MindScribe: Innovation Report

Introduction & Background: When my mom's asthma flares, even a short sentence costs her breath. A close friend's grandmother was recently intubated near the end of life. She could not tell us if she was in pain or if she just needed water. Those moments made the gap obvious: in urgent care, people don't need paragraphs, they need a few essential words, spoken quickly with almost no effort. That insight shaped MindScribe. I studied work on silent-speech systems and eye-controlled communication. Many sEMG projects try to reconstruct full speech or support very large vocabularies with dense electrodes; strong in the lab, but slow to set up at a bedside. Other studies show small channel counts can still carry useful information, yet practical, repeatable placement and across-day stability are not always clear. Gaze typing needs many long fixations and can tire already-ill eyes; dry eye is common in ICU settings (*See App Z).* From this literature, I chose to avoid continuous eye control and use one voluntary blink only when the model is uncertain: a low-effort confirm step. **<u>My guiding question.</u>** <u>*What is the smallest useful word list and the smallest number of facial electrodes that still let a person say what matters, and can a single blink safely confirm only when needed?*</u>

Methods & Procedure: I standardized an 8-channel facial/submental montage with fixed landmarks, polarity, and pre-flight checks so signals are repeatable across days (muscles: ZYG, RIS, LLS, DAO, DLI, MAS, ABD, SLH) (*See App A).* Signals were recorded with an OpenBCI Cyton at 250 Hz; for every trial I saved a 2.0 s window that spans the "mind preview" and the silent articulation (*See App B, C).* Each trial followed one fixed loop: READY 1.0 s → preview 0.8 s → silent speak 1.2 s → REST 1.0 s, drawn from eight words (yes, no, help, pain, up, down, hmmm, blank). Targeted follow-ups produced 5,294 total trials with documented per-class counts (*See App C).* I filtered 20–120 Hz with 60/120 Hz notches, then cropped each word to an energy-peak segment so every example has the same length (*See App D).* For learning, I applied per-session z-score normalization and used a compact temporal model fed by a simple feature stack (raw, envelope, RMS). I evaluated generalization by holding out entire sessions (across-day test) with saved splits *(See App E)*. To minimize electrodes, I ran a greedy channel-ablation on the held-out test sessions and then retrained on Top-K. Zygomaticus (ZYG) and Risorius (RIS) ranked highest most often, with DLI/LLS as common next choices, evidence for a 3–4 electrode cheek-patch path (*See App F).* The demo GUI speaks automatically when confidence ≥ 0.70 and arms a single blink only when confidence is lower, keeping eye effort minimal for the design rationale.

Observations & Results: Using the fixed protocol at 250 Hz with a 2.0 s save window, I collected 5,294 trials with balanced timing and documented per-word count on 8 – channels (*See App B, C*). On a representative 5-word held-out-day test, the compact model achieved ≈ 0.90 accuracy (± 0.005) with similar macro-F1, showing the patterns generalize across days rather than only within one session *(See App E)*. Ablation confirmed that most of the predictive power sits on a few cheek channels: ZYG/RIS lead; DLI/LLS often complete effective Top-3/Top-4 sets, supporting the tiny cheek-patch prototype *(See App F, G)*.

Conclusion & Why It Matters: MindScribe shows that few electrodes and a small, high-value word list can restore urgent communication with almost no eye work. When the model is confident, it speaks right away; when it is not, one blink confirms. This design choice directly serves people who are intubated, short of breath, or fatigued people who need fast, clear words, not long interactions *(See App Z)*.

What's Next: I will finalize the cheek-patch using the best Top-3/Top-4 channels and mark quick re-attachment points (for nurses and home caregivers) *(See App A, F)*. I will tune class-specific confidence gates in the GUI (e.g., stricter for "silence") while keeping blink-to-confirm only for low-confidence cases. *(See App E, Z*). I will run a small user pilot with the patch to check comfort and across-day stability using the same trial protocol; the acquisition pipeline (board settings, filtering, cropping, exports) will remain unchanged for reproducibility. *(See App B, C, D)*.

Key Results

Channel	Muscle	5-word set: Yes, No, Pain, Hmm, Silence				8-word set: Yes, No, Pain, Hmm, Silence, Help, Up, Down				
		8 Channel Accuracy 90%				8 Channel Accuracy 74.60%				
1	Masseter		○	○	○		○	○	○	○
2	Zygomaticus	1	●	○	○	1	●	○	○	○
3	Anterior Belly of Digrastic		○	○	○	5	○	○	○	●
4	Depressor Labii Inferiorus		○	○	○	4	○	○	●	○
5	Levator Labii Superioris	3	○	●	○	3	○	●	○	○
6	Risorius	2	●	○	○	2	●	○	○	○
7	Depressor Anguli Oris	4	○	○	●		○	○	○	○
8	Stylohoid		○	○	○		○	○	○	○
			56.8%	77%	85.6%		49.3%	54.8%	63.8%	69.1%

Table 1: 8 and 5 word accuracies on 8 Channels and Channel Ablation Accuracies (Detailed Results in App G)

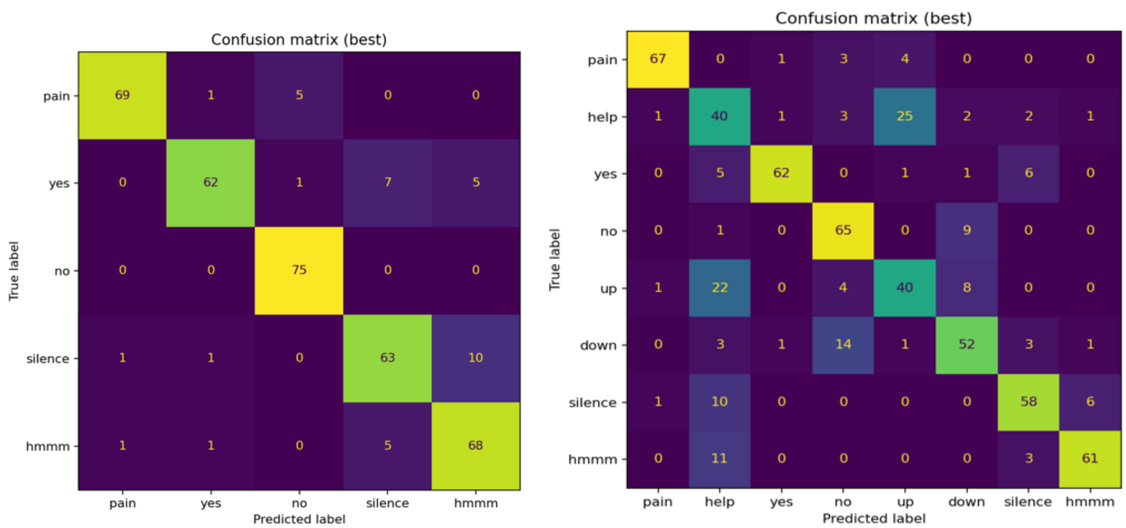
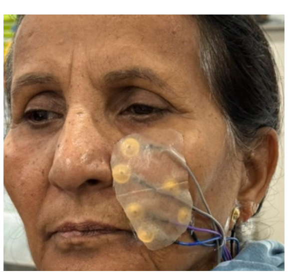


Figure 1 and 2: Confusion Matrices for 5 and 8 words.

Patch Design



This is a simple homemade version of a tiny 3 electrode patch targeting three key muscles for urgent words with Silicon tape. Factory made patches will be even more durable, cheap and comfortable. This can make urgent communication fast, comfortable, cost effective, reliable and assessable for all caregivers.

References

Wang, Y., Tang, T., Xu, Y., Bai, Y., Yin, L., Li, G., Zhang, H., Liu, H., & Huang, Y. (2021). All-weather, natural silent speech recognition via machine-learning-assisted tattoo-like electronics. *npj Flexible Electronics, 5*, Article 20. https://doi.org/10.1038/s41528-021-00119-7

Wu, J., Zhang, Y., Xie, L., Yan, Y., Zhang, X., Liu, S., An, X., Yin, E., & Ming, D. (2022). A novel silent speech recognition approach based on parallel inception convolutional neural network and Mel frequency spectral coefficient. *Frontiers in neurorobotics*, *16*, 971446. https://doi.org/10.3389/fnbot.2022.971446

Song, R., Zhang, X., Chen, X., Chen, X., Chen, X., Yang, S., & Yin, E. (2023). Decoding silent speech from high-density surface electromyographic data using transformer. *OpenReview.* 80 (1), https://doi.org/10.1016/j.bspc.2022.104298.

Chen, X., Zhang, X., Chen, X., & Chen, X. (2023). Decoding Silent Speech Based on High-Density Surface Electromyogram Using Spatiotemporal Neural Network. *IEEE transactions on neural systems and rehabilitation engineering: A publication of the IEEE Engineering in Medicine and Biology Society*, *31*, 2069–2078. https://doi.org/10.1109/TNSRE.2023.3266299

Lai, W., Yang, Q., Mao, Y., Sun, E., & Ye, J. (2023). Knowledge Distilled Ensemble Model for sEMG-based Silent Speech Interface. *Cornell University.* https://doi.org/10.48550/arXiv.2308.06533

Li, H., Lin, H., Wang, Y., Wang, H., Zhang, M., Gao, H., Ai, Q., Luo, Z., & Li, G. (2021). Sequence-to-Sequence Voice Reconstruction for Silent Speech in a Tonal Language. *Cornell University.* https://doi.org/10.48550/arXiv.2108.00190

Salomons, I., del Blanco, E., Navas, E., & Hernáez, I. (2025). Electrode Setup for Electromyography-Based Silent Speech Interfaces: A Pilot Study. *Sensors*, *25*(3), 781. https://doi.org/10.3390/s25030781

Li, Z., Ma, B., Mao, W., Zhang, J., Yu, Z. and Lu, Y. (2024). SVIT-SSR: A sEMG-based vision transformer approach for silent speech recognition. Electronics Letter, 60: e13285. https://doi.org/10.1049/ell2.13285

Inoue, M., Sato, M., Tomeoka, K., Nah, N., Hatakeyama, E., Arulkumaran, K., Horiguchi, I., & Sasai, S. (2025). A silent speech decoding system from EEG and EMG with heterogeneous electrode configurations. *Cornell University.* https://doi.org/10.48550/arXiv.2506.13835

Mohapatra, P., Pandey, A., Zhang, X., & Zhu, Q. (2025). Can LLMs understand unvoiced speech? Exploring EMG-to-text conversion with LLMs. In *Proceedings of the 63rd Annual Meeting of the Association for Computational Linguistics (Volume 2: Short Papers)* (pp. 703–712). Vienna, Austria: Association for Computational Linguistics. https://aclanthology.org/2025.acl-short.56/

Majaranta, P., Ahola, U.-K., & Špakov, O. (2009). Fast gaze typing with an adjustable dwell time. In *Proceedings of the 27th ACM Conference on Human Factors in Computing Systems (CHI 2009)* (pp. 357–360). Association for Computing Machinery. https://doi.org/10.1145/1518701.1518758

Bodet-Contentin, L., Messet-Charrière, H., Gissot, V., Renault, A., Muller, G., Aubrey, A., Gadrez, P., Tavernier, E., & Ehrmann, S. (2022). Assessing oral comprehension with an eye tracking based innovative device in critically ill patients and healthy volunteers: a cohort study. *Critical care (London, England)*, *26*(1), 288. https://doi.org/10.1186/s13054-022-04137-3

McLaughlin, D., Peters, B., McInturf, K., Eddy, B., Kinsella, M., Mooney, A., Deibert, T., Montgomery, K., & Fried-Oken, M. (2021). Decision-making for access to AAC technologies in late stage ALS. In B. T. Ogletree (Ed.), *AAC Challenges and Solutions: Improving Everyday Service Delivery* (pp. 169–199). Plural Publishing.

Chen, Y., He, J., Wu, Q., Pu, S., & Song, C. (2024). Prevalence and risk factors of exposure keratopathy among critically ill patients: A systematic review and meta-analysis. *Nursing open*, *11*(1), e2061. https://doi.org/10.1002/nop2.2061

Willcox, M. D. P., Argüeso, P., Georgiev, G. A., Holopainen, J. M., Laurie, G. W., Millar, T. J., Papas, E. B., Rolland, J. P., Schmidt, T. A., Stahl, U., Suarez, T., Subbaraman, L. N., Uçakhan, O. Ö., & Jones, L. (2017). TFOS DEWS II Tear Film Report. *The ocular surface*, *15*(3), 366–403. https://doi.org/10.1016/j.jtos.2017.03.006

Karakus, S., Mathews, P. M., Agrawal, D., Henrich, C., Ramulu, P. Y., & Akpek, E. K. (2018). Impact of dry eye on prolonged reading. *Optometry and Vision Science, 95*(12), 1105–1113. https://doi.org/10.1097/OPX.0000000000001303

Chang W. D. (2019). Electrooculograms for Human-Computer Interaction: A Review. *Sensors (Basel, Switzerland)*, *19*(12), 2690. https://doi.org/10.3390/s19122690

Królak, A., & Strumiłło, P. (2012). Eye-blink detection system for human–computer interaction. *Universal Access in the Information Society, 11*, 409–419. https://doi.org/10.1007/s10209-011-0256-6

Appendix List

Number	Appendix Name
Z	Review of Literature and the Gap
A	Standardizing Electrode Placement
B	Cyton (Device) Configuration and Settings
C	Data Collection Protocol
D	Pre- Processing and Data Export
E	Model Training and Retraining
F	Channel Ablation – finding the best electrodes
G	Key Results

Acknowledgements

I thank the experts I consulted, my parents, my grandparents, my teachers, mentors, my dog and YIS organizers who helped me push boundaries and work hard towards the completion of this project.

I do acknowledge that I still have a long road ahead of me.