in scoring jaw movements, comparing traditional scoring with scoring that is assisted through objective kinematic measurements, relative to expert opinion. Two questions were posed: 1. Is there a difference in percentage agreement between traditional and assisted scoring methods, relative to expert opinion? 2. Is there a relationship between perceived usefulness of the kinematic measurements and percentage agreement with the expert opinion?

Method Computer vision techniques were utilised to localise the positions of standard anthropometric clinical landmarks and derive objective 3D spatial measurements, obtained from video-recordings of children completing the Motor Speech Hierarchy Probe Wordlist (MSH-PW; Namasivayam et al., 2021). Spatiotemporal charts, characterising measurements of interest in the MSH-PW scoring system stage III mandibular control (e.g., range of jaw opening, jaw control) were then generated and visually displayed to clinicians to support their clinical decision making.

A randomised two-arm AB/BA cross-over study design was utilised, entailing scoring condition (A = traditional, B = assisted) and child case study diagnosis: typical development (TD) and motor speech disorder (MSD) across three sessions. Percentage agreement with expert opinion across jaw range and jaw control were calculated. Based on Viera and Garrett (Viera & Garrett, 2005) scores $> 64\%$ were considered "substantial" agreement. Spearman's rank correlations were computed to assess the relationship between perceived assistance (binary scale) of the kinematic graphs and percentage agreement with the expert.

Results Fourteen S-LPs participated, with experience in the assessment of MSD ranging from new graduates to 10 years (median 7 years). For TD cases, mean percentage agreement with expert opinion in traditional and assistive conditions, respectively, was 60% (SD 13%) and 59% (SD 9%), for jaw range; and 62% (10%) and 64% (13%), for jaw control. Substantial agreement with the expert opinion was achieved by 64% of clinicians for at least one TD case. For MSD cases, mean percentage agreement with expert opinion in traditional and assistive conditions, respectively was 64% (SD 21%) and 74% (SD 13%) for jaw range; and 71% (SD 21%) and 69% (SD 27%) for jaw control. Substantial agreement (i.e., scored >64% agreement) with the expert opinion was obtained by 86% of S-LPs in the assistive condition, and 64% in the traditional condition. A strong, positive correlation between perceived usefulness of the visually displayed kinematic data (spatiotemporal charts) and mean percentage rating with expert opinion was demonstrated for jaw range but not control in case study TD: jaw range r = .74 (p =.006); and MSD: jaw range r = .80 (p =.002).

Discussion These early pilot data suggest spatiotemporal kinematic charts could facilitate clinicians in their identification of atypical jaw movements in speech; and useful in supporting the diagnosis of a SSD. The data from this study are being used to inform further research focused on improving measurement accuracy of the articulatory kinematic charts and training material.



Example of facial landmarks (left) and articulatory kinematic chart illustrating jaw opening in the word "bob" (right).

References

Iuzzini-Seigel, J., Allison, K., & Stoeckel, R. (2022, early release). A tool for the differential diagnosis of childhood apraxia of speech and dysarthria in children: A tutorial. *American Journal of Speech - Language Pathology.*

Namasivayam, A. K., Huynh, A., Bali, R., Granata, F., Law, V., Rampersaud, D., ... & Hayden, D. (2021). Development and Validation of a Probe Word List to Assess Speech Motor Skills in Children. *American Journal of Speech-Language Pathology, 30*(2), 622-648.

Viera, A. J., & Garrett, J. M. (2005). Understanding interobserver agreement: the kappa statistic. *Fam med, 37*(5), 360-363.

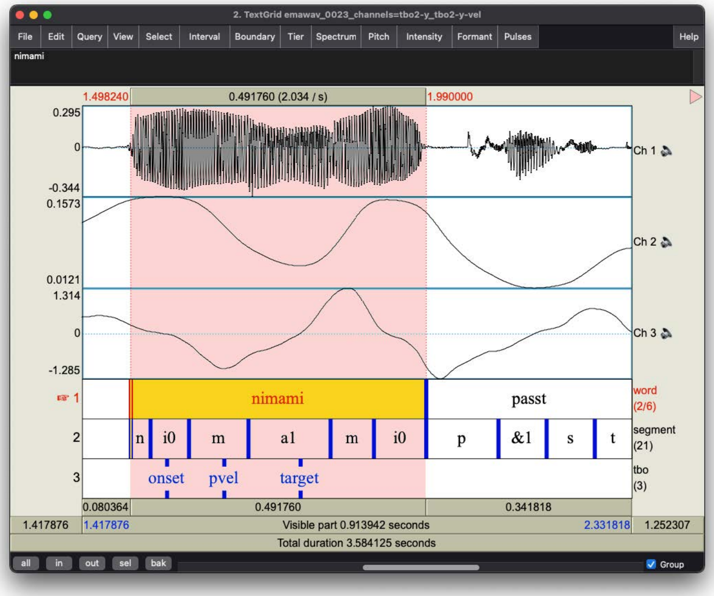
MAKING ARTICULATION ACCESSIBLE IN PRAAT

Philipp Buech[1],Simon Roessig[2], Lena Pagel[2], Anne Hermes[1]
[1]*CNRS/Sorbonne Nouvelle, Paris, France*
[2]*University of Cologne, Cologne, Germany*

Data from electromagnetic articulography (EMA) requires particular software for its post-processing, its display as well as its annotation. Users have the choice between different software packages (e.g., MVIEW (Tiede, 2005), VISARTICO (Ouni et al., 2012), EMUR (Winkelmann et al., 2021)), but their use may be limited since they require to purchase licenses, are no longer maintained or are difficult to access for non-technical users. PRAAT (Boersma & Weenink, 2022) on the other hand is widespread in phonetics and speech science for the display, annotation and analysis of acoustic speech signals as it is freely available, well maintained, easy to use, and available on different operating systems (e.g., Mac, Windows, Linux). Furthermore, it provides interfaces with other environments (e.g., R (R Core Team, 2020), PYTHON (Van Rossum & Drake, 2009)). Despite its large community of users from multiple areas in speech science, especially phonetics, there is no way to display and analyze EMA data in PRAAT.

We present EMA2WAV, a lightweight open-source software for the post-processing and the conversion of EMA data to multi-channel audio files to work within other programs. Primarily intended for PRAAT, but not limited to it, this conversion allows the reduction of the number of software packages in a workflow and further makes the data easily accessible to non-technical users. The conversion tool is implemented in PYTHON and it provides (i) smoothing methods (e.g., moving average filter, butter low pass filter) and (ii) calculations/processing (e.g., 1st and 2nd derivative, tangential velocity). The EMA tracks can be exported as multi-channel WAV files (with or without the audio signal) and/or CSV files (one column per channel). The resulting WAV file can thus be used for the workflow in PRAAT. If the EMA data is stored alongside the audio signal, the first channel represents the audio data and the other channels the desired EMA tracks. Figure 1 provides an example for a display of the waveform, the tongue body position and velocity, and textual annotations. EMA2WAV provides user-friendly post-processing and conversion of EMA data into a common format – granting the advantages of free programs such as PRAAT, which are easily accessible and include many experienced users making articulatory data available for everyone.

Up to now, this converter is written for articulatory data collected with the electromagnetic articulograph models AG500 and AG501 of Carstens Medizinelektronik GmbH. We are planning to extend the converter to older models of Carsten's EMA devices (AG100, AG200). We aim to specifically target the speech motor control community to get feedback about the current version of the converter, to discuss its advantages and shortcomings, and to collect ideas on additional features for a smooth and fluent annotation procedure.



Screenshot of sample EMA data in PRAAT (tongue body y-position and velocity) along with annotations in a TextGrid.

References

Boersma, P., & Weenink, D. (2022). Praat: Doing phonetics by computer [Computer program].

Ouni, S., Mangeonjean, L., & Steiner, I. (2012). VisArtico: A visualization tool for articulatory data.

R Core Team. (2020). R: A language and environment for statistical computing. Vienna, Austria, R Foundation for Statistical Computing.

Tiede, M. (2005). MVIEW: Software for visualization and analysis of currently recorded movement data. Haskins Laboratories.

Van Rossum, G., & Drake, F. L. (2009). Python 3 reference manual. *CreateSpace*.

Winkelmann, R., Jaensch, K., Cassidy, S., & Harrington, J. (2021). EmuR: Main Package of the EMU Speech Database Management System.

THE COLLABOVERSE: A COLLABORATIVE DATA-SHARING AND SPEECH-ANALYSIS PLATFORM

Justin D. Dvorak[1,2], Frank R. Boutsen[1,3]
[1] *Communication and Audio Technology Laboratory (CATLAB)*
[2] *The University of Oklahoma Health Sciences Center, Oklahoma City, Oklahoma, United States of America*
[3] *New Mexico State University, Las Cruces, New Mexico, United States of America*

Introduction Collaboration among speech-language researchers is often thwarted by difficulties in synchronizing data and analysis workflows. Current methods tend to be limited to data sharing, using techniques such as email, (S)FTP, network drives, cloud storage, and even portable devices. Critically, none of these methods independently support synchronization and collaboration across all stages of a project, including study design, data collection & management, signal processing, acoustic analysis, statistical modeling, and dissemination of results. An additional burden to collaboration is the need to transform data and conduct handshaking among the go-to acoustic tools, programming languages, machine-learning frameworks, and computing environments, such as Praat (Boersma & Weenink, 2022) and Praat scripts, MATLAB (The Mathworks, 2022), R (R Core Team, 2022), and Python (Python Software Foundation, 2022).

Solutions to this problem are out of reach for most researchers, due to lack of expertise in coding and data manipulation skills. Finally, transferability and reproducibility of analysis workflows, whether manual or automated, can be limited because of differences in platform (e.g., PC vs Mac) and configuration (e.g., installed packages, library dependencies, system paths), not to mention competencies among lab members (e.g., executing code in an IDE). In this presentation, we propose a solution to the above problem, which allows synchronization of data and workflows, and collaborative analysis through a simple, web-based graphical user interface (GUI). While streamlining analysis for point and click usage, our solution also allows advanced users to develop their own modules and analysis workflows.

Methods Common procedures in the laboratory (such as stimulus generation, signal processing, and statistical analyses) are implemented in the front-end interface and back-end module architecture. Data uploading into the Collaboverse is handled through a secure WebSocket interface, with users' files stored on an encrypted server. A server interface is implemented in NodeJS (OpenJS Foundation, 2022), which communicates object-state updates with the user-side JavaScript web client using JSON, and handles interfacing with back-end analytics and compute resources. Analytic functionality within the Collaboverse is implemented as a module system, optionally using independent and parallel hardware. Modules interface with the ecosystem as virtual users in a non-blocking flow-control paradigm, so an error in one module does not jeopardize other ongoing operations. GUI component management (state updates and callbacks) is handled through CSS-styled elements and canvas objects, which can be remotely updated according to each module's requirements (e.g., displaying updates in an analysis, or visualizing results in a custom figure).

Results We present a live demonstration of the capabilities of the Collaboverse. Audience members are shown how to import and manage speech data, conduct selected acoustic analyses, export results, assemble toolchains of modules, and automate workflows in a collaborative environment, presented via the metaphor of a virtual shared desktop and whiteboard.

References

Boersma, P., & Weenink, D. (2022). Praat: doing phonetics by computer [Computer program]. Version 6.2.10, retrieved 17 March 2022 from http://www.praat.org/

The Mathworks (2022). MATLAB [Computer program]. Version 9.12.0 (R2022a), retrieved from https://www.mathworks.com/products/matlab.html

R Core Team (2022). R: A Language and Environment for Statistical Computing. Vienna, Austria [Computer program]. Version 4.1.3, retrieved from https://www.R-project.org/

Python Software Foundation (2022). Python [Programming language]. Version 3.10.3, retrieved from https://www.python.org/

OpenJS Foundation (2022). Node.js [Computer program]. Version 17.5.0, retrieved from https://nodejs.org/en/.

ON TIMING AND PRONUNCIATION METRICS FOR INTELLIGIBILITY ASSESSMENT IN PATHOLOGICAL ALS SPEECH

Jackson Liscombe, Michael Neumann, Hardik Kothare, Oliver Roesler, David Suendermann-Oeft, Vikram Ramanarayanan
[1]*Modality.AI, Inc., San Francisco, California, United States of America*

We investigate three speech metrics – goodness of pronunciation (GoP), percent pause time (PPT), and a new measure of canonical timing alignment (CTA) – with respect to how well they characterize the temporal and spectro-acoustic aspects of pathological speech intelligibility in people with amyotrophic lateral sclerosis (pALS). We selected 2174 read speech utterances of varying lengths spoken by 40 distinct participants (comprising both pALS and healthy controls) from a large corpus of speech and video data collected remotely via an interactive dialog agent. In order to capture primarily timing-related aspects of intelligibility, we propose and compute the CTA feature, which is a measure of how similar the temporal sequencing of forced alignments of words is between a given speaker and a canonical healthy speaker. We also computed PPT, which captures the pausing duration relative to the speaking duration, and GoP, a measure that captures both spectral and temporal information and has been used widely in the literature to capture intelligibility. We find that in general, CTA is as, if not more, informative than GoP and PPT in distinguishing controls from bulbar pre-symptomatic and bulbar symptomatic ALS patients in our cohort. Moreover, both CTA and GoP displayed moderate to high correlations with human listener effort, potentially highlighting the relative importance of timing over spectral information in characterizing ALS pathological speech.

*Medians of each of the metrics considered for different ALS cohorts (*Mann Whitney pairwise tests showed that the means were significantly different for each metric and each cohort at $p \leq 0.00001$ except for GOP between PRESYM and CONTROL which was less significant at $p = 0.01$*

	BULBAR	PRESYM	CONTROL
GOP	-0.31	-0.17*	-0.15*
CTA	66.71%	77.31%	80.72%
PPT	4.15%	2.72%	1.14%

Results of 3-way cohort classification experiments with 10-fold cross validation using a Random Forest classifier

feature set	F-measure*	relative improvement*
GOP+CTA+PPT	0.496	49.40%
GOP+CTA	0.477	43.67%
CTA+PPT	0.455	37.05%
GOP+PPT	0.431	29.82%
CTA	0.417	25.60%
GOP	0.412	24.10%
PPT	0.332	NA

Mean Absolute Error (MAE) and relative improvement in correlations between different feature combinations and human-scored listener effort

feature set	correlation	MAE*	relative improvement*
GOP+CTA+PPT	0.8004	15.15	81.66%
GOP+CTA	0.7920	15.46	79.75%
GOP+PPT	0.7261	16.71	64.80%
CTA+PPT	0.6969	17.95	58.17%
GOP	0.6806	17.97	54.47%
CTA	0.6786	18.50	54.02%
PPT	0.4406	23.03	NA

CROSS-LANGUAGE GENERALIZABILITY OF ACOUSTIC FEATURES FOR ALZHEIMER'S DISEASE DETECTION MODELS

Arian Shamei[1], Yadong Liu[1], and Bryan Gick[1,2]
[1] *University of British Columbia, Vancouver, Canada*
[2] *Haskins Laboratories, New Haven, CT, United States of America*

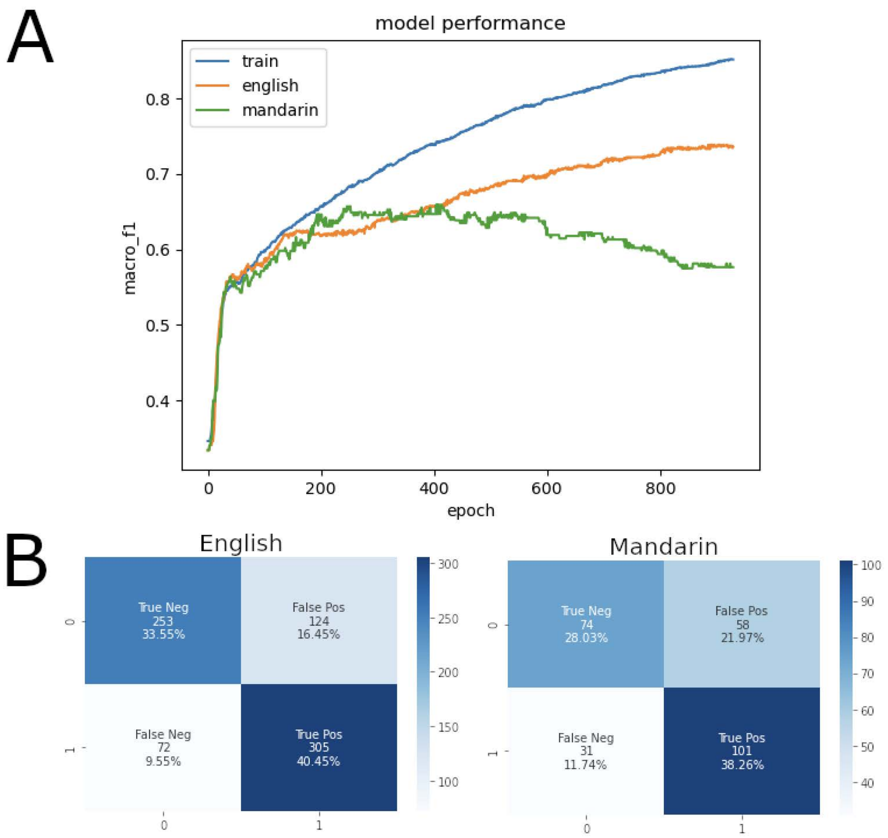
Background Alzheimer's disease (AD) is a neurodegenerative disorder with progressive loss of cognitive and motor func-tions and the main cause of dementia (Breiyeh & Karaman, 2020). The Interspeech ADreSS challenges have produced substantial work on detecting AD from speech through machine learning, with top-performing models exceeding 80%accuracy (Luz et al., 2021). Similar challenges have been conducted for other languages (e.g. Mandarin; Bao & Mong, 2021), yet there has been little discussion of cross-language generalizability of acoustic training data.

While we have found work assessing the cross-language generalizability of syntactic, semantic, and paralinguistic features for AD de-tection (Lindsay et al., 2021), we have found no work assessing the generalizability of acoustic features. We hypothesize that motor control impairments from AD should result in similar physiological alterations to speech motor control regardless of language, therefore, we expect some degree of generalizability. Here we train a Convolutional Neural Network (CNN) for binary classification of AD versus control speech using English data, and compare model performance on two validation sets: 1) a speaker-independent subset of the English data; and 2) a small parallel Mandarin-language dataset of AD and control speech.

Methods
Data: English data consisted of speech from female controls (N=46) and female AD participants (N=85) completing the cookie-theft description task within the Dementiabank[4] Pitt corpus. Mandarin data came from the 16th National Conference on Man-Machine Speech Communication[3] and included 14 females (control N=7, AD N=7) completing the cookie-theft task and free talks. Both datasets were resampled to 22K/16-bits. Using Librosa (McFee, 2022), we extracted 1000ms chunks of continuous speech, then generated a mel-formant cepstral coefficient (MFCCs) from each chunk using a 100-4000Hz range, 64 filters, n fft of 2048, hop-length of 64. In total, 3350 English MFCCs were extracted, with 2596 used in training (24 control, 55 AD speakers) and 754 for validation (13 control, 30 AD).

264 MFCCs were extracted for Mandarin validation (7 control, 7 AD). No speakers were shared between sets. The number of speech samples were equal across AD and control sets, but the number of speakers was not. Experiment: Using tensorflow (Abadi et al., 2015), we constructed a 2-layer CNN as follows: 1) Pooling at each convolutional layer. 2) Batch normalization and 15% dropout on the first layer. 3) ReLu activation; sigmoid on dense layer. 4) Adam optimizer with low learning-rate (1×10-6). Model was trained to 90% training macro-F1, with validation macro-F1 for English and Mandarin recorded each epoch. Best-epoch predictions for validation sets were plotted using confusion matrices.

Results & conclusions For English validation data, best-epoch performance was 73.8% macro-F1 at epoch 912. For Mandarin, best-epoch performance was 65.9% macro-F1 at epoch 404. Figure 1 illustrates performance at each epoch (A) for the English training set (blue), English validation set (orange), and Mandarin validation set (green), and best-epoch validation predictions via confusion matrices (B). Peak performance for Mandarin occured early and exceeded chance by 16%, then decreased due to overfitting, while English validation performance continued to improve. In sum, these results suggest that the cross-language generalizability of acoustic features for AD detection is high, though subject to overfitting. Replication with other languages and larger datasets is necessary for further insight.



Model performance and best epoch validation predictions

References

Breijyeh, Z., & Karaman, R. (2020). Comprehensive review on Alzheimer's disease: Causes and treatment. *Molecules, 25*, p. 5789.

Luz, S. et al. (2021). Speech Analysis for Alzheimer's Dementia Recognition. *Frontiers in Computer Science*, p. 96.

Bao, C., & Mong, Y. (Eds). Applied Sciences (2021): Selected Papers from 16th National Conference on Man-Machine Speech Communication (NCMMSC2021).

Lindsay, H., Tröger, J., & König, A. (2021). Language Impairment in Alzheimer's Disease — Robust and Explainable Evidence for AD-Related Deterioration of Spontaneous

Speech Through Multilingual Machine Learning. *Frontiers in aging neuroscience, 13*, p. 228.

Brian McFee et al. (2022). librosa/librosa: 0.9.1. Version 0.9.1. Feb. 2022. `https://doi.org/10.5281/zenodo.6097378`

Abadi, M., et al. (2015). TensorFlow: Large-Scale Machine Learning on Heterogeneous Systems. Software available from tensorflow.org. `https://www.tensorflow.org/`