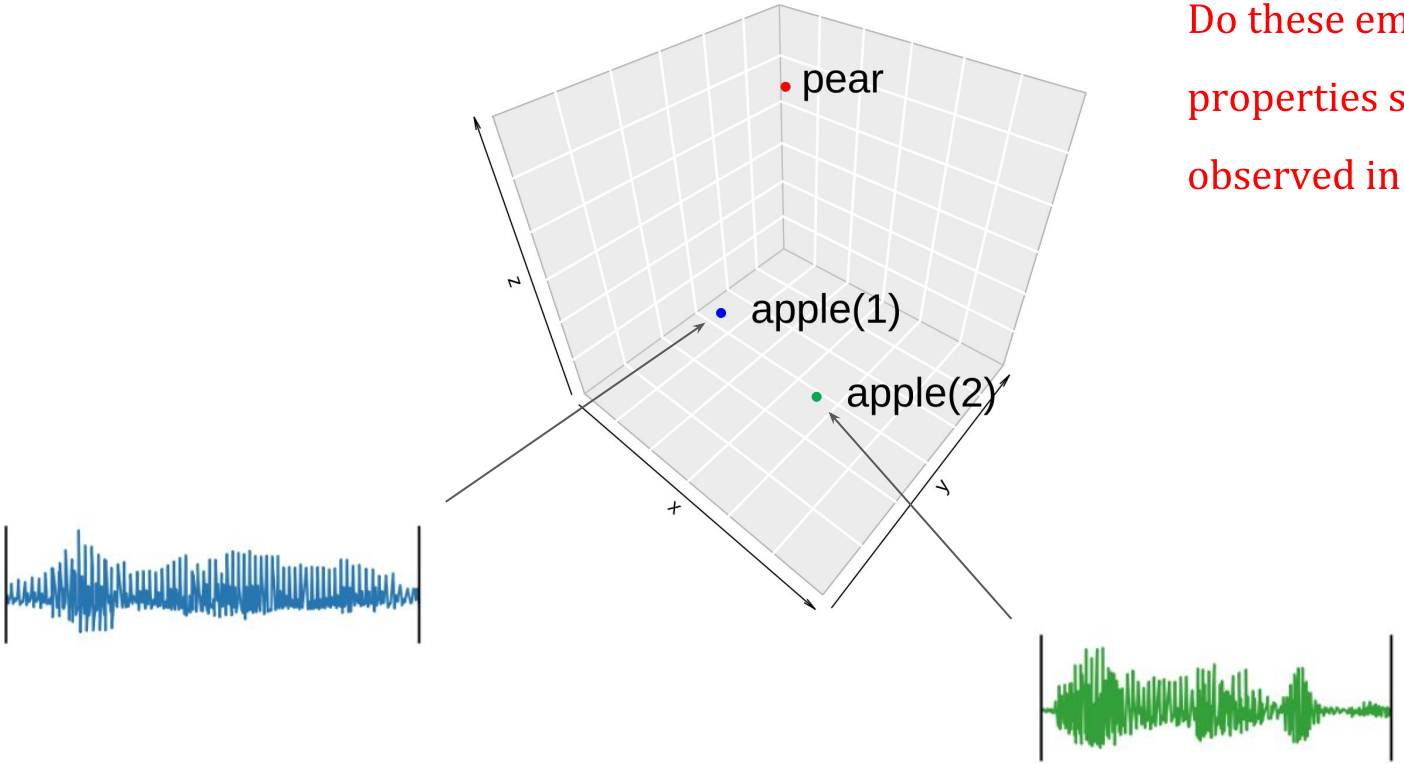
Analyzing Autoencoder-based Acoustic Word Embeddings

Yevgen Matusevych [1]
Herman Kamper [2]
Sharon Goldwater [1]

[1] University of Edinburgh, UK
[2] Stellenbosch University, South Africa

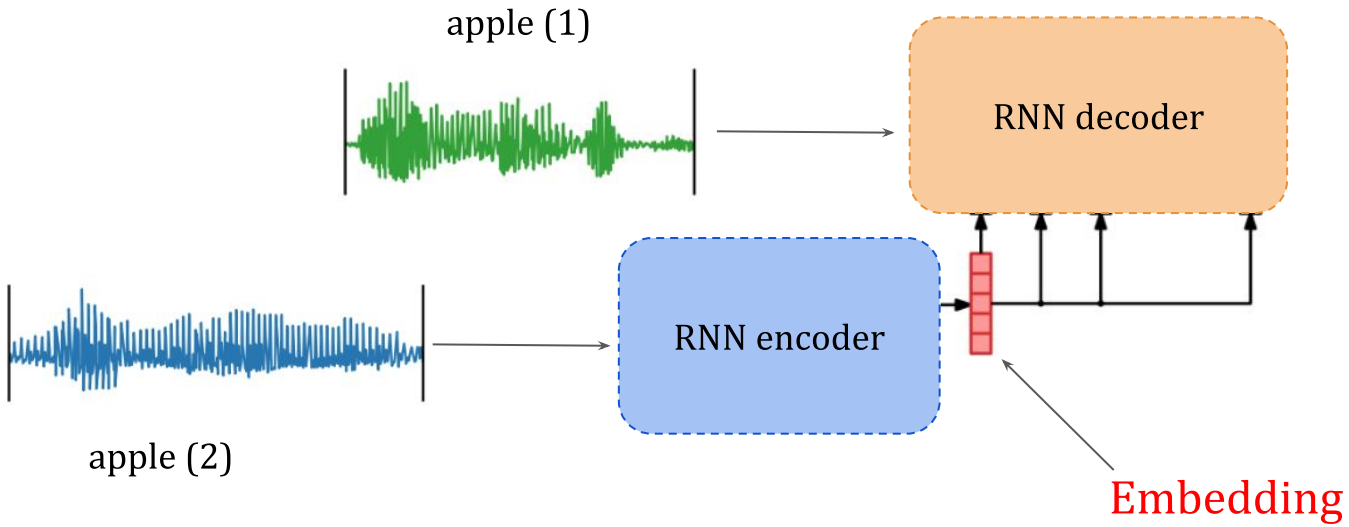
Acoustic word embeddings



Do these embeddings have properties similar to those observed in human speakers?

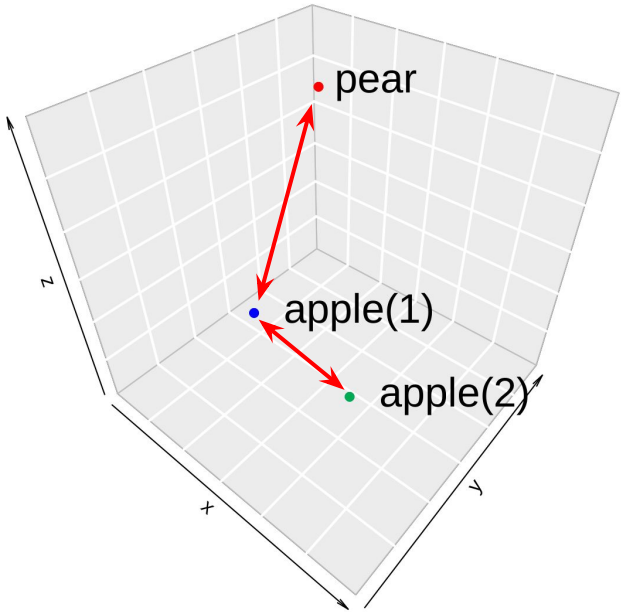
Matusevych, Kamper, Goldwater. Analyzing autoencoder-based acoustic word embeddings.

Correspondence-autoencoding recurrent neural network



Matusevych, Kamper, Goldwater. Analyzing autoencoder-based acoustic word embeddings.

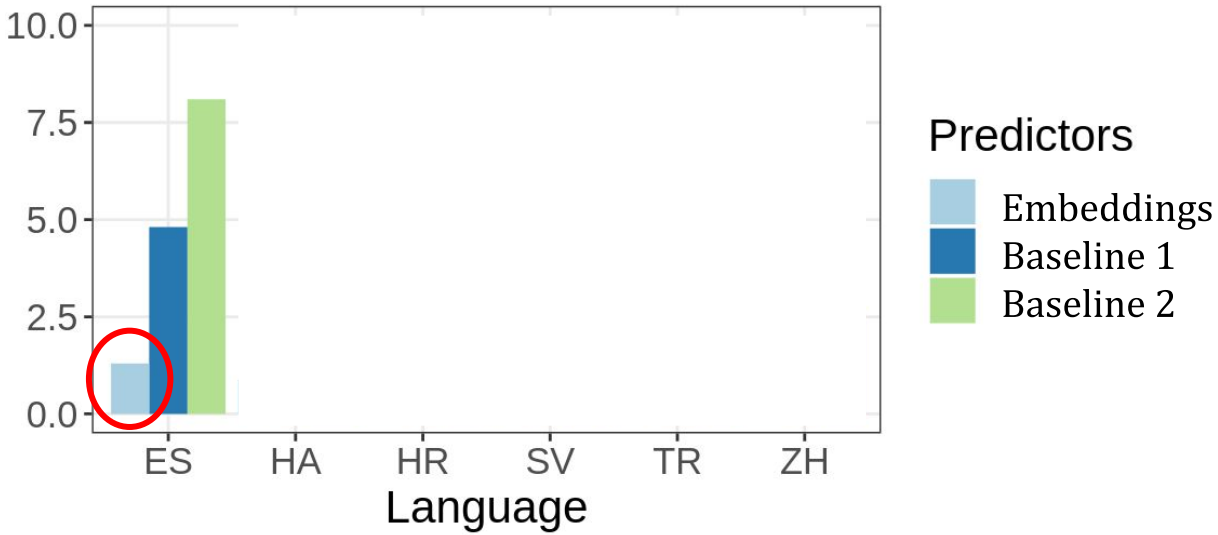
Probing the embedding space



Matusevych, Kamper, Goldwater. Analyzing autoencoder-based acoustic word embeddings.

Linear regression: Predict number of phones

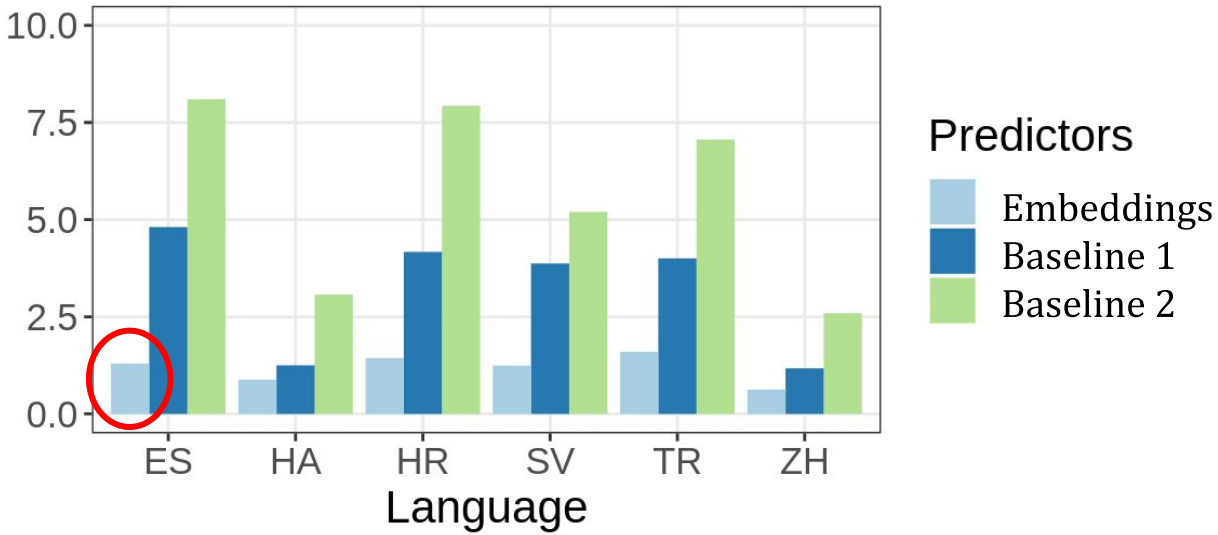
Mean squared error
(lower is better)



Matusevych, Kamper, Goldwater. Analyzing autoencoder-based acoustic word embeddings.

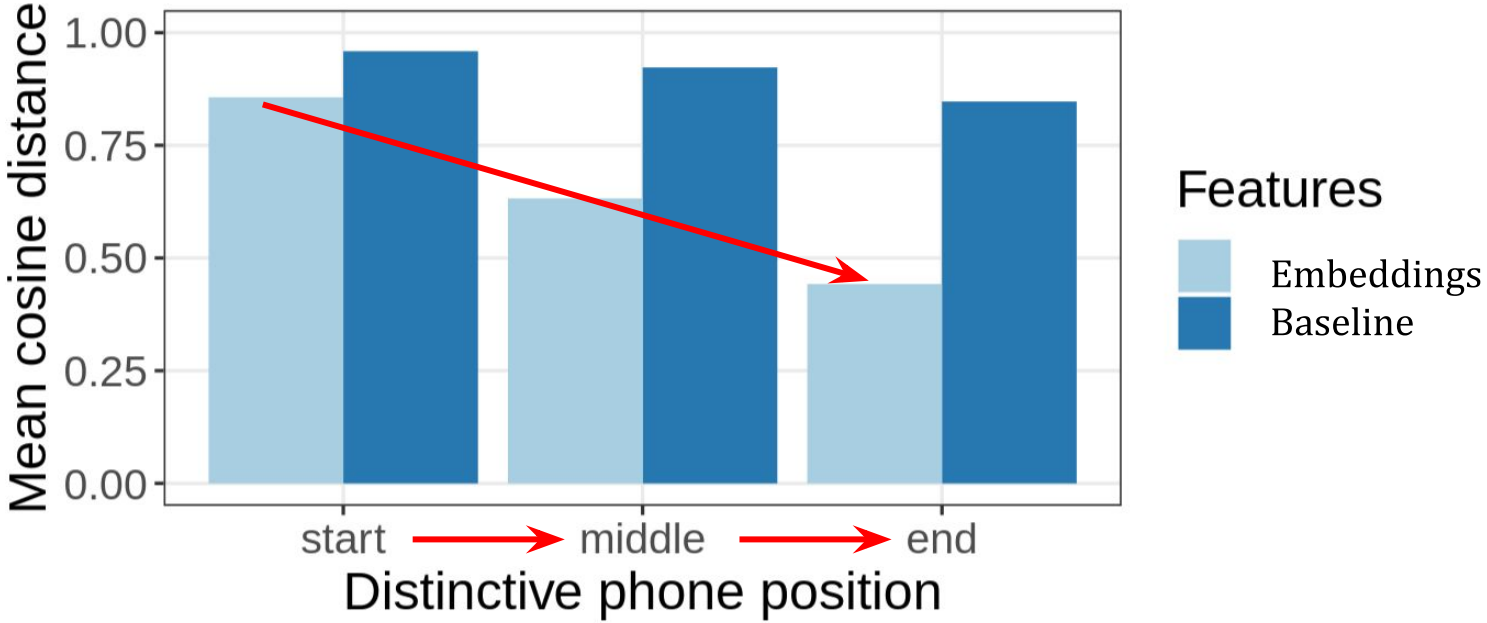
Linear regression: Predict number of phones

Mean squared error
(lower is better)



Matusevych, Kamper, Goldwater. Analyzing autoencoder-based acoustic word embeddings.

A cognitive word onset bias: First phone is more prominent



Matusevych, Kamper, Goldwater. Analyzing autoencoder-based acoustic word embeddings.

Conclusion

- Acoustic embeddings show some promise for cognitive science.
- Spoken words of variable duration are embedded into the same space that is easy to probe.
- They can provide a link between speech processing and lexical access.

Matusevych, Kamper, Goldwater. Analyzing autoencoder-based acoustic word embeddings.