EEG Evidence for Audiovisual Speech Perception Deficits in Parkinson's Disease: A Research Proposal

Geena Bruno

Introduction:

EEG stands for electroencephalogram, has 32 electrodes on the HydroCel Geodesic Sensor Net (HCGSN), and is available with densities of 32, 64, 128, or 256 channel net. It is connected to an amplifier and amplified electrical signals coming from the brain are recorded. Researchers are interested in the brain's response to a specific stimuli (Khamis-Dakwar, 2012). EEG does not show where activation originates, but rather, the timing of how a person's brain responds to certain stimuli in a task presented. An event-related potential (ERP) is the average electrical activity correlated with the specific type of stimulus presentation; researchers have to segment parts of EEG to get ERPs. ERP components are defined in terms of a combination of polarity, latency, and scalp distribution, and are sensory cognitive and motor processes relating to behavior and thought (Kappenman & Luck, 2011). Mismatch negativity (MMN) is an ERP component that peaks to 150–250 milliseconds when there is a change of stimulus (Khamis-Dakwar, 2012); for example, if the stimulus is an oddball paradigm, having a standard and deviant sound (ta, ta, ta, da, ta). MMN also has the largest amplitude at frontal and central electrode sites of the skull, and the patient does not need to respond in order to get an MMN response on the EEG (Fonaryova Key et. al, 2005). This is great for babies who cannot respond on their own, or for patients who have cognitive deficits.

Motor speech disorders are difficulties relating to problems of movement, resulting from a neurological disorder or injury that affects motor planning, programming, coordination, timing, and execution used for speech; these disorders can also affect respiration, resonation, phonation, and articulation (Owens Jr. et. al, 2014) The basal ganglia (BG) is a subcortical area of the brain that regulates motor functioning and maintains posture and muscle tone. It has direct and indirect pathways, and if damaged it results in either slowed movements as in Parkinson's disease or involuntary movements as in Huntington's disease (Owens Jr. et. al, 2014). The BG is also involved in online monitoring of auditory feedback during speech (De Keyser et. al, 2016). Damage to the neural circuitry connecting the prefrontal and BG areas of the brain causes an increase in syntactic and voice onset timing errors on single-word production tasks (Walsh & Smith, 2011).

Dysarthria is a group of speech disorders resulting from disturbances in the central and peripheral nervous systems that control muscles of speech production. Specifically, in Parkinson's disease it is described as hypokinetic dysarthria. Hypokinetic dysarthria is a BG control circuit pathology evident in voice, articulation, prosody. It has effects of rigidity, difficulty initiating movement, reduced force and range of motion in articulators, slow but sometimes fast repetitive movements on speech, reduced vocal loudness with harsh–hoarse quality, slow speaking rate with burst of rapid-fire articulation, excessive in long pauses, prolonged syllables, mono-loudness, and reduced phonation time (Freed, 2000; Owens et. al, 2014). Individuals with Parkinson's disease also have higher lip aperture variability, portraying less consistent oral motor coordination than the control participants (Walsh & Smith, 2011). Sensorimotor integration deficits may result from dysfunctions of

feedforward and feedback control in people with this disease (De Keyser et al, 2016).

Importantly, people diagnosed with Parkinson's disease have poor temporal discrimination

for tactile, auditory, and visual stimuli, also known as proprioceptive deficits (Duffy, 2005),

which will have an effect on auditory-visual integration.

Research shows that multisensory integration increases speed of detection,

enhances sensory sensitivity, and correctly identifies events: "Multisensory integration is

most effective and therefore elicit maximal behavioral enhancements when less intense or

weak and ambiguous individual stimuli are applied" (Freiherr et. al, 2013). This is not only

evident in people with a disorder or disease, but also in normal hearing people; when you

are at a bar or loud restaurant, you rely on visual cues like looking at someone's mouth to

help you understand what the person is saying to you. This is also called the Principle of

Inverse Effectiveness; multisensory integration helps counteract the consequences of

unisensory deterioration (Freiherr et. al, 2013). Also, substantial work in the speech

production and speech perception literature supports the notion that speech perception

and production have a bidirectional influence on one another (De Keyser et. al, 2016). To

add to this, visual articulatory information integrates with the motor system during speech

production (Venezia et. al, 2016).

One way to elicit an MMN response is to use the McGurk effect, which is what is

being used in this research proposal. The McGurk effect is a perceptual illusion that shows

the influence of visual speech information on the perception of speech (Cienkowski &

Carney, 2002). What the patient *sees* influences what the patient *hears* (Francisco et. al,

2017). The McGurk Effect shows that manner of articulation and voicing are transmitted

most efficiently from the auditory portion of the stimulus and place of articulation is transmitted best by the visual portion (Hessler, et. al, 2013).

This study will investigate the possible differences in audiovisual speech perception processing in individuals with Parkinson's disease and healthy individuals as indexed by the MMN component recorded using an EEG net.

Methods

A. *Participants*

There will be 10 individuals in the PD group and 10 healthy individuals in the control group. Healthy individuals are matched with patients with PD by gender and age. Patients with PD were diagnosed by a neurologist using the Unified Parkinson's Disease Rating Scale. Hypokinetic dysarthria was diagnosed by speech language pathologists using the Sentence Intelligibility Test and Arizona Battery for Communication Disorders of Dementia. Each participant received a speech evaluation that included a case history of his or her speech and language, an oral mechanism examination, perceptual judgments of respiration, phonation, resonance, articulation, and prosody during sustained phonation, a diadochokinetic task, and reading and spontaneous speech. Participant interviews and review of PD participants' medical records were used to determine that all participants had negative histories for head trauma and for pre-existing communication, memory, neurologic, or psychiatric problems.

B. EEG

High-density EEG will be recorded from participants while they are exposed to experimental stimuli. In the AV McGurk paradigm, participants view a standard presentation of congruent auditory and visual information (e.g. articulation of /ba/) interspersed with a deviant presentation of incongruent visual (e.g. articulation of /ga/) dubbed over the original audio stimulus. In EEG experiments utilizing the McGurk effect, the congruent audiovisual presentation of /ba/ is presented repeatedly as the standard stimulus. Infrequent presentation of the McGurk stimulus, the incongruent audio presentation of /ba/ paired with visual presentation of /ga/, generates the MMN (Saint-Amour et. al, 2007). In this paradigm the presented audio is consistently /ba/ and only the visual stimulus changes.

An additional AV condition utilizing an inverse McGurk deviant will be implemented to explore the effects of incongruence as mediated by modality. The inverse McGurk condition (AV inverse, labeled AI) utilizes congruent presentation of auditory and visual /ba/ as a standard stimulus with a change to auditory /ga/, while maintaining visual /ba/ for the deviant stimulus. Healthy individuals perceive the auditory aspect of the stimulus, or do not fuse responses into a single percept, but rather perceive both sounds simultaneously (/b-ga/). The AI condition in the present study would provide a direct contrast with the AV condition, elucidating the influence of modality on incongruent deviance detection.

A visual-only (VO) control condition is also necessary to ensure that the derived MMN is due to AV integration processes (visual information changing the auditory percept) rather than responses to change in visual stimulus (Saint Amour et al., 2007).

The data will be analyzed within groups and between conditions as well as across conditions and between groups, within the time windows of interest, to evaluate whether there were significant differences in MMN peak amplitude.

Condition	Standard 80%	Deviant 20%	Deviant Percept
AV	auditory /ba/ and visual /ba/	auditory /ba/ and visual /ga/	/da/
AI	auditory /ba/ and visual /ba/	Auditory /ga/ and visual /ba/	/ba/ or /b-ga/
VO	Visual /ba/	Visual /ga/	/ga/

Figure 1

This experiment will utilize a 128-channel EEG system (Electric Geodesics, Inc.). The 128 electrodes are arranged in a predictable geodesic position relative to one another in a sensor net. The electrodes are held together by a fine elastomer and contain a silver chloride-plated carbon fiber embedded in a plastic substrate. Each electrode has sponge inserts that are soaked in an electrolyte solution of potassium chloride and water before use to ensure optimum conductivity. Each participant will wear a net that fits the circumference of their head, which will have been measured by the researcher. In addition, measurements are taken to locate the vertex of the skull in order to accurately place the

net. After the net is placed on the participant, it is ideal to have the individual sit in a chair in a sound-attenuated room. The participant sits in front of a computer monitor that presents the stimuli and the sensor net connects to a calibrated amplifier.

Figure 2

C. Stimuli

The auditory-visual (AV) stimuli throughout this experiment were created to evoke the McGurk MMN. Stimuli were created by digital recording of a female native speaker of American English saying /ba/ and /ga/. Digital video (Canon Vixia HFR50) and correlating audio (Blue Mic Yeti Pro, www.bluemic.com) were recorded at a sampling rate of 44.1KHz and a frame rate of 24 images/second; they later trimmed for a total duration of 300 ms per token. The places of articulation for /ba/ and /ga/ differ greatly. Since the auditory distinction in this paradigm depends on place of articulation, video segments began in the preparatory articulatory position—closed lips for /ba/ and open lips for /ga/. The speaker was instructed to open her mouth minimally. Visual inspection of video segments ensured that jaw coordination was consistent between the /ba/ and /ga/ video segments. Video frame was cropped using Apple iMovie to reveal only the speaker's mouth in order to constrain the visual presentation and to avoid eye-movement artifacts during EEG

recording. The audio tracks were separated from the video and edited in Praat with 50ms

rise/fall to avoid click artifacts in the recording, and amplitudes normalized to 70 dB. The

vowel segment from one /ba/ recording was removed in Praat and the spliced /a/ segment

was used with the onsets for both /b/ and /g/, so the only difference in the audio is the

consonant segment. The audio track for /ba/ was dubbed over the video tracks of both

/ba/ and /ga/, creating congruent (auditory /ba/, visual /ba/) and incongruent McGurk

(auditory /ba/, visual /ga/) AV stimuli in Apple iMovie. Onset of the AV stimuli begins with

the contrastive articulatory position, closed lips for /ba/ and open mouth for /ga/. The AI

condition was similarly created, with an inverse McGurk deviant (auditory /ga/ dubbed

onto visual /ba/).

Visual-only (VO) stimuli consisted of the same 300 ms /ba/ and /ga/ video tracks

with audio removed. EEG epochs were segmented to coincide with the onset of the

auditory component of stimulus to more specifically examine the neurophysiological

response to the visual influence on auditory perception (Hessler et al., 2013).

Figure 3

  a. *Oddball Paradigm*

Stimuli for the AV, AI, and VO conditions will be presented in an oddball paradigm. This is when the standards were presented for 80% of trials and the deviants for 20% of trials. Each condition has 450 total trials, with 360 standards and 90 deviants. Stimuli will be presented pseudorandomly in order to ensure that at least two standards came before every deviant and that deviants were not played consecutively. The interstimulus interval (ISI) for all conditions is 600 ms. Presentation of each condition is counterbalanced between participants within each group.
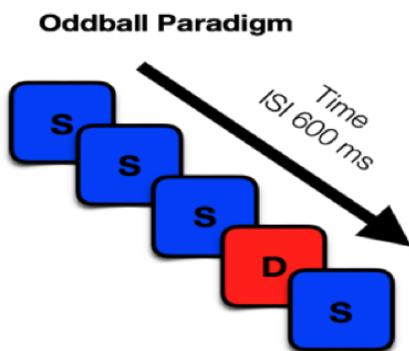

Figure 4

B. Experiment/Behavioral Measures and Procedures

a. Behavioral Measures

1. The Sentence Intelligibility Test (SIT) has 1,100 sentences (100 sentences ranging from five to 15 words in length). Participants will be evaluated using 10 to 15 sentences to test intelligibility and will be asked to repeat sentences read by the researcher. This will help to make sure participants have a high intelligibility rate to understand what they need to do throughout the experiment and the instructions they need to follow (Yorkston, Beukelman, & Tice, 1996).

2. The Arizona Battery for Communication Disorders (ABCD) is a comprehensive assessment and screening of dementia. It includes 14 subtests that evaluate verbal episodic memory, linguistic comprehension, visuospatial construction, etc. The ABCD was standardized on Alzheimer's and Parkinson's disease patients (Bayles & Tomoeda, 1991). It is important that patients have at least 80% or more on the ABCD in order to participate in this study.

c. Participation in the experiment should involve two visits:

<u>Visit 1:</u>

1. Participants will be shown the lab, equipment, and procedures by an experienced speech-language pathologist with experience diagnosing and treating individuals with PD. Caregivers will accompany patients with PD and questions will be encouraged throughout this familiarization period.

2. Participants will be presented with a consent form that the participant, caregiver (if necessary), and researcher have to sign. The speech-language pathologist gives time for participants to read the form and also provides a verbal explanation. Last, the participant and caregiver are informed of any risk factors and of their option to withdraw at any time during the experiment.

3. Following completion of consent procedures, participants participated in administration of standardized tests

4. Visit 1 will vary in time depending on each participant's needs

<u>Visit 2:</u>

1. After assessing if the participants were appropriate for the study, the participants will be invited back to complete the experimental tasks. On the second lab visit, a consent form (the same as for visit 1) will be presented again and there will be opportunities for questions.

2. Hearing must be screened at 500, 1000, 2000, and 4000 Hz within a range of 40–20 dB before the start of the experiment.

3. The circumference of the head will be measured to choose the correct net size and the vertex will be marked on the skull to ensure that all the electrodes are correctly placed.

4. The participant will be seated 80 cm from the computer monitor, in a sound-attenuated room. The participants will be given insert headphones; the red-coded insert goes into the right ear and the blue in the left. A video camera will be in the room to give the researcher visual information about the participant during the experiment. The participant will be reminded to signal at any time during the experiment if he or she does not wish to continue. The amplifier must be checked and calibrated before the net is connected, and impedances (loss of signal between scalp and sensor) must be measured. In order to improve impedances, the electrodes need to be adjusted as necessary so that they are in good contact with the participant's scalp.

5. Experimental EEG tasks will be presented in random order and counterbalanced across participants. Tasks will be presented in short runs of less than 10 minutes to minimize fatigue and reduce habituation that interferes with MMN elicitation. Participants will be encouraged to take short breaks between runs.

6. After completing the experimental tasks, the speech-language pathologist will remove the EEG net and the participant will be asked questions. Visit 2 may last approximately 60 minutes or longer, depending on participant's needs and number and duration of breaks.

Proposed Results:

Speech-production impairment in PD may impact higher-level linguistic and cognitive processing as indexed by the McGurk MMN. The comparison group is expected to show the classic MMN in response to the McGurk effect in the AV condition but not the AI condition. It is anticipated that responses in the PD group will be more variable than in the comparison group. This means that individuals with PD not getting an early automatic fusion MMN response in the AV may have a deficit in audiovisual integration. We expect the PD group's response to be similar in the AV and AI conditions, indicating that they can detect the oddball deviant but do not experience AV integration. Results for VO stimuli will be consistent within the PD group and comparison group, demonstrating that the results are not due to simple change detection of the visual stimulus.

Results will also reveal a smaller amplitude and longer latency response in individuals with PD as compared to controls. This means that patients with PD will have a slower response time and have a tendency toward later going response. This will change their MMN response to look like a P300 response (another ERP component derived from EEG measurements). These results explain that individuals with PD have a deviance in detection of stimuli and an insensitivity to modality.
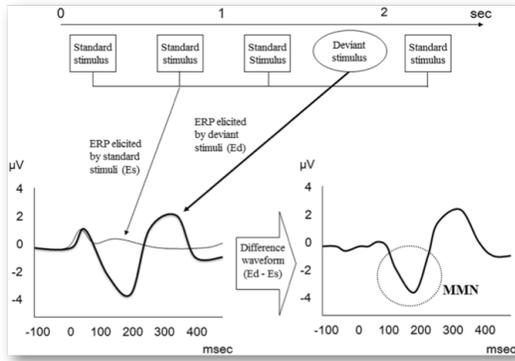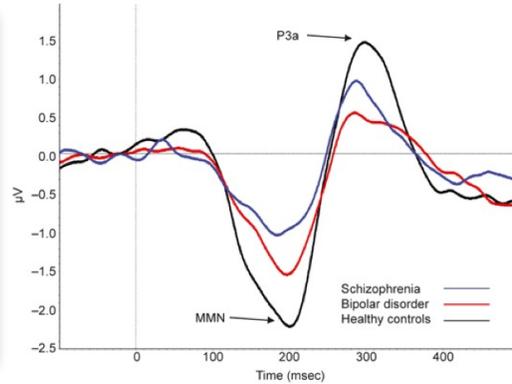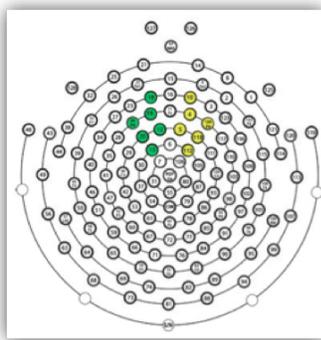
Figure 5 (examples of MMN and P300 responses)

Figure 6 (Jahshan et. al, 2012)

MMN is recorded at largest over the fronto central electrodes, regardless of where the process is happening. The image below is a representation of this implication.

Left montage/Right montage

Figure 7

Discussion and Clinical Implication:

The results of the proposed study will help us understand if deficits in AV integration are part of Parkinson's disease or a result of the ongoing experience of living with a speech impairment. As stated in literature, patients with PD have proprioceptive deficits that cause poor temporal discrimination for tactile, auditory, and visual stimuli (Duffy, 2005). This can have an effect on AV integration, and therapy may not change these deficits. Also, in many therapy techniques for patients with speech, language, and hearing

disorders, speech therapists emphasize the importance of the principle of inverse effectiveness, or using multiple senses to counteract the negative effects of weakening one sensory system (Freiherr, 2013). While we tend to use visual articulatory cues during speech therapy, we don't know if visual articulatory information is facilitative or overloading the system in patients with PD.

References

Bayles, K. A., & Tomoeda, C. K. (1991). Arizona Battery for Communication Disorders of Dementia. Tucson, AZ: Canyonlands Publishing.

Cienkowski, K. M., & Carney, A. E. (2002, November). Auditory-visual speech perception and aging. *EAR & HEARING*, *23*(5), 439-449.

De Keyser, K., Santens, P., Bockstael, A., Botteldooren, D., Talsma, D., De Vos, S.,... De Letter, M. (2016, October). The relationship between speech production and speech perception deficits in Parkinson's disease. *Journal of Speech, Language, and Hearing Research, 59,* 915-931.

Duffy, J. R. (2005) *Motor Speech Disorders: Substrates, Differential Diagnosis, and Management* (2nd ed.). Elsevier Mosby.

Fonaryova Key, A. P., Dove, G. O., & Maguire, M. J. (2005). Linking Brainwaves to the Brain: An ERP Primer. DEVELOPMENTAL NEUROPSYCHOLOGY,27(2), 183-215. DOI: 10.1207/s15326942dn2702_1

Francisco, A, A., Jesse, A., Groen, M. A., & Mcqueen, J. M. (2017, January). A general audiovisual temporal processing deficit in adult readers with dyslexia. *Journal of*

Speech, Language, and Hearing Research, 60, 144-158

Freed, D. (2000). *Motor speech disorders: diagnosis and treatment.* San Diego, CA: Singular

Thomson Learning.

Freiherr, J., Lundstrom, J. N., Habel, U., & Reetz, K. (2013, December). Multisensory integration

mechanisms during aging. *Frontiers in Human Neuroscience, 7,* 1-6.

Hessler, D., Jonkers, R., Stowe, L., & Bastiaanse, R. (2013). The whole is more than the sum of its

parts–Audiovisual processing of phonemes investigated with ERPs. *Brain and

language, 124*(3), 213-224.

Kappenman, E. S., & Luck, S. (2011). ERP Components: The Ups and Downs of Brainwave

Recordings. *The Oxford Handbook of Event-Related Potential Components,* 3-30.

doi:10.1093/oxfordhb/9780195374148.013.0014

Khamis-Dakwar, R. (2012). Translational display of neurophysiological investigations. In

Communication Sciences and Disorders . In: Goldfarb, R. (Ed.). *Translational SLP/A:

Essays in honor of Dr. Sadanand Singh.* Plural Publishing, Inc. IntroductionToERP(2012)

Owens, R. E., Jr., Metz, D. E., & Farinella, K. A. (2014). *Introduction to Communication

Disorders: A Lifespan Evidence-Based Perspective, Fifth Edition.* (5th ed.). Pearson.

Saint-Amour, D., De Sanctis, P., Molhoim, S., Ritter, W., & Foxe, J. J. (2007, February). Seeing

voices: High-intensity electrical mapping and source-analysis of the multisensory

mismatch negativity evoked during the McGurk Illusion. *Neuropsychologia, 45*(3),

587-597

Venezia, J.H., Fillmore, P., Matchin, W., Isenberg, A.L., Hickock, G., & Fridriksson, J. (2016).

Perception drives production across sensory modalities: A network for sensorimotor

integration of visual speech. *NeuoroImage, (126),* 196-207.

Walsh, B., & Smith, A. (2011, June). Linguistic complexity, speech production, and

comprehension in Parkinson's Disease: Behavioral and physiological Indices. *Journal of*

Speech, Language, and Hearing Research, 54, 787-802.

Yorkston, K. M., Beukelman, D., & Tice, R. (1996). Sentence Intelligibility Test for Windows

[Computer software]. Lincoln, NE: Tice Technology.