***Dear Fellow ECoG Researcher,***

These files form a set of speech data for lists of nouns that were first read directly, and then used to produce associated action verbs. There were 2 noun lists, and each list was presented 3 times for noun reading, then 3 times for verb production. The experiment was meant to test working memory and cognitive load in the language system.

Please keep in mind that these anonymized data are from real patients who donated time in a difficult period of their lives to advance our understanding of the brain. Any publication involving these data **MUST** include the following in the methods section of the manuscript, without modification:

**Ethics statement:** All patients participated in a purely voluntary manner, after providing informed written consent, under experimental protocols approved by the Institutional Review Board of the University of Washington (#12193). All patient data was anonymized according to IRB protocol, in accordance with HIPAA mandate. It was made available through the XXXXName of database/manuscriptXXX repository [insert citation here], freely available at XXXURL.

Note that the audio traces of patient speech are not included so as to preserve their anonymity. Patients participated in a simple verb-generation task, where written nouns (approximately 2.5 cm high, and 8–12 cm wide) were presented on a screen approximately 1 m from the patient, at the bedside. In between each 1.6-second cue was a blank-screen 1.6-second interstimulus interval. The experiment consisted of 12 distinct experimental runs, grouped into 4 phases:

***Phase 1:*** For the first portion of the task, the patient read the 1st list of 40 words (“LIST 1” reproduced below and also present in the file titled ‘list1.mat’). This list was presented for 3 consecutive experimental runs in succession.

***Phase 2:*** For the second portion of the task, the patient was again presented with the same 1st list of 40 nouns (“LIST 1”) and instructed to produce a verb that was connected to the noun (it could be something they might do to that noun, or that the object/animal might do itself): for example, if the cue read “ball”, the patient might say “kick”, or if the cue read “bee”, the patient might say “fly”. This list was presented for 3 consecutive experimental runs in succession. The patients were neither instructed to state the same verbs, nor to generate new ones during each repetition. When they asked, they were instructed to do whatever they preferred. Patients repeated the same verb associated with each noun in the vast majority of cases. Because the verbs were initially generated during the first run, and later repeated during the second run, one might assume that (1) cognitive load diminishes during each verb run and (2) there is a working memory component of the task during the 2nd and 3rd task runs.

***Phase 3:*** Noun reading as in phase 1, but reading nouns for the 2nd list of 40 nouns (“LIST 2” reproduced below and also present in the file titled ‘list2.mat’).

***Phase 4:*** Verb generation as in phase 2, but for “LIST 2”.

The basic datafiles (in MATLAB format) are named “##\_nouns\_LX\_RY.mat” and “##\_verbs\_LX\_RY.mat” in the folders data/##, where ## denotes the 2 letter patient code, X denotes the list number (1 or 2), and Y denotes the run number (1-3). This code is not the patient’s initials.

To explicitly avoid any confusion -- the order of 12 task runs, and therefore the chronological sequence of data collection, is:

1 - “##\_nouns\_L1\_R1.mat”,

2 - “##\_nouns\_L1\_R2.mat”,

3 - “##\_nouns\_L1\_R3.mat”,

4 - “##\_verbs\_L1\_R1.mat”,

5 - “##\_verbs\_L1\_R2.mat”,

6 - “##\_verbs\_L1\_R3.mat”,

7 - “##\_nouns\_L2\_R1.mat”,

8 - “##\_nouns\_L2\_R2.mat”,

9 - “##\_nouns\_L2\_R3.mat”,

10 - “##\_verbs\_L2\_R1.mat”,

11 - “##\_verbs\_L2\_R2.mat”,

12 - “##\_verbs\_L2\_R3.mat”,

Note that patient ZT only performed the first half of the task (e.g. “LIST 1” runs only).

Each “##\_verbs.mat” or “##\_nouns.mat” datafile has 4 variables:

“stim” (timex1): Timing of visual cues to read noun, or produce associated verb. The values 1-40 correspond to the 40 words in each list. These lists are saved in the files ‘list2.mat’ of the main directory.

"data" (time x number of channels): These are the data.

* sampled at 1000Hz
* scale factor: 1 amplifier unit = .0298 microvolts
* built-in band pass 0.15 to 200 Hz,

- but a 1 pole band pass, so there is no sharp corner at 200Hz.

-The amplitude roll-off function is in the file “ns\_1k\_1\_300\_filt.mat”

In the folder stim/, there are scripts titled “##\_stim.m”, where ## denotes the 2 letter patient code. When called, they produce two variable with information about electrocortical stimulation (ECS) result for language.

“stimsites” (1xN): “N” channels which were stimulated as part of a stimulation pair.

“ecssites” (Mx2): Channel pairs where ECS produced interruption of naming during clinical mapping.

Patients WC and JC also had ~3 minutes of baseline fixation data. This is stored in files called “##\_base.mat” in the folders data/##, where ## denotes the 2 letter patient code.

"data" (time x number of channels): These are the data.

* sampled at 1000Hz
* scale factor: 1 amplifier unit = .0298 microvolts
* built-in band pass 0.15 to 200 Hz,

- but a 1 pole band pass, so there is no sharp corner at 200Hz.

-The amplitude roll-off function is in the file “ns\_1k\_1\_300\_filt.mat”

Patient brain surfaces were rendered from, and electrodes were localized with either CT or X-Ray. In some cases, MRIs were sliced thickly (~6mm), so surface quality was not good. This MRI-CT localization approach (CTMR) is described in *“Automated electrocorticographic electrode localization on individually rendered brain surfaces” by D Hermes, et al in Journal of Neuroscience Methods, 2009*. The MRI-X-Ray localization approach (XRMR) is described in *“Brain surface electrode co-registration using MRI and x-ray.” by Miller, et al in IEEE Eng Med Biol Soc, 2010*. The individualized MRI anatomy files are in the folder brains/, with files named – “##\_brain.mat”, with variables:

* + "brain": This is a structure representing the tessellated brain surface. It can easily be plotted with the CTMR package *(see “Automated electrocorticographic electrode localization on individually rendered brain surfaces” by D Hermes, et al in Journal of Neuroscience Methods, 2009)*
  + "locs" (number of channels x 3): Electrode locations, for plotting on the rendered brain.

**LIST 1:**

'JET'

'LAWN'

'MOVIE'

'FLAG'

'LADDER'

'BELL'

'RULER'

'FINGER'

'LAKE'

'DOLL'

'BIKE'

'CIGAR'

'FORK'

'BATON'

'CAR'

'DOLLAR'

'DOOR'

'SEED'

'RADIO'

'HAMMER'

'CANE'

'GIFT'

'SHIRT'

'KNIFE'

'MILK'

'BOOK'

'TOWEL'

'BALL'

'FAN'

'WHEEL'

'GRAVE'

'GUM'

'SONG'

'CAT'

'STOVE'

'PEN'

'OAR'

'BENCH'

'RIFLE'

'YARN'

**LIST 2:**

'SOAP'

'GUN'

'JOB'

'SCHOOL'

'CRAYON'

'BOAT'

'FOOT'

'PURSE'

'LENS'

'BEER'

'PLANE'

'BLANKET'

'BEE'

'DOG'

'LETTER'

'SCALE'

'FIST'

'BED'

'FOOD'

'MATCH'

'PILL'

'HOSE'

'CHAIR'

'TOY'

'TREE'

'NEEDLE'

'OVEN'

'MONEY'

'PHONE'

'PAPER'

'KEY'

'BROOM'

'ICE'

'BRICK'

'HORN'

'FIRE'

'POOL'

'BASKET'

'RAZOR'

'GLOVE'

Best Wishes!

Kai Miller,

Stanford University, 2015

[kai.miller@stanford.edu](mailto:kai.miller@stanford.edu), [kjmiller@gmail.com](mailto:kjmiller@gmail.com)