

README

Welsh Auditory Masked Priming Experiment

The research team on NSF Grant 1453724:
Experimental and Descriptive Investigations of Welsh Consonant Mutation

September 18, 2018

1 Associated files

- 1 raw, anonymized data file (`results_for_archive.csv`);
- 128 WAV files; each is a unique stimulus in the experiment consisting of the masked prime and target;
- 1 file indicating what words occurred as masks for which prime-target pairs (`master_input.tsv`).

2 Design of the experiment

Participants judged the lexicality of 64 real Welsh words (all masculine) and 64 phonotactically-legal Welsh non-words. Real word and non-words were all presented with a preceding monosyllabic function word (either *y* [ə] ‘the’ or *ei* [i] ‘his’), as were their associated masked primes. Real word targets occurred in the following priming conditions:

- **Unrelated:** *ei baith* ‘his desert’ primed by *ei wên* ‘his smile’.

Prime and target share no phonemes and are not morphologically or semantically related. The function word preceding both prime and target was always the same for a given stimulus.

- **Phonologically related:** *ei baith* ‘his desert’ primed by *ei daith* ‘his journey’.

Prime and target share all but the first consonant, but are not morphologically related and the correspondence between the two initial phonemes does not constitute a possible mutation relationship.

- **Morphologically related:** *ei baith* ‘his desert’ primed by *y paith* ‘the desert’ or *y paith* primed by *ei baith*.

Prime and target are morphologically related via Soft Mutation, e.g. one is the unmutated form of a word and the other is the Soft Mutation form of the same word. They have the same amount of phonological overlap as the phonologically related prime-target pairs in the word itself, but the prime and target differ in the preceding function word (one occurs with *y* and one with *ei*).

- **Identity:** *ei baith* primed by *ei baith* ‘his desert’.

Prime and target are identical (as are the preceding function words).

Four lists were constructed in a Latin Squares design such that all participants saw all real word targets and all real word targets occurred in every priming condition but each subject responded to each word only once. Half of the real word targets always occurred with the mutation triggering function word *ei* ‘his’ in all lists and half always occurred with *y* ‘the’.

All non-word targets had unrelated real word primes; half of them occurred with the preceding function word *ei* ‘his’ and half with *y* ‘the’.

3 Contents of the data file

- **subject:** A unique identifying code for each participant;
- **gender:** Participants’ gender (m = male, f = female);
- **item:** A unique identifying code for each stimulus;
- **trial:** Integer ranging from 1–128, indicates the order in which stimuli were presented (which was fully randomized for each subject);
- **rt_raw:** Reaction time in milliseconds measured from stimulus onset; negative values indicate incorrect button responses (and correspond to 0’s in the accuracy column). For actual reaction time take the absolute value of this column;
- **stimulus:** Corresponds to the name of the WAV file responded to on each trial. The stimulus names contain the following information: The target, the prime, lexicality of the target (R = real, N = nonce), lexicality of the prime, priming condition (M = morphologically related, PC = phonologically related, ID = identity, UR = unrelated), mutation status of the target (M = mutated, U = unmutated, F = filler, e.g. non-word trials);

- **targetOnset**: Time in milliseconds between the onset of the stimulus and the onset of the target word;
- **targetOffset**: Time in milliseconds between the onset of the stimulus and the offset of the target word;
- **duration**: Duration in milliseconds of the target word;
- **accuracy**: Accuracy of button response (1 = correct, 0 = incorrect);
- **list**: Factor indicating the experimental list (1, 2, 3, 4). Stimuli for non-word trials were the same across all lists and therefore have NA values in this column.
- **prime**: Priming condition (ur = unrelated, id = identity, mr = morphologically related, pc = phonologically related). See section ?? for a detailed explanation and example prime-target pairs;
- **lexicality**: Lexicality of the target word;
- **target**: A code identifying the target word (this is unique for real word targets, ranging from 01–64; the non-word targets also have target numbers ranging from 01–64, so if you wish to analyze real and non-words together you should be aware that the current target codes are not unique across real and non-words);
- **targetMut**: A binary factor indicating whether the target word appeared in mutated form with the preceding morpheme *ei* [i] or unmutated form with the preceding morpheme *y* [ə]. Half of the non-word targets appeared with the former and half with the latter (non-words appearing after the mutation trigger *ei* were consistent with soft mutation).
- **age**: Integer indicating participants' ages in years.
- **hist_welsh**: The numerical score computed from the BLP for participants' history with Welsh;
- **hist_eng**: The numerical score computed from the BLP for participants' history with English;
- **use_welsh**: The numerical score computed from the BLP for participants' usage of Welsh;
- **use_eng**: The numerical score computed from the BLP for participants' usage of English;
- **profic_welsh**: The numerical score computed from the BLP for participants' proficiency in Welsh;

- **profic_eng**: The numerical score computed from the BLP for participants' proficiency in English;
- **tude_welsh**: The numerical score computed from the BLP for participants' attitude toward the Welsh language;
- **tude_eng**: The numerical score computed from the BLP for participants' attitude toward the English language;
- **tot_welsh**: Participants' composite Welsh language score;
- **tot_eng**: Participants' composite English language score;
- **dominance**: Participants' aggregate language dominance score.

4 Language dominance scores

The Bilingual Language Profile (BLP) was used to collect and compute language dominance scores. Please consult documentation on this tool for details of how the dominance scores are computed.

5 A few notes on participants

- Demographic information that could lead to the identification of participants has been excluded from the raw data file.
- During the debrief phase participant 30 identified a handful of the masked primes, indicating some level of prime awareness.
- Due to an error during data collection, language dominance scores cannot be definitely associated with Subjects 49, 50, and 51. For that reason these subjects have "NA" values for the language dominance variables.