How long is long? An MEG study of M100 response in tracking word length in Bangla/Bengali

Swarnendu Moitra[1], Dustin A. Chacón[2], Linnaea Stockall[3]
[1,3][Queen Mary University of London], [2][New York University Abu Dhabi, University of Georgia],
[1]s.moitra@qmul.ac.uk, [2]dustin.alfonso@nyu.edu, [3]l.stockall@qmul.ac.uk

1 Introduction

Electrophysiological responses demonstrate an early effect of word length which localizes to posterior left fusiform gyrus around 100ms (M100). Stem-to-whole word transition probability (TP) evokes a later response around 170ms (M170) in anterior left fusiform gyrus (Pylkkänen and Okano, 2010; Grainger et al., 2008; Rastle et al., 2000), suggesting rapid lexical and morphological analysis. Anterior left fusiform gyrus is also known as the visual word form area (VWFA). The M100 is usually taken to reflect lower level visual processing, while the M170 has been interpreted as reflecting early stages of morphological analysis. However, most research on these responses has focused on languages written in alphabetic scripts, with one character roughly corresponding to one phoneme. How universal are the M100/M170 responses, and does the M100 also reflect linguistic analysis, or a psychophysical response to visual complexity? We report on an MEG study on Bangla, a language with an abugida script. In abugidas, one character corresponds to a consonant. Vowel ligatures are written on either side of the consonant, above, or below it. Some consonants have an unwritten "implicit" vowel, allowing some consonant clusters to be represented as a single complex character. Consequently, there are several distinct ways to quantify word length, e.g. প্র(prô) consists of 1 grapheme, 2 glyphs: প(p) and ্র(r), and 3 phonemes: /p/,/ɹ/,/o/. We ask two questions: (1) Which measure of word length best correlates with the M100 response? and (2) Is a stem-to-whole word TP-effect observed in the M170 response? Preliminary results suggest that M100 responses reflect the number of phonemes in the word, suggesting that the M100 indexes rapid abstract linguistic analysis.

2 Methods

MEG recordings were obtained from 22 Bangla speakers (18-62 years,$\bar{X}$=28) at New York University Abu Dhabi. The participants performed a Visual Lexical Decision Task, lying in a dimly-lit magnetically shielded room while MEG data were recorded continuously using a 208-channel axial gradiometer system (Kanazawa Institute of Technology, Kanazawa, Japan). All were literate in standard Bangla which included native dialects from West Bengal, India

and Bangladesh. All participants were recruited either from NYU Abu Dhabi, or the UAE Bangla speaking community and reported normal or corrected vision, normal hearing, and no history of neurological disorders. Written informed consent was provided by all participants prior to the experiment and were compensated for their time. The experiment reported here is a subset of a larger design which investigated the neural differences between morphosyntactic processing and morphosemantic processing. We are reporting on the stimuli which consisted of 152 morphologically complex grammatical words and 152 pseudowords (See table 1). Three different measures of length were calculated from indicNLP (966 million words)(Kunchukuttan, 2020): number of graphemes, number of glyphs (unicode derived, nchar() fn in R), and number of phonemes (native speaker annotated). TP was estimated as the log of lemma frequency to stem frequency.

3 Results

600ms epochs were extracted for each word. We used a two-stage regression analysis in which regressions were fit at each time-point and source point per subject for factors of TP and word length. There were six regressions – TP + word length for 3 definitions of word length in two separate time windows. Spatio-temporal cluster-based permutation tests were conducted on the one-sample t-test values derived from the beta coefficient of the regressions in left fusiform gyrus in the M100 (100 - 130ms) and M170 (170 - 200ms) time window. After correction for multiple comparisons, only one word length cluster was significant ($p = 0.01$), corresponding to the number of phonemes in posterior left fusiform gyrus from 100ms-130ms. One TP cluster was significant and negatively correlated in anterior left fusiform gyrus from 170ms-200ms ($p = 0.04$). No other significant clusters sensitive to any other word length measures were identified.

4 Discussion

Early visual responses to words suggest rapid analysis of abstract linguistic structure. Our results leverage the complexity of an abugida system to demonstrate that the M100 response is modulated by the number of phonemes in a writing system without a clear one-to-one phoneme-grapheme correspondence, and that the M170 reflects stem-to-word TP, consistent with results for other languages (Wray et al., 2022).

Prefix	Pseudowords		Grammatical	
	SemViol	**CatViol**	**Gramm**	**Fillers**
prôti	*প্রতিরক্ত *prôti-rɔktô* prôti-blood 'trans-blood'	*প্রতিনীল *prôti-nil* prôti-blue 'trans-blue'	প্রতিহিংসা *prôti-hiṁsa* prôti-violence 'revenge'	মানুষজন *manuṣ-jôn* people-CLF 'people'
duḥ / dur	*দুর্নাক *dur-nãk* duḥ-nose 'bad-nose'	*দুঃকালো *dus-kalô* duḥ-black 'bad-black'	দুর্ঘটনা *dur-gʰɔṭôna* duḥ-event 'accident'	গাজরগুলো *gajôr-gulo* carrot-PL 'carrots'

Table 1: Example of Stimuli

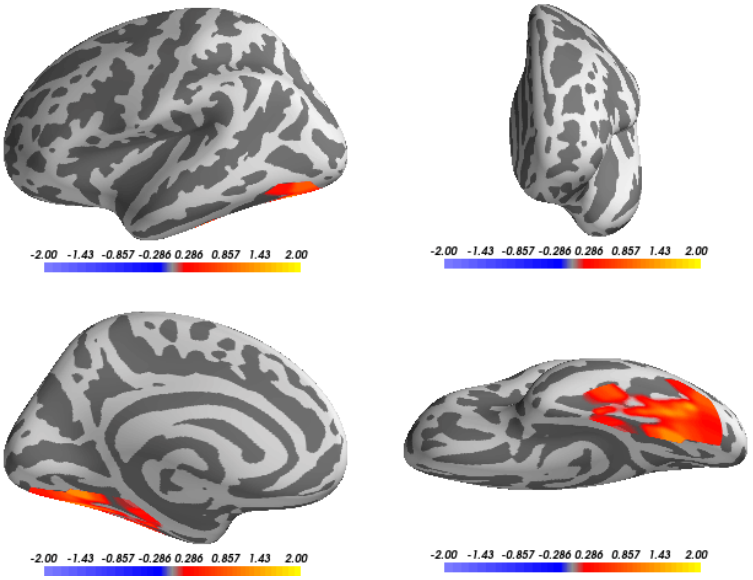

Fig 1. Posterior Left Fusiform Gyrus cluster M100 (100-130ms)

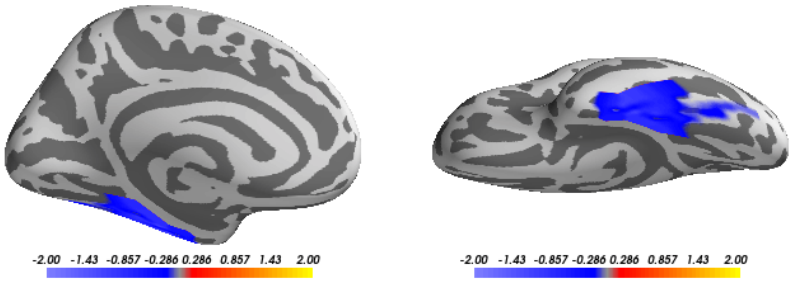

Fig 2. Anterior Left Fusiform Gyrus cluster M170 (170-200ms)

References

Grainger, J., Rey, A., & Dufau, S. (2008). Letter perception: From pixels to pandemonium. Trends in cognitive sciences, 12(10), 381–387.

Kunchukuttan, A. (2020). The indicnlp library. Indian language NLP Library.

Pylkkänen, L., & Okano, K. (2010). The nature of abstract orthographic codes: Evidence from masked priming and magnetoencephalography. PloS one, 5(5), e10793.

Rastle, K., Davis, M. H., Marslen-Wilson, W. D., & Tyler, L. K. (2000). Morphological and semantic effects in visual word recognition: A time-course study. Language and cognitive processes, 15(4–5), 507–537.

Wray, S., Stockall, L., & Marantz, A. (2022). Early Form-Based Morphological Decomposition in Tagalog: MEG Evidence from Reduplication, Infixation, and Circumfixation. Neurobiology of Language, 3(2), 235–255.