Study of articulatory movement from the single slice dynamic imaging of the vocal tract in Parkinsonism

S Senthil Kumaran[1], Sunita Gudwani[1], Mohit Saxena[2], and Madhuri Behari[2]

[1]*Department of NMR, All India Institute of Medical Sciences, New Delhi, Delhi, India,* [2]*Department of Neurology, All India Institute of Medical Sciences, New Delhi, Delhi, India*

Introduction: Real-time dynamic MRI has been used as non-invasive method to give valuable spatial and temporal information of vocal tract during speech production [1, 2]. Parkinsonism has perceptual features as reduced articulatory displacements for consonant production due to rigidity and improper coordination of agonist-antagonist muscle groups [3]. Speech performance in Parkinson's Disorder (PD), multiple system atrophy (MSA) and progressive supranuclear palsy (PSP) is observed to be significantly different [4]. In this study, we compare velar and alveolar natural rate syllable production in normal healthy control, PD, MSA and PSP with MRI single slice dynamics to visualize relevant clinical features in minimum time with adequate temporal & spatial information.

Methods: The subjects (native Hindi speakers) studied were normal healthy control, PD, MSA and PSP of age range 45-70 years. Imaging was done at 1.5T MR scanner (Avanto, Siemens, Germany) using 32-channel head coil. The subjects had to read aloud the velar and alveolar Hindi syllables (unaspirated, voiceless & voiced consonant-short vowel combination) at normal speech rate. The visual stimuli were presented (SuperLab4.2, Cedrus Inc. USA) using MR compatible audio-visual system with binocular LCD goggles, microphone and headphones (NordicNeuroLab, Norway). The dynamics of articulators during natural speech production was studied by 2D-single slice trufi sequence in sagittal view, with parameters: flip angle = 77°; TR = 1.90 ms; TE = 1 ms.; echo-span 2.7 ms; bandwidth = 1395 Hz/ pixel; phase = 40 freq/ s.; FOV = 360 mm; slice thickness = 10 mm (10mm X 241mm X 360mm); temporal resolution 5 frames per second or 1 frame in 200 ms; total acquisition time = 80 s. Audio recording of verbal output was done during the MRI and outside (Sony recorder) and processed in Wavepad sound editor, praat and wavesurfer software's to analyze the spectrogram.

Results and Discussion: The parameters calculated were tongue tip contact location (TTCL = distance from lip aperture for that subject in baseline position), distance between tongue-tip/ dorsum from palate (PC= precision of contact), distance of groove from lip aperture (DGLL), grooving (distance of groove base from tongue-tip), velic aperture (VA for nasality), opening of epiglottis & elevation of larynx (voicing) [2]. Mid-sagittal images show that the constriction in velar consonants /ka/ was imprecise (PC ≥ 1mm) (Fig-1; A) and the release was incomplete as well as slow (not completed till next frame of image) in MSA and PSP subject as compared to control. The velar contact was imprecise but equal rate in PD and control similar to kinematic study of articulatory movements [5]. The syllable rate calculated from contact/ release image number and images per second was observed as 2-3 syllables per second (double the normal speech rate) in MSA and PSP subjects. In PD and control, it was 4-5 syllables per second, in concordance with earlier reports [3-5]. Similar was observed by spectrogram analysis MSA (mean = 407 ± 24.35 ms), PSP (416 ± 18.05 ms), PD (mean = 228 ± 20.68 ms) and control subject (mean = 202 ± 15.75 ms) of audio recording inside/outside the MR-scanner (Fig-1; F). In PD during alveolar voiceless affricate /cha/ production the contact was posterorized (ratio of mean TTCL of voiceless alveolar consonant production in PD and control =1.92), incomplete (PC ≥ 1mm) and the tongue body grooving was less (ratio of mean grooving in PD & control 0.42). In MSA the location of tongue-tip contact was incomplete/ imprecise, constriction was anteriorized (ratio of mean TTCL in MSA & control = 0.68) along less grooving (ratio = 0.64) (Fig-1; A & D). In PSP, during /dza/ syllable the contact was posteriorized (ratio of TTCL = 1.68) with more grooving (ratio= 1.55) (Fig-1; 4D). In PD the VA was observed as ≥ 1mm in many non-nasal consonant productions showing hypernasality (Fig-1; 3B). During voiced velar and alveolar syllable (Fig-1C; 'dza' = alveolar affricate;) production the pharyngeal and laryngeal shaping revealed the vocal cord vibration [2] but the precision of contact at vocal-cords would have been better with axial view [1]. The spectrogram analysis showed that the less grooving resulted in more intensity/frication (73.26 ± 6 dB) in PD as compared to control (63 ± 4 dB) subject and more grooving resulted in less intensity/frication (59 ± 5 dB) in PSP subject (Fig-1; F), also reported by other studies [3, 4].

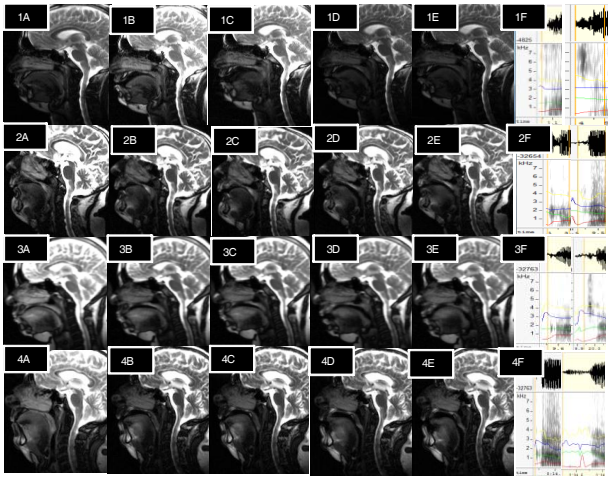

Figure 1. Mid-sagittal vocal tract representative images and spectrogram in 1: normal healthy control; 2: MSA subject; 3: PD subject; and 4: PSP subject. Images representing as (A): baseline (neutral position) and CV production (B) velar stop 'ka' contact; (C): 'ka' release; (D) alveolar affricate 'dza' contact and (E) 'dza' release; (F) spectrogram of syllable 'ka' and 'cha' in same subject.

Conclusion: Sagittal-view single slice dynamic imaging gave sufficient clinical information to be correlated with acoustic parameters for consonant articulation, although the laryngeal/ vocal vibration would have been added with axial slice information.

References: [1] Kim et al, 2012. *J Magn Reso Imag*, 35:943-948. [2] Ventura, Freitas & Tavares, 2011. *repositorio-aberto.up.pt/bitstream/10216/56115/2/29378.pdf*. [3] Gentil, Pollack & Perret, 1995. *Rev Neurol*, 151:105-112. [4] Skodda, Visserb & Schlegel, 2011. *J Voice*, 25:725-731. [5] Wong, Murdoch & Whelan, 2010. *Int. J Speech Lang Pathol*, 12:414-425.