Beyond Responding Fast or Slow: **Improving Cognitive Models of Memory Retrieval** using **Prosodic Speech Features**

Thomas Wilschut[1]
Florian Sense[1]
Odette Scharenborg[2]
Hedderik van Rijn[1]

[1]Experimental Psychology, University of Groningen
[2]Multimedia Computing Group, Delft University of Technology

1. Background

- Adaptive learning systems have improved fact learning.
- The effectiveness of such systems depends on their ability to estimate and predict learner performance.
- In current systems, **accuracy** and **response times** (RTs) are frequently used to estimate memory strength.

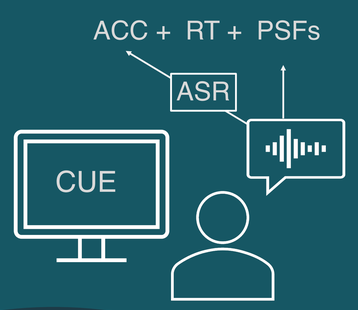
- Recent developments in automatic speech recognition (ASR) allow for the transition from *typing*-based systems to *speech*-based systems.
- **Prosodic Speech Features** (PSFs), such as intonation, rhythm, and stress, may carry additional information about individual learning processes which could be used to improve speech-based adaptive learning systems.

2. Research questions

1. Which PSFs are associated with retrieval accuracy and retrieval speed?

2. Can PSFs be used to improve predictions for future learning performance?

3. Methods



PROCEDURE

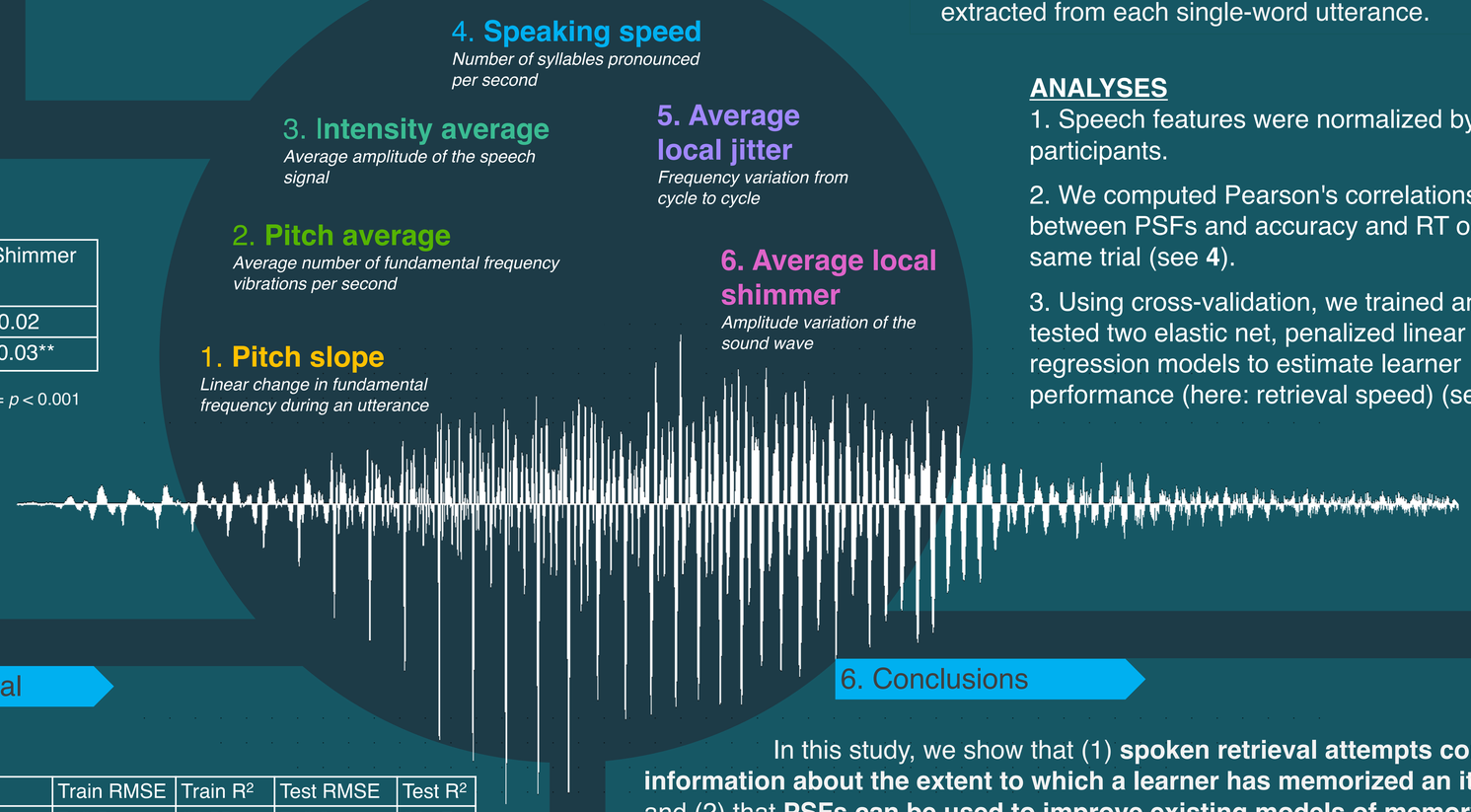
- Participants (n = 50) studied Swahili-English vocabulary items using the **SlimStampen** adaptive scheduling system – which uses **accuracy** and **RTs** to predict learner performance and determine item scheduling.
- Participants responded by pronouncing the English translation of the Swahili word.
- All utterances were transcribed to text in real-time using Google's speech-to-text ASR.
- After the experiment, PSFs were automatically extracted from each single-word utterance.

ANALYSES

1. Speech features were normalized by participants.

2. We computed Pearson's correlations between PSFs and accuracy and RT on the same trial (see **4**).

3. Using cross-validation, we trained and tested two elastic net, penalized linear regression models to estimate learner performance (here: retrieval speed) (see **5**).



4. PSFs correlate with memory measures…

	RT	Pitch slope	Pitch average	Intensity average	Speaking speed	Jitter	Shimmer
Accuracy	-0.37 ***	**-0.10*** **	-0.03*	**0.05*** **	**0.07*** **	-0.01	-0.02
RT		**0.11*** **	0.01	**-0.05*** **	**-0.07*** **	0.03*	0.03**

* = $p < 0.05$, ** = $p < 0.01$, *** = $p < 0.001$

Higher retrieval accuracy and **faster response times** were associated with **falling pitch** (negative pitch slope), **higher intensity** and **higher speaking speed.**

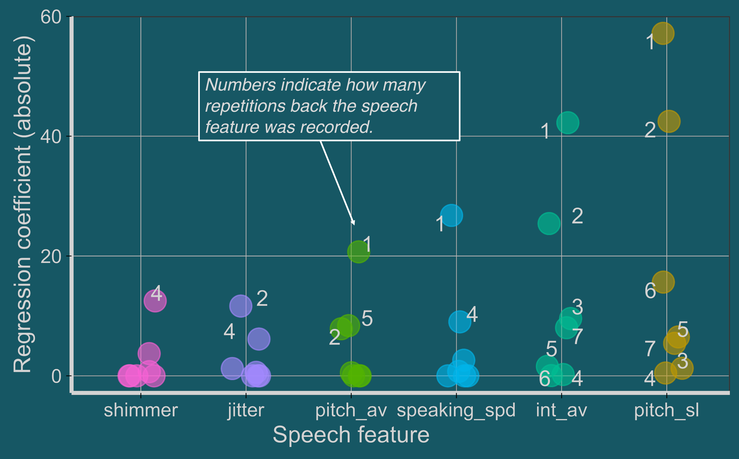
5. … and can be used to improve models of memory retrieval

- **Using past-repetition PSFs in addition to the original memory model estimations resulted in better estimations of current-repetition retrieval speed.**

- Past-repetition **pitch slope** was the most important predictor of performance, followed by **average intensity, speaking speed**, and **average pitch**.

- PSFs for the **past two** repetitions were the most important predictors of current-repetition performance for the same item.

	Train RMSE	Train R^2	Test RMSE	Test R^2
Original model	748.5	68.55	748.3	68.47
Original model + Prosodic speech features	670.6	74.69	671.0	74.56



6. Conclusions

In this study, we show that (1) **spoken retrieval attempts contain information about the extent to which a learner has memorized an item**, and (2) that **PSFs can be used to improve existing models of memory retrieval**.

Our results are important in two ways:

1. They show that specific speech characteristics are associated with learning performance. These features can be a valuable new tool in further research.

2. They can contribute to the development of speech-based adaptive learning systems or cognitive tutors. Employing PSFs can be a computationally inexpensive way to improve educationally relevant learning applications.

More information

www.slimstampen.nl

www.thomaswilschut.com

Contact

@thomas_wilschut

t.j.wilschut@rug.nl

university of groningen