

## Background and objective

Although the P300 spelling grid has been modestly successful even for patients, the visual flashes that are needed are not accessible for those with vision deterioration. For these potential users, tactile or auditory stimulation are the modalities of choice, but their implementations are mostly binary in nature. In [1,2] we presented the offline results for a new auditory multi-class paradigm to overcome some of these difficulties. Here we present the results from an online trial with 21 healthy subjects.

## Methods

### AMUSE Paradigm

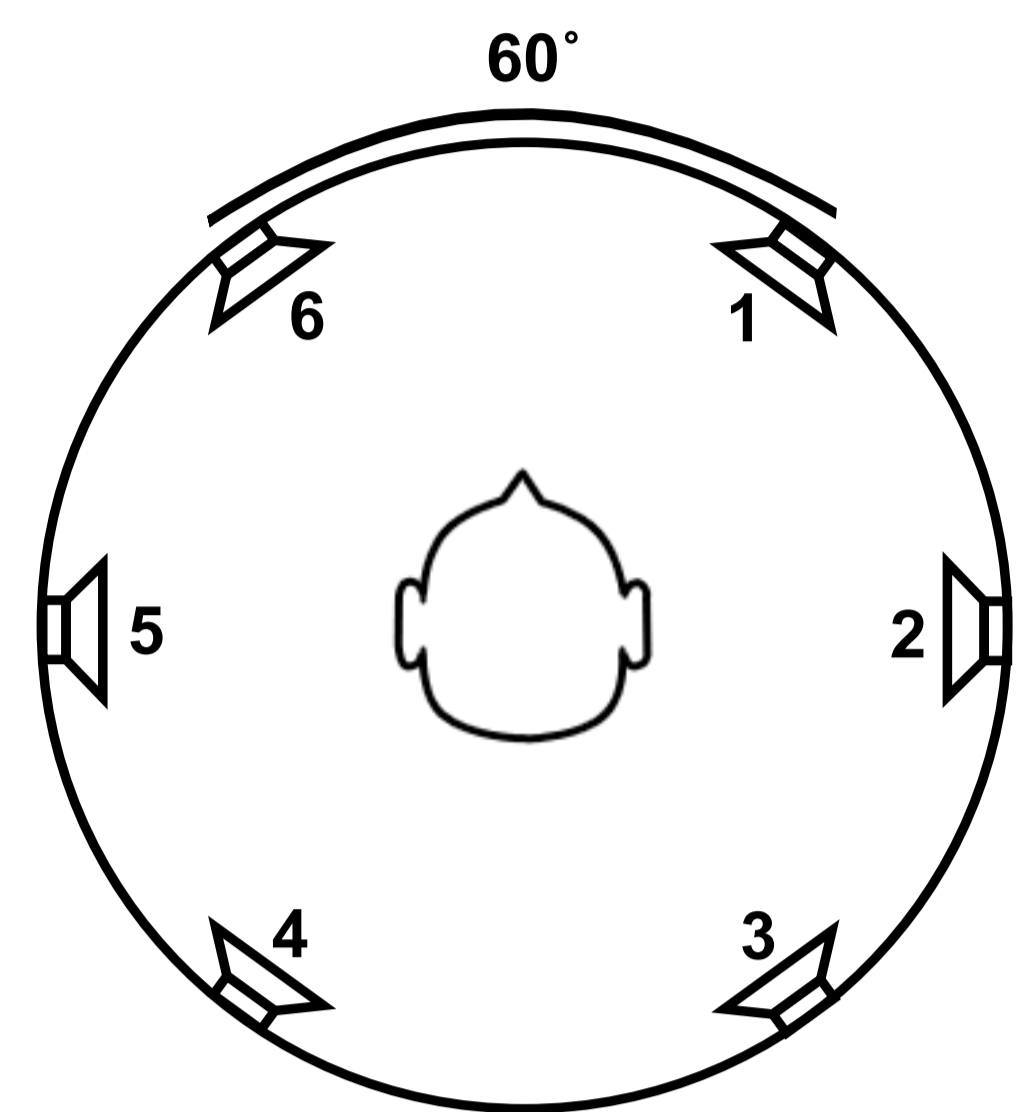


Figure 1. Experimental setup

The paradigm, as presented in [1], was adapted for online text writing. It consists of a ring with 6 audio speakers around the subject (Figure 1) which produce stimuli in a pseudo random order. Focusing attention to any of these spatial directions resulted in the reliable elicitation of P300 signals in the EEG after a stimulus from that direction. This effect was used for driving a BCI. This paradigm is dubbed Auditory MULTiclass Spatial ERP, or AMUSE.

An adapted version of the hex-o-spell speller [3] was created in which characters can be selected in a two step process. First a group of letters is selected (example: 'A-E') by focusing on the corresponding direction. In the second step the characters are divided over five of the directions and an individual letter can be selected. Choosing the sixth direction returns the user back to the first selection step.

## Experiments

A study was performed with 21 naïve, healthy subjects in two sessions. Each session consisted of a calibration part (30 min) and a spelling part. Six subjects did not gain sufficient control in the first session and were excluded from the second session. One subject did not manage to write in the second session.

### Session 1

Using a fixed number of 15 iterations, subjects are asked to write one of two pangrams (30 and 33 letters). Erroneously written letters had to be corrected by the user by performing a backspace action. Which (group of) letter (s) corresponds to a direction is read to the user before each selection. These labels are generated dynamically by speech synthesis technology. Label-, stimulus- and result presentation for a single selection took around 34 seconds. The maximal theoretical speed is thus .89 char/min, with a 2 step letter selection process.

### Session 2

To increase spelling speed, the labels were no longer read but subjects learned them by heard. Also, a dynamic stopping method was implemented to limit the number of iterations. Subjects were asked to write two pangrams. Erroneously written letters had to be corrected by the user by performing a backspace action. When considering 15 iterations for a selection, stimulus- and result presentation took around 27 seconds. The maximal theoretical speed is thus 1.10 char/min.

## Dynamic stopping criterion

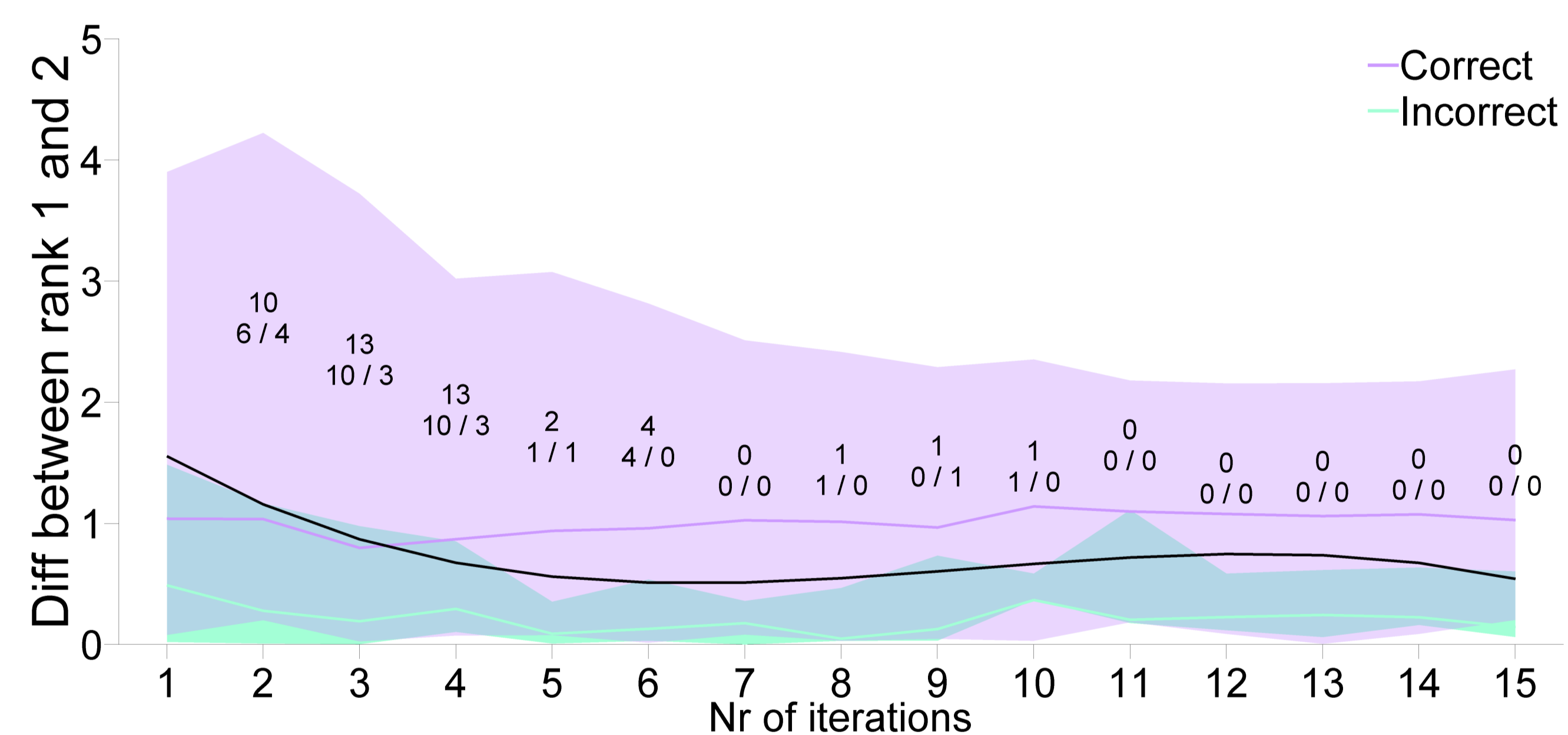


Figure 2. Estimation of stopping criteria based on the calibration data.

We define the target direction as the direction with the lowest median classification score over all performed stimulations. Evidence for true target direction is accumulated with each iteration. For iterations 4 to 15, an early stop could be elicited if there is already high evidence for a target. High evidence is expressed as less than 5% chance of a false positive decision (based on calibration data). For this we used a *difference rank criterion (DRC)*, which is the difference between the lowest and the second lowest ranked direction.

### Heuristic to determine thresholds for the DRC for early stopping

An early stop should be made with less than 5% probability of a false positive. Therefore two curves are estimated:

(a) For every iteration, the 95-percentile of the incorrect decisions is estimated. A polynomial curve of order 3 is fitted to these values (smoothing). The resulting polynomial is shown in Figure 2 as a black line.

(b) The median (50-percentile) values of the correct decisions is estimated (ilac line in Figure 2)

Strictly speaking (a) is the decision boundary. However, due to a low number of samples this is difficult to estimate reliably. Therefore, the threshold for the DRC for every iteration is fixed to the larger value of (a) and (b).

## Results

### Physiology

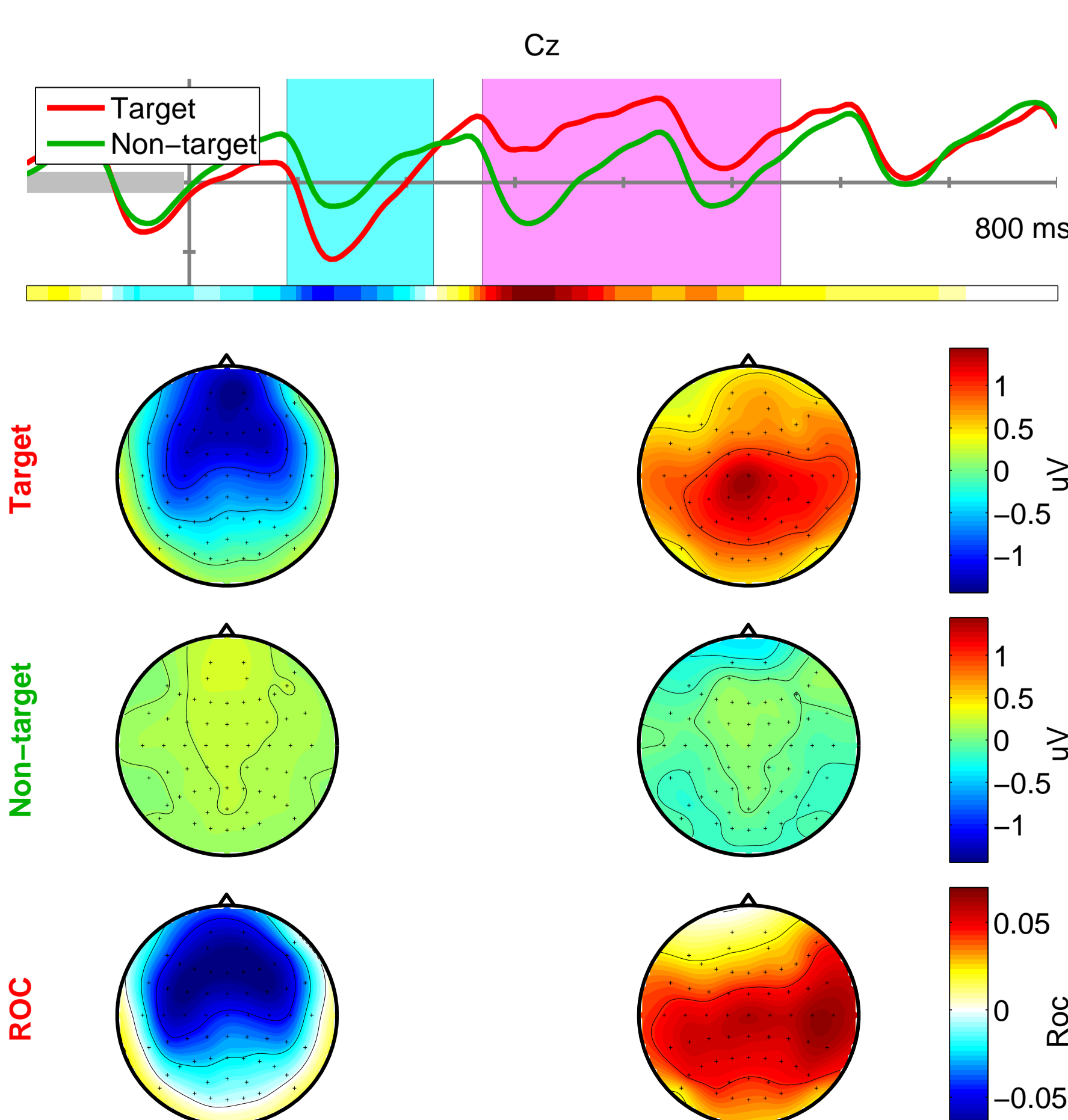


Figure 3. Grand average ERP (top) and scalp maps (left) for calibration data of session 1. Grand average ERP of the calibration data show that both an early component—frontal areas, and bilateral over auditory perception areas—and a late component—located over the midline, with peak values over Cz electrode—were attention modulated. Scalp maps (left) are averages over the two intervals marked in the ERP plot (top). The ROC scalp map (left, bottom) gives an estimate of the separability of both classes.

## Spelling performance

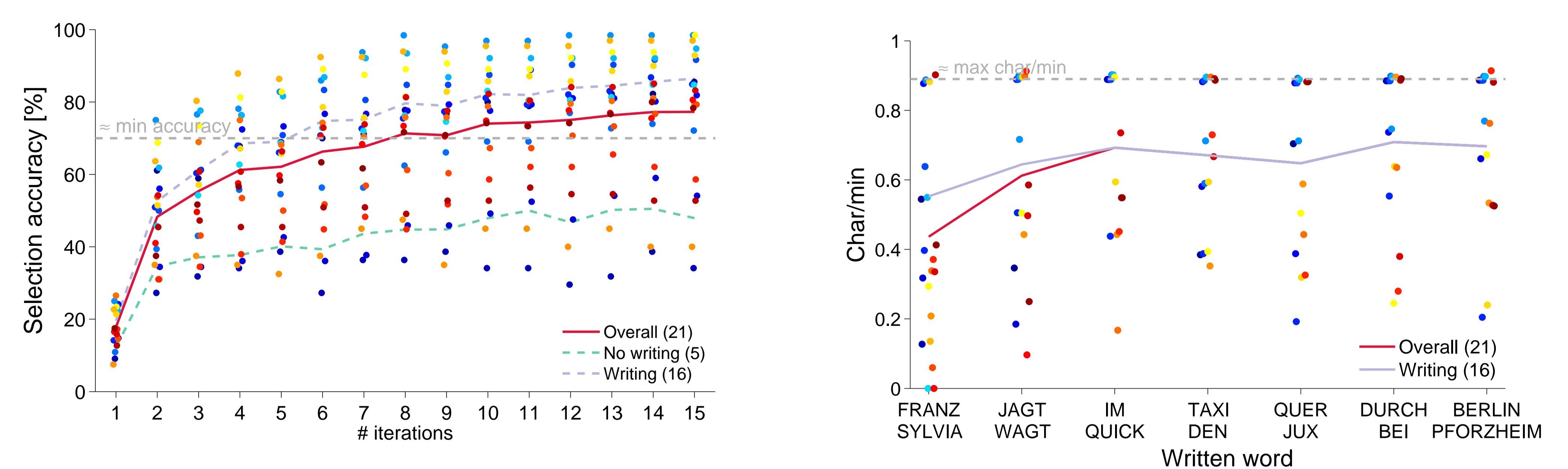


Figure 4. **Left:** selection accuracy plotted against the number of iterations, for the online data of session 1. Single subject values are represented by colored dots. Most subjects quickly reach an accuracy above 70% which is a minimum for this application. **Right:** spelling performance in char/min for individual words written in the first session. As words are plotted chronologically, the X axis represents time (increasing to right). The relative flat learning curve ("Writing (16)") matches the subjective reports of the subjects that the application (adapted Hex-o-spell [3]) was intuitive to use.

## Session transfer and improvements

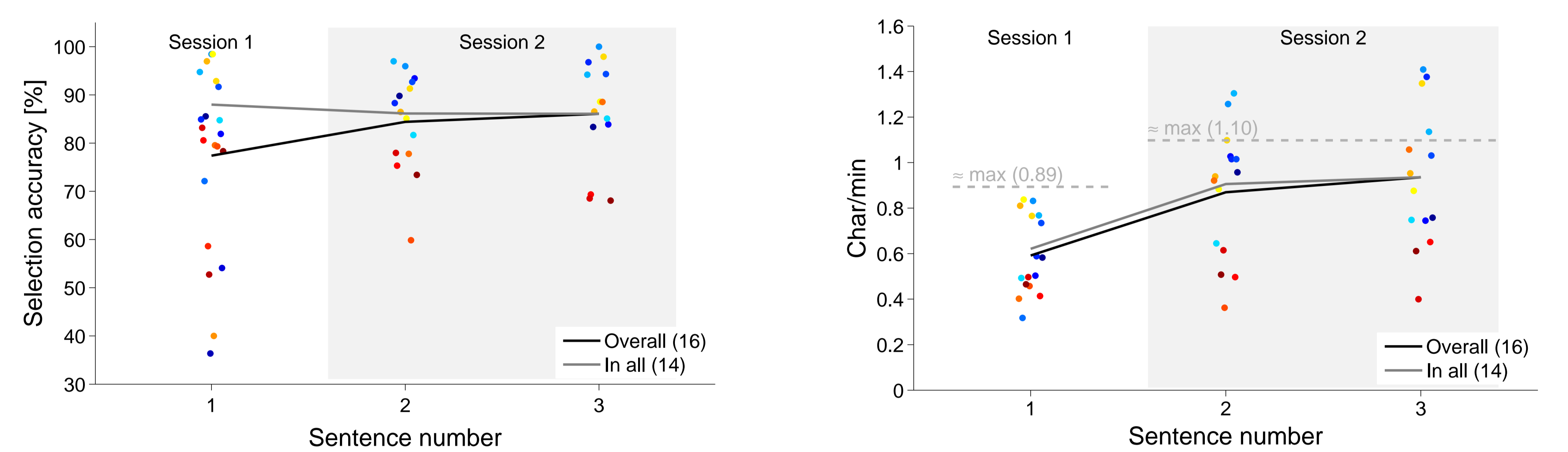


Figure 5. **Left:** selection accuracy for subject plotted against sentence number (session). When only considering the subjects that successfully wrote all sentences, the performance is stable over sessions (line "In all (14)"). **Right:** char/min plotted against sentence number (session). In session 2 the max char/min value is increased by removing the label presentation, thus leading to a performance increase. Further increase in performance is the result of introducing the dynamic stopping heuristic, with multiple subjects reaching scores higher than 1.1 char/min.

## Early stopping

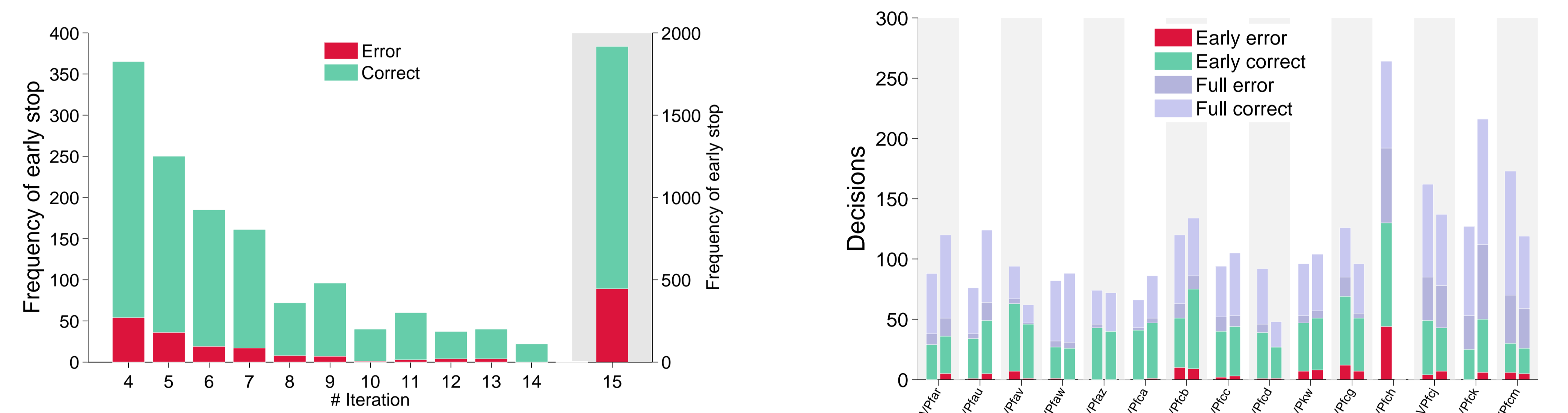


Figure 6. **Left:** frequencies of early stops for each number of iteration (grand average). The number of erroneous early stops is low (red bars) and linearly correlated to the total number of early stops at each iteration. As every trial is ended latest at iteration 15, the number of stops is a magnitude bigger (see right axis). **Right:** number of decisions per sentence for subjects in the second session. Four categories of selection are plotted, errors and correct decisions (both with- or without early stopping).

## Discussion

The spatial auditory paradigm (AMUSE), as presented offline in [1-2], was shown to be successful in an online setting. From 21 subjects, 16 were able to write a full, German sentence in the first session after only 30 minutes of calibration. Users had no prior experience with BCI, AMUSE, or the adapted version of Hex-o-spell [3].

The performance (as measured in char/min) is relatively low compared to visual paradigms. However, as AMUSE can be used strictly auditorily it may better serve the need of late-stage ALS patients with vision deterioration. With an average performance of 0.90 char/min in the second session, AMUSE outperforms most state-of-the-art auditory BCIs.

Disregarding labels and a simple, but effective early stopping method improved the average performance in session 2. As seen in Figure 6-left, the number of erroneously early stopped trials is relatively small compared to the correct early stops. On an individual level this may be different. For subject VPfc the first sentence of the second session proved so difficult that the experiment was stopped.

Currently, experiments with an ALS patient are ongoing.

## Acknowledgments

This work was partly supported by the European Information and Communication Technologies (ICT) Programme Project FP7-224631 and 216886, by grants of the Deutsche Forschungsgemeinschaft (DFG) (MU 987/3-1) and Bundesministerium für Bildung und Forschung (BMBF) (FKZ 01B001A, 01GQ0850) and by the FP7-ICT Programme of the European Community, under the PASCAL2 Network of Excellence, ICT-216886. This paper only reflects the authors' views and funding agencies are not liable for any use that may be made of the information contained herein.

## References

- [1] E.M. Schreuder, M. Tangermann, B. Blankertz (2009) Initial results of a high-speed spatial auditory BCI. *Int J Bioelectromagnetism*, 11, pp. 105-109
- [2] M. Schreuder, B. Blankertz, M. Tangermann (2010) A New Auditory Multi-Class Brain-Computer Interface Paradigm: Spatial Hearing as an Informative Cue. *PLoS ONE* 5(4): e9813. doi:10.1371/journal.pone.0009813
- [3] B. Blankertz, G. Dornhege, M. Krauledat, M. Schröder, J. Williamson, R. Murray-Smith, K.-R. Müller (2006) The Berlin Brain-Computer Interface presents the novel mental typewriter Hex-o-Spell. *Proceedings of the 3rd International BCI Workshop and Training Course*, pp. 108-109