Voice-related Biomarkers in Neurodegenerative Diseases

The 9th UEP Biomarkers Committee Meeting

Dott.ssa CAMESASCA VALENTINA

MD, ENT – Head & Neck Surgeon

Higher Education in Phoniatry: Voice, Swallowing and Speech disorders

Ospedale Grande Metropolitano Niguarda - Centro Clinico NeMO (Milano, Italy)

Neurodegenerative Diseases (NDD)

*The term "neurodegeneration" originated from the word "**neuro-**," which indicates nerve cells, and "**degeneration**," which describes a process of losing structure or function in tissues or organs.*

→ ***Any degenerative disease that predominantly affects neurons***

NDD are a vast cluster of neurological disorders with various clinical and pathological manifestations that impact particular subsets of neurons in distinct functional anatomic systems, which debilitate patients' cognitive health and physical abilities, leading to a decline in autonomy and quality of life.

15% of the global populations **Unexplained begin** **Insidious course**

Despite their prevalence, there are **few disease-modifying therapies** available to prevent or treat NDD

Neurodegenerative Diseases

the most prevalent neurodegenerative diseases:

Parkinson’s disease (PD)

amyotrophic lateral sclerosis (ALS)

including Alzheimer’s disease (AD)

Amyotrophic Lateral Sclerosis (ALS)

NDD characterized by progressive impairment of muscles' strength and function due to progressive loss of motor neurons (upper UMNs and lower LMNs).

Multiple speech subsystems (i.e., respiratory, resonatory/velopharyngeal, voice/phonatory, and articulatory) may be affected, causing devastating effects in **communication**.

Different Phenotypes (**Bulbar**, **Spinal**, **Respiratory**) across individuals

Over the course of the disease, most individuals with ALS will experience ***bulbar symptoms*** affecting **speech**, **feeding**, and **swallowing.**

Amyotrophic Lateral Sclerosis (ALS)

Over 80% of individuals with ALS develop **dysphonia** and **dysarthria**, primarily a mixed **spastic-flaccid** subtype resulting from the deterioration of both UMNs and LMNs involved in speech production.

48% patients with ALS present with dysphonia **at onset** (mainly Bulbar phenotype)

Speech impairment may **begin up to 3 years prior to diagnosis of ALS**, and as ALS progresses over time there is significant deterioration in speech.

Detectable changes in **voice quality** (hoarseness, roughness, strain, and breathiness) or **loudness**

NEED: ***Valid, reliable, and easily implemented measures of dysphonia***

→ *identify bulbar involvement,*

→ *track bulbar decline over time,*

→ *monitor progress during clinical trials*

Amyotrophic Lateral Sclerosis (ALS)

Traditional (Perturbation/Noise-Based) Acoustic Measures

Among the most commonly used acoustic measures for quantifying dysphonia in ALS have been **jitter** (a measure of cycle-to-cycle variation in frequency), **shimmer** (a measure of cycle-to-cycle variation in amplitude) and **harmonics-to-noise ratio** (**HNR**, a measure of turbulent noise present in the voice signal)

Acoustic perturbation measures $\rightarrow$ *reflect the* ***amount of involuntary variation (perturbation)*** *in the vocal signal.*

Acoustic correlates of perceptual measures of voice*: both* ***jitter*** *and* ***shimmer*** *have primarily been associated with perceived* ***roughness*** *and* ***overall voice quality****, whereas* ***HNR*** *has been associated with perceived* ***breathiness*** *and* ***roughness.***

Amyotrophic Lateral Sclerosis (ALS)

Traditional (Perturbation/Noise-Based) Acoustic Measures

However, existing Literature related to acoustic characterization of dysphonia in ALS is **limited** by several methodological factors (small sample sizes, psychometric limitations of these measurements).

→ *The **reliability** of jitter values may **decrease as the severity of dysphonia increases***

→ *Perturbation analysis is limited to **sustained vowels produced at a steady pitch**, because the characteristics of continuous speech (such as short vowel durations, fundamental frequency variation, pauses, and voiceless consonants) may significantly impact these measures*

→ **psychometric limitations of these measurements**

Amyotrophic Lateral Sclerosis (ALS)

Cepstral/Spectral Measures of Dysphonia

A growing body of literature demonstrated the utility of **cepstral and spectral approaches** (e.g., cepstral peak prominence [CPP], low-high spectral ratio [L/H ratio], and related features) as an alternative to traditional acoustic measures for objectively measuring of dysphonia.

***ADVANTAGES**: they can be extracted from **continuous speech** and that they may be **more reliable across the range of impairment severity** than traditional features.*

- **CPP: higher** in voices with strong periodicity and **lower** in voices characterized by aperiodic noise. Correlation with perceptual features of overall dysphonia and breathiness.
- **CPP SD:** reflects frequency and amplitude variations in normal speech patterns
- **L/H ratio: differentiates** normal from **dysphonic voices**. Correlation with perceptual features of overall dysphonia and breathiness.
- **L/H ratio SD**: reflects variability across the duration of a voice sample. A **higher** value corresponds to **better laryngeal support** for dynamic adjustments of the larynx particularly in continuous speech

Amyotrophic Lateral Sclerosis (ALS)

Cepstral/Spectral Measures of Dysphonia

The nature of a speech stimulus can affect cepstral/ spectral values.

CPP values differed significantly when measured during a **sustained vowel** or during **continuous speech** from the same speaker.

Different **sentences** included in the protocol for the CAPE-V have different levels of correlation with the **CSID**, and none of those correlations are as strong as the correlation with a **sustained vowel**.

→ *It is recommended the use of a **variety of stimuli during clinical voice assessment** to overcome the effect of glottal and supraglottal differences on acoustic measurements.*

Amyotrophic Lateral Sclerosis (ALS)

LIMITATIONS

The ***complexity*** of NDD, in which ***multiple speech subsystems*** may be affected, complicates the direct assessment of single subsystems since each may have a significant confounding effect on the measurement of the others.

Voice analysis in ALS is also complicated by the ***variable*** manifestation of the disease ***across individuals*** and within an individual ***across time***.

Moreover, **UMN** and **LMN** involvement results in vastly different phonatory disorders, with UMN dysfunction producing a ***spastic dysphonia*** and LMN dysfunction producing a ***flaccid dysphonia***.

!!!

→ ***Auditory-Perceptual evaluations***

→ ***Cepstral/Spectral measures***

Amyotrophic Lateral Sclerosis (ALS)

Sustained Wovel task

jitter, **shimmer**, and **HNR** are **robust acoustic measures of dysphonia**

Strongly correlated with auditory-perceptual features of vocal dysfunction: **roughness**, **breathiness**, **strain**, and **overall dysphonia severity**.

Roughness and **breathiness** are related to increased **Jitter** and **Shimmer** and represents the beginning of *laryngeal weakness* and the *decreased laryngeal control*.

Increased **HNR** reflects the turbulent noise due to *incomplete glottal closure*.

Increased **F0** due to *dysfunction of intrinsic laryngeal muscles*.

Decreased **MPT (Maximum Phonation Time)** due to *reduced control system of vocal fold movement* and progressive *weakness of respiratory support* for voice production.

→ Discriminate better ***Bulbar involvement.***

Amyotrophic Lateral Sclerosis (ALS)

Sustained Vowel task

Strong and significant correlations also between **cepstral/spectral acoustic features** and perceptual ratings.

CPP strongly correlates with ***breathiness***

CSID correlate with both ***strain*** and ***overall dysphonia severity***

L/H ratio SD correlate with ***roughness***, ***strain***, and ***overall dysphonia severity***.

→ **Better discrimination between dysphonic and nondysphonic voices**

→ **Perform well even with low levels of dysphonia**

Amyotrophic Lateral Sclerosis (ALS)

Continuous Speech

CPP and **CPP SD** correlate with only one perceptual feature (**CPP** with **Strain** and **CPP SD** with **Breathiness**).

L/H ratio significantly correlates with **breathiness**, **strain**, and **overall dysphonia**.

→ *During continuous speech, cepstral measures are affected by **fluctuations in a variety of laryngeal and articulatory factors** including vocal intensity, fundamental frequency, sound pressure level, syllable stress, vowel context, and vowel type.*

→ *Continuous speech, on the other hand, is **more representative of habitual voice use patterns** containing pitch and loudness variations that serve as important perceptual indicators of vocal dysfunction.*

CPP SD is sensitive to features of continuous speech that impact the variability in the voice signal's periodicity, such as changes in the vowel spectrum and changes in the frequency spectrum due to intonation patterns

Amyotrophic Lateral Sclerosis (ALS)

Continuous Speech

→ *Acoustic assessment of continuous speech is complicated by various factors related to* ***dysarthric speech.***

→ *The deterioration of the* ***resonatory*** *(i.e., velopharyngeal function) and* ***phonatory subsystems*** *occurs* ***prior*** *to* ***articulatory subsystem*** *decline.*

The particular characteristics of **motor speech disorders** (e.g., **reduced articulation rate**, **reduced vowel space**, and **increased frequency and length of pausing**) confound cepstral/spectral acoustic measures when applied to individuals with complex neurodegenerative diseases.

Additionally, the **mixed spastic-flaccid dysarthria** associated with ALS has the potential to "cancel out" straightforward associations among individuals with complex neurodegenerative diseases.

→ ***Hypernasality, reduced intelligibility,*** *and* ***overall speech impairment*** → ***BIAS !!***

Amyotrophic Lateral Sclerosis (ALS)

Hypernasality

Hypernasality is a distinctive perceptual feature of speech resonance caused by **velopharyngeal dysfunction**.

Acoustically, hypernasality is characterized by the

- reduction in the amplitude of the first formant (**F1**),
- presence of an **extra pole (or nasal formant)-zero pair** between the first and second formants,
- **shifts of the center of the low-frequency** spectral prominence,
- increased amplitude of the **bands between F1 and F2,**
- decreased amplitude in the F2 region.

Amyotrophic Lateral Sclerosis (ALS)

Hypernasality

Because hypernasality is an excessive nasal resonance that occurs on voiced segments (i.e., vowels and voiced consonants), several acoustic indices based on **voice spectral analysis** have also been proposed to correlate with the perceptual evaluation of hypernasality.

- difference between the amplitude of F1 and the amplitude of the extra peak (A–P1) where P1 locates between F1 and F2
- difference between the amplitude of F1 and the amplitude of the first nasal peak P0, which compared to P1 is lower in peak frequency and locates below the F1
- voice low tone (below 600 Hz) to high tone (above 600 Hz) spectral ratio
- and the one-third octave spectral analysis

Amyotrophic Lateral Sclerosis (ALS)

Hypernasality - Cepstral/spectral measures

Cepstral/spectral measures better reflect the underlying pathophysiology and demonstrate stronger associations with perceptual ratings of voice quality.

CPP and **L/H ratio** are used to quantify other aspects of speech production and may reflect not only the **periodicity of vocal folds vibration**, but also the **filtering effects and the resonant function of the vocal tract.**

Concurrent impairments in velopharyngeal mechanism and voice subsystem yield confounding effects on spectral energy, causing misclassification.

For instance, ***voice*** impairments are often characterized by increased energy (noise) in the higher frequencies, yielding **lower L/H spectral ratio** values than in unimpaired speakers. In contrast, ***hypernasality*** causes reduced amplitude of spectral energy at higher frequencies potentially resulting in **higher L/H spectral ratio.**

Amyotrophic Lateral Sclerosis (ALS)

Hypernasality - One-third octave analysis

One-third octave analysis using the frequency band centered at 1600 Hz is effective at differentiating **Asymptomatic** patients, predominantly **Hypernasal**, predominately impaired **Voice**, and **Mixed)**

→ *potential to be used as an early indicator!*

Demonstrates subsystem-specific alterations in the power of distinct frequency bands:

- one-third octave frequency bands **below the band centered at 1600 Hz** captures pronounced impairment in the resonance subsystem (velopharyngeal mechanism)
- one-third octave frequency bands **centered at 1600 Hz and above** captures impaired voice quality as well as impairment in voice and resonance subsystem

Amyotrophic Lateral Sclerosis (ALS)

Hypernasality - CPP

Sensitive acoustic measure that varies depending on speech task/stimuli, vocal tract configuration, and intensity.

→ ***Nasalized vowels*** *are associated with* ***reduced CPP****, while* ***nasal sentences*** *(senteces loaded with nasal consonants) are characterized with* ***increased CPP !!***

CPP decreased in patients with predominately **voice impairment**, while it **increased** in patients with predominately **hypernasality** (nasality impairment), which may explain the **lack of differentiation** of the mixed from asymptomatic group (reduced CPP due to voice impairment obfuscated by an increase due to hypernasality).

Hypernasality facilitates phonation and improves the synchronized harmonic organization of voice.

→ nasalization decreases the open quotient of the vibratory cycle by inducing increased muscular adductory forces of vocal folds, independent of intensity or vocal loudness.

→ presence of a strong low-frequency spectral energy during nasal resonance may lead to increased CPP.

→ ALS patients commonly have impairments across all speech subsystems, including articulatory and respiratory, which may influence CPP measure.

Amyotrophic Lateral Sclerosis (ALS)

Hypernasality – L/H spectral ratio

Some Authors demonstrated the relations of L/H spectral ratio with nasalance and hypernasality rating at a specific cutoff frequency (i.e., 600 Hz), while others did not find such associations.

Hypernasality causes increased spectral energy (amplitude) of the frequency bands between F1 and F2, and reduced amplitude around 2500 Hz.

Voice impairment is often characterized by increased energy (noise) in the higher frequencies, yielding lowered L/H spectral ratio values than unimpaired speakers.

CONTROVERSIES

!!!

Amyotrophic Lateral Sclerosis (ALS)

Hypernasality – Sum up

One-third octave analyses and the CPP measure have high classification **accuracy** for identifying asymptomatic samples from those perceived with predominantly nasality impairment.

CPP: higher predictive power for differentiating ***asymptomatic*** from perceived ***voice impairment***.

One-third octave analyses: greater accuracy for distinguishing ***asymptomatic*** from perceived ***mixed hypernasality and voice impairment***.

The power (dB) of several one-third octave frequency bands more specifically the frequency band centered at 1600 Hz may be a robust acoustic biomarkers with the potential to distinguish asymptomatic samples from samples with hypernasality, voice, and mixed hypernasality, and voice impairment with high accuracy.

Amyotrophic Lateral Sclerosis (ALS)

Voice Onset Time (VOT)

VOT is a measure of the temporal difference between an articulatory stop release and the onset of vocal fold vibrations.

Index frequently used to describe intergestural coordination between the articulatory and laryngeal systems during speech and important acoustic cue for the voiced–voiceless distinction.

Speakers in the **late-stage ALS** have a significantly **longer** lead VOT values, whereas there are no differences in VOT between the **early-stage ALS** and **healthy** speaker.

Speakers with **more severe ALS** showed greater occurrence of **voicing lead** and **longer voicing lead**.

Voicing precedes **articulatory** onset with disease progression in the production of bilabial stops, which suggests that the relative timing of *coordination* between the *supralaryngeal structures* and the *phonatory system* is affected in the late stage of ALS.

Amyotrophic Lateral Sclerosis (ALS)

Voice Onset Time (VOT)

*There is a direct relationship between **respiratory** and **phonatory** subsystems, possibly because of **shared neural circuitry**, where modifying respiration showed more impact on voicing than articulation.*

The respiratory subsystem, particularly low lung volume has been found to **shorten VOT .**

*For the production of stops, the difference between intraoral pressure and subglottal pressure determines when the voicing occurs. In **ALS patients**, during the production of the voiced word-initial stops, **glottal pulsing precedes the stop release**, suggesting that the **laryngeal system is activated prior** to the build-up of intraoral pressure for the stop*.

*Three possible **adjustments to the articulatory-laryngeal subsystems** could occurre while maintaining or initiating phonation during the closure in voiced stops: a **passive supraglottal expansion**; an **active enlargement of the cavity** by adjustments like larynx lowering, tongue root advancement, tongue body lowering; **nasal airflow** through an incomplete velopharyngeal closure*

In the group with **late-stage ALS**, **longer negative (lead) VOT** may be due to **impaired lip and jaw movements** or due **to supraglottal leakage from weakness in the velopharyngeal closure.**

Amyotrophic Lateral Sclerosis (ALS)

Voice Onset Time (VOT)

*During the production of **bilabial stop consonants**: lips move at high velocity at the onset of oral closure, resulting in tissue compression and airtight seal for intraoral pressure buildup. There are adjustments that may be enabling healthy speakers to have flexible control using either voicing lead or short lag voicing mode*

Findings of reduced occurrence of a **flexible mode of voicing** (**lead** vs. **short lag**) and **inverse correlation between VOT and intelligible speaking rate** may be indicative of the neurodegenerative changes that occur with **disease severity in ALS**.

*Changes in VOT with speech **decline in ALS** could be considered as **progressive loss of phonetic features**, such as **voiced-voiceless contrast** and **oral** versus **nasal contrasts** in later stages of ALS.*

*The coordination between the articulatory and phonatory speech subsystems may be influenced by individual **differences in neural degeneration localization:** different phenotypes (**bulbar**, **spinal**, **respiratory**) different prevalence in late-stage ALS (**bulbar symptomatic - high speech function** vs. **low speech function**).*

Amyotrophic Lateral Sclerosis (ALS)

Voice Onset Time (VOT)

Patients in **late-stage ALS** have **longer lead VOT** than patients in early-stage ALS and healthy speakers.

VOT duration increases with increasing **severity of speech** symptoms in ALS.

VOT reflects impaired temporal coordination between the laryngeal and supralaryngeal structures.

***Longer lead VOT** evident in **late-stage ALS** reflects neurodegenerative changes by the disease progression.*

***Voicing lead** occures predominantly in **late-stage ALS**, suggesting that laryngeal and articulatory coordination changes with disease progression.*

Amyotrophic Lateral Sclerosis (ALS)

Machine Learning Voice Analysis

JMIR MEDICAL INFORMATICS Tena et al

Original Paper

Detection of Bulbar Involvement in Patients With Amyotrophic Lateral Sclerosis by Machine Learning Voice Analysis: Diagnostic Decision Support Development Study

Alberto Tena[1], MSc; Francec Claria[2], PhD; Francesc Solsona[2], PhD; Einar Meister[3], PhD; Monica Povedano[4], PhD

[1]Information and Communication Technologies Group, International Centre for Numerical Methods in Engineering, Barcelona, Spain
[2]Department of Computer Science, Universitat de Lleida, Lleida, Spain
[3]Institute of Cybernetics, Tallinn University of Technology, Tallinn, Estonia
[4]Motoneuron Functional Unit, Hospital Universitari de Bellvitge, Barcelona, Spain

A total of 15 acoustic features were extracted: jitter(absolute), jitter(relative), jitter(rap), jitter(ppq5), shimmer(relative), shimmer(dB), shimmer(apq3), shimmer(apq5), shimmer(apq11), pitch(mean), pitch(SD), pitch(min), pitch(max), HNR(mean), and HNR(SD).

Acoustic analysis of the vowels elicited from patients with ALS may be used for **early detection of bulbar involvement**. This could be done **automatically** using supervised classification models.

Better performance by applying **PCA (principal component analysis)** previously to the obtained features.

Bulbar involvement can be detected using automatic tools before it becomes perceptible to human hearing. The results point to the importance of obtaining objective measures to allow an early and more accurate diagnosis, given that humans may often misdiagnose this deficiency.

Alzheimer's Disease (AD)

*NDD characterized by **memory loss**, **language dysfunction**, and various **cognitive disorders**, which ultimately lead to patients' **loss of independent living ability**.*

*With the **aging of world population**, this disease will bring a great burden to the families concerned and may cause great social and economic pressure. Number of patients will reach 152 million in 2050 !*

The early-stage Mild Cognitive Impairment (MCI) is a keyphase to decide whether AD can be controlled or not, and to stop or slow down disease developing into dementia later stage

DIAGNOSIS:

- **Cerebrospinal fuid**: Aβ and protein Tau → expensive, painful, and risky
- **MRI – PET**
- **Scales** and **Measures**: Montreal Cognitive Assessment (MoCA), Clinical Dementia Rating (CDR), Mini-Mental State Exams (MMSE), and Alzheimer's Disease Assessment Scale Cognitive-Subscale (ADAS-cog) → not achieve high diagnosis accuracy or reliability.

→ ***Urgent need to develop a more effective and less risky method to diagnose AD !!!***

Alzheimer's Disease (AD)

*Subtle changes in voice and language can be observed **years before** the appearance of prodromal symptoms of Alzheimer's disease.*

- Alterations of verbal fluency, reflected by the patient's hesitation to speak and slow speech rate
- Word finding difficulties, leading to circumlocution and frequent use of filler sounds (e.g., uh, um)
- Semantic errors, indefinite terms, revision, repetitions, neologisms,
- Lexical and grammatical simplification,
- Loss of semantic abilities in general
- Discourse characterized by reduced coherence, with implausible and irrelevant details.
- Alterations perceived in prosodic features (pitch variation and modulation, speech rhythm)

*Voice features have the potential to become simple and noninvasive biomarkers for the **early diagnosis** of AD and, more generally, conditions associated with dementia.*

Alzheimer's Disease (AD)

Formants F1 F2 F3 and Voice Onset Time (VOT)

F3 of wovel /u/

- *Usually related to **lip rounding**.*
- *The production of **/u/** requires **higher spatio-temporal constraints** and its production necessitates **two simultaneous constrictions**: a bilabial constriction and a tongue-dorsum one in the velar region.*

Higher VOT values of consonant /t/

*The consonant /t/ is likely to be the most affected because it requires high **gestural celerity**, together with a very **precise tongue-tip contact for occlusion**.*

Alzheimer's Disease (AD)

Automatic Voice Analyzer

able to automatically analyze temporal and acoustic voice markers to flag the onset of preclinical AD.

Automatic Speech recognition (ASR)

Able to acoustically distinguish early-stages of AD using prosodic cues related to **pitch** (e.g., rate ofvocal fold vibration during voiced segments of speech), **voicing** (e.g.,percentage of speech produced utilizing vocal folds such as with vowel sounds as opposed to unvoiced harsher sounds usually associated with consonants), and **speaking rate and formant energy** (e.g., spectral shape of energy in voiced sounds)

Machine Learning and Deep Learning-based Speech Analysis

Lexical-semantic Index: sensitive to amyloid positivity, higher diagnostic accuracy, associated with disease progression

Acoustic scores : sensitive to cognitive status, higher diagnostic accuracy compared, more sensitive to amyloid status in prodromal versus preclinical AD.

→ *Preliminary findings suggest that digital voice biomarkers are not only able to detect cognitive impairment using brief audio recordings, but distinguish AD biomarker status and disease progression*

Neurodegenerative Diseases (NDD)

Promising voice-related digital Biomarkes...

...need large Clinical Trials

Thanks for your attention

Dott.ssa CAMESASCA VALENTINA

MD, ENT – Head & Neck Surgeon

Higher Education in Phoniatry: Voice, Swallowing and Speech disorders

Ospedale Grande Metropolitano Niguarda - Centro Clinico NeMO (Milano, Italy)