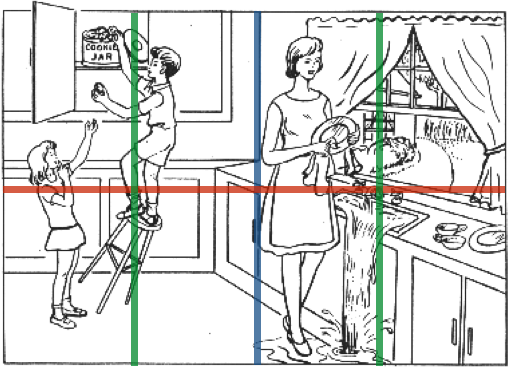
**Automatic Detection of Alzheimer’s from Speech using Hemispatial Neglect Markers**

**Background:** Machine learning has been used to diagnose Alzheimer’s disease from speech samples of patients undergoing the “Cookie Theft Test” component of the Boston Diagnostic Aphasia Exam. Previous work has achieved an accuracy of 81.92% using lexical and acoustic markers extracted from patient interviews. We propose and evaluate three sets of hemispatial neglect markers.

**Methods:** We use the publically available DementiaBank dataset and as baseline extract 353 lexical and acoustic markers (derived from previous work) from 499 interviews (257 AD). We evaluate three divisions of the Cookie Theft image: *Halves*, *strips* and *quadrants*, as seen in figure 1. For a given division, we compile a list of *information units* (info-units) that are contained in each region (e.g.: “stool” and “mother” are contained by the left and right halves, respectively). For each region we then record four measures from a given transcript: 1) Number of info-units mentioned 2) ratio of info-units to all words 3) ratio of unique info-units to all possible info-units in the region 4) ratio of unique info-units to total mentioned info-units (a measure of redundancy). We also include quadratic interaction terms between regions. We then perform a 10-fold cross validation procedure with a correlation-based feature selection preprocessing phase and train a logistic regression model with the three sets hemispatial neglect markers, and compare against baseline.

**Results:** The halves model achieved the best accuracy (0.842 ± 0.061), beating the baseline by 2% (0.822 ± 0.05). The strips model (0.829 ± 0.057) also beat baseline but fell short of the halves model by 1.3%. The quadrants model (0.813±0.070) had the worst accuracy and largest variation of the four models.

**Conclusion:** Including halves hemispatial neglect markers (and their quadratic interaction terms) improves Alzheimer’s classification accuracy by 2 percent.

****

**Figure 1:** Cookie Theft image showing three spatial divisions. “Halves” divides the image in two along the blue line, “quadrants” divides the image along the red and blue lines, and “strips” divides the image along green and blue lines. Here we see the “mother” info-unit in the right half, centre-right strip, and in the NE and SE quadrant. Models are trained with one of the divisions.