

## Objective

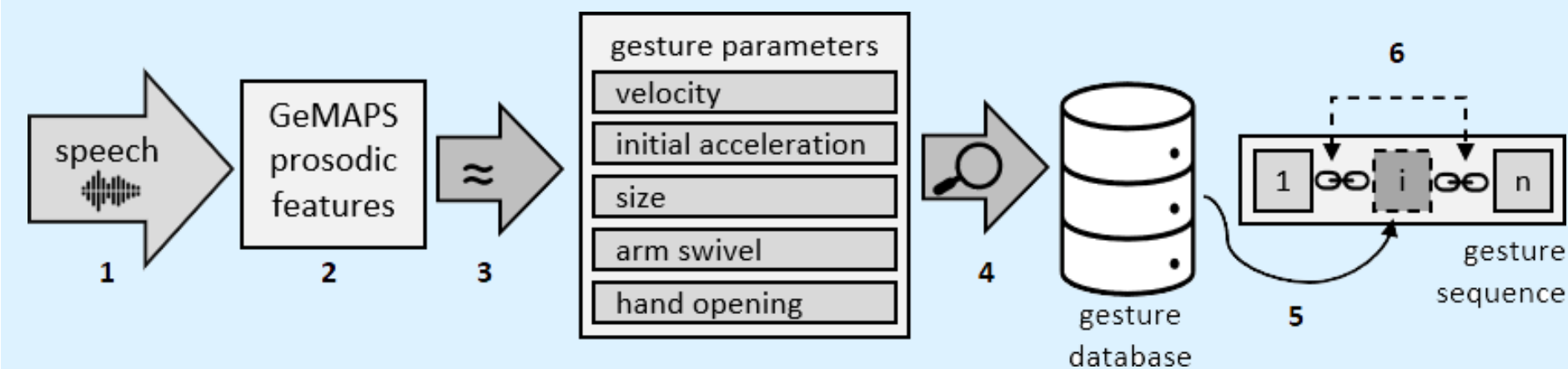
A system for automatic gesture generation from speech. We combine machine learning and database sampling, guaranteeing defined gesture motion.

## Gesture Database

**10 hours** of motion-capture of natural non-dyadic conversational speech (dataset [3] and dataset used in [2])  
Motion segmentation into stroke phases [2] => **23,700 gestures**  
The gesture stroke segmentation informs our **gesture timing**



## Matching System



1. The system input is the speech signal and the gesture timing
2. Prosodic (GeMAPS) features are extracted automatically
3. Gesture parameters are estimated automatically [1]
4. The best matching gesture is found in the database
5. The matched gesture is inserted into the gesture sequence
6. Gestures are linked by synthesized preparations, retractions, & transitions

## Perceptual study

Online experiment, 54 participants  
Gestures animated on the VHTK Brad character [5] (*right*)  
7-point Likert scale rating:  
*"How well did the expressive quality of the gestures match the expressive quality of the speech?"*



### Conditions

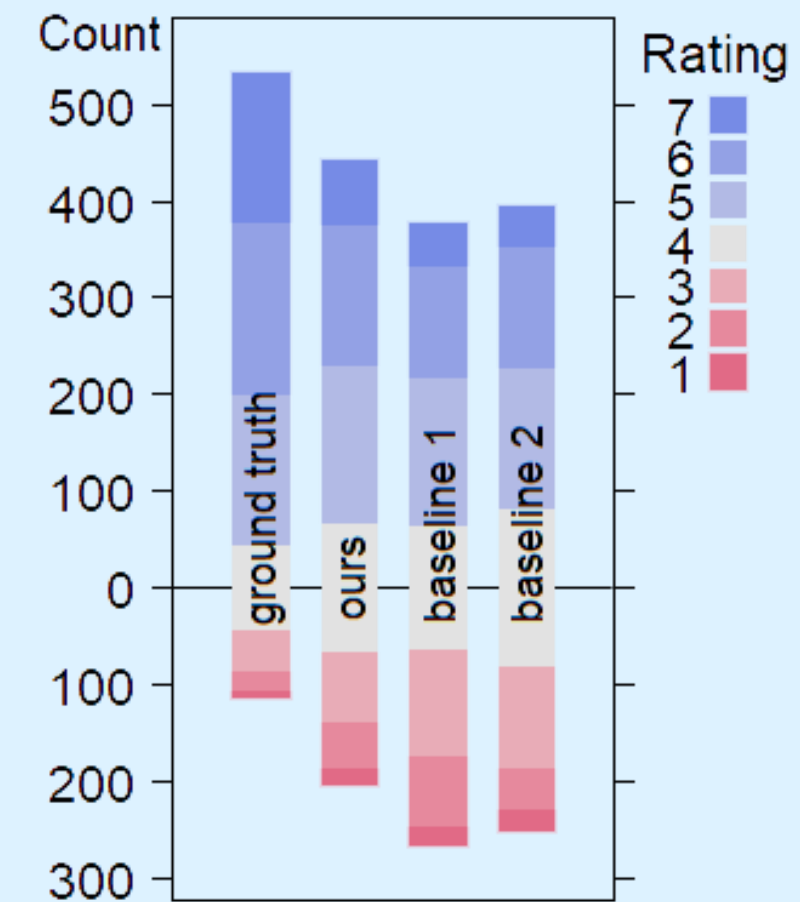
1. Ground truth gesture strokes
2. Parameter-matched gestures
3. Baseline 1: Timed & unmatched
4. Baseline 2: Untimed & unmatched  
(Gesture transitions are always synthesized)

### Results

Ground truth <\*\*\* ours <\*\*\* baselines

### Conclusions

- Parameter-matching gestures improves speech-gesture match
- Gesture timing alone does not improve speech-gesture match



[1] Ferstl, Y., Neff, M., & McDonnell, R. (2020). Understanding the predictability of gesture parameters from speech and their perceptual importance. In *Proceedings of the 20th ACM International Conference on Intelligent Virtual Agents* (pp. 1-8).

[2] Ferstl, Y., Neff, M., & McDonnell, R. (2020). Adversarial gesture generation with realistic gesture phasing. *Computers & Graphics*, 89, 117-130.

[3] Ferstl, Y., & McDonnell, R. (2018). Investigating the use of recurrent motion modelling for speech gesture generation. In *Proceedings of the 18th International Conference on Intelligent Virtual Agents* (pp. 93-98).

[4] Hartholt, A., Traum, D., Marsella, S. C., Shapiro, A., Stratou, G., Leuski, A., Morency, L.P & Gratch, J. (2013, August). All together now. In *International Workshop on Intelligent Virtual Agents* (pp. 368-381). Springer, Berlin, Heidelberg.