SEMI-AUTOMATIC ANALYSIS OF SPONTANEOUS SPEECH IN HISTORY INTAKE FROM APHASIA-RELATED STROKE PATIENTS TO RANK AND CATEGORIZE FAST SYMPTOMS

E. Rigas[1], C. Spyrou[2], M. Papoutsoglou[1], H. Proios[2]

1 Aristotle University of Thessaloniki, Department of Informatics, Thessaloniki, Greece.
2 University of Macedonia, Department of Educational and Social Policy, Thessaloniki, Greece.

Background and Aims:

Spontaneous speech analysis from aphasia-related stroke patients during a history intake is often compromised due to either syntactic, morphological or semantic difficulties. Statistics and machine learning techniques were used to analyze spontaneous responses to the question “why are you here today?” from patients with stroke-related aphasia. We aim to rank and categorize Face, Arm, Speech, Time (FAST) symptoms.

Methods:

We use a dataset consisting of 58 participants, 68.96% male and 31.04% female, aged between 26 and 78 with mean 58.29 and median 61, the majority having suffered an ischemic stroke. All participants have responded orally, and audio recordings had been transcribed for subsequent analysis of the text. We used two standard natural language processing techniques, namely term frequency (tf) which shows the number of times a word occurs in a document and latent Dirichlet allocation (LDA) which distinguishes topics within documents.

Results:

Regarding tf, we rank single and double- word sequences (Tables 1-2) and we observe that “speech” and “arm” appear frequently, while words related to face, or other conditions also exist. Regarding LDA, we selected 3 topics which are clearly distinguished from each other and we observe that “speech” and “arm” (Figure 1) are categorized in two of them, while the third one is vaguer as it contains words related to the general condition of the patients.

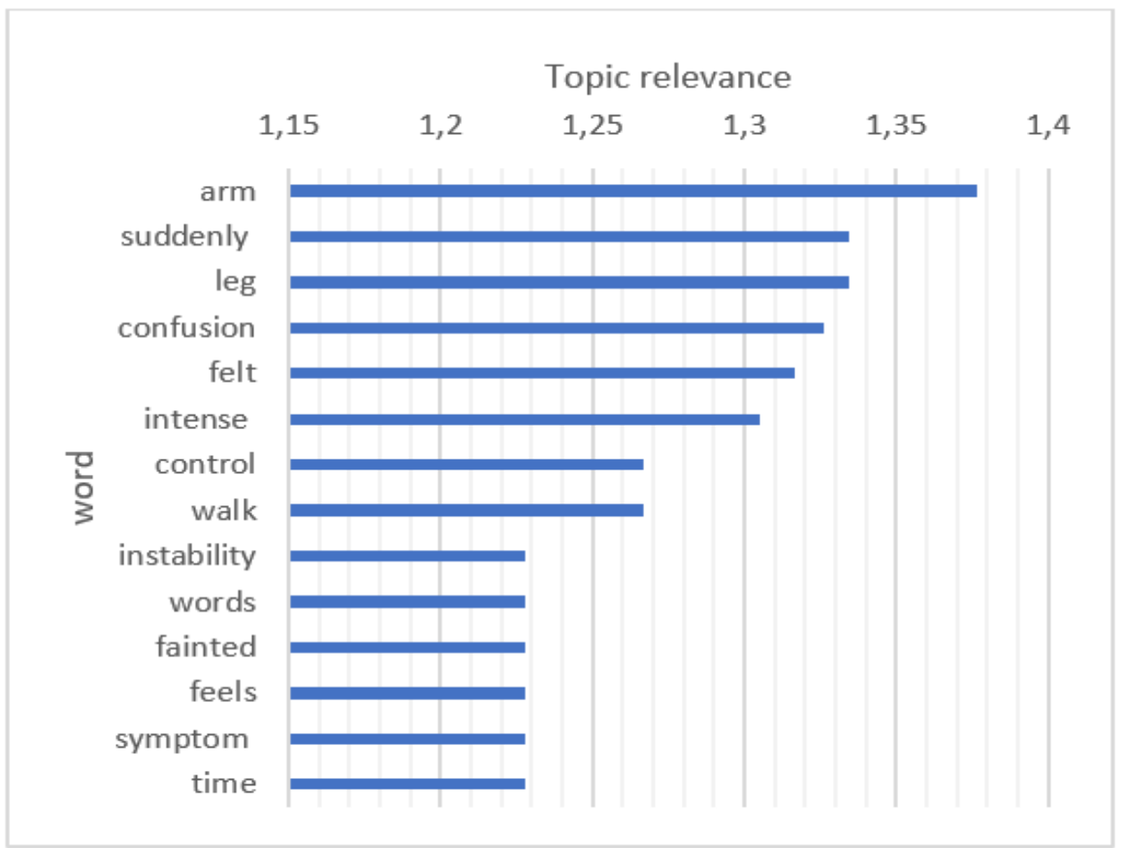

Figure 1 Topic 1

Word	N	Word	N	Word	N	Word	N
speech	43	incident	8	pain	5	body	3
weakness	24	stroke	7	felt	5	dizziness	3
arm	18	confusion	7	after	5	ability	2
felt	14	transferred	6	sense	5	headache	2
hospital	11	paralysis	6	symptom	4	surgery	2
leg	11	started	5	lost	4	claims	2
remember	9	numbness	5	followed	4	instability	2
intense	9	reports	5	sickness	4	alone	2
suddenly	8	eye	5	problem	4	immediately	2

Table 1 Single word frequency (N = number of times a word exists in dataset)

Word	N	Word	N	Word	N
inability talk	52	weaken end	8	that fell	4
weakness arm	26	walk well	8	mentions is	4
can not	24	incident started	4	eye weaken	4
transferred hospital	18	arm started	4	spastic movement	4
speech confusion	12	first symptom	4	suddenly felt	2
started weaken	10	had control	4	lose consciousness	2
didn’t have	8	started feeling	4	senses arm	2
clamp lightly	8	leg started	4	profound weakness	2
arm leg	8	according saying	4	could still	2

Table 2 Double word frequency

Conclusions:

Results from the history intake for individuals even during rehabilitation do support existing research that suggests the two most common stroke symptoms are "arm/speech." This is important for understanding human cognitive rehabilitation also. Our results show that the pre-experience event (episodic memory) of FAST symptoms are maintained in mental representations as seen in discourse analysis during rehabilitation. These verbal memories of symptoms can affect post-stroke life span. Also, we show that Artificial Intelligence techniques such as machine learning can be efficient in analyzing, categorizing and comprehending text of aphasic patients

References:

1. Aroor, S., et al. (2017). BE-FAST (Balance, Eyes, Face, Arm, Speech, Time) Reducing the Proportion of Strokes Missed Using the FAST Mnemonic. *Stroke*, *48*(2), 479-481.
2. Blei, D. M., et al. (2003). Latent dirichlet allocation. Journal of machine Learning research, 3(Jan), 993-1022.
3. Salton, G., & Buckley, C. (1988). Term-weighting approaches in automatic text retrieval. Information processing & management, 24(5), 513-523.