DARTMOUTH

Looking at Sounds
Visualizing Time Series Data

Dr. Simon Stone

Interview Day for the Position of **Research Data Specialist**

August 1, 2022

Hello, nice to meet you!

Simon Stone

- Senior Researcher and Data Scientist at Technische Universität Dresden, Germany
- In the process of relocating to the US with my family

Graduate degree

- Electrical Engineering and Information Technology
- Minor in Medical Engineering

Doctorate degree

- Electrical and Computer Engineering
- Focus on Signal Processing, Data Science, and Speech Technology

Outline

- What makes Time Series Data special?
- Speech signals as a special kind of Time Series
- German articulatory phonetics 101
- Undoing time: Dimensionality reduction
- Visualizing emerging patterns
- Take-aways and final words

Outline

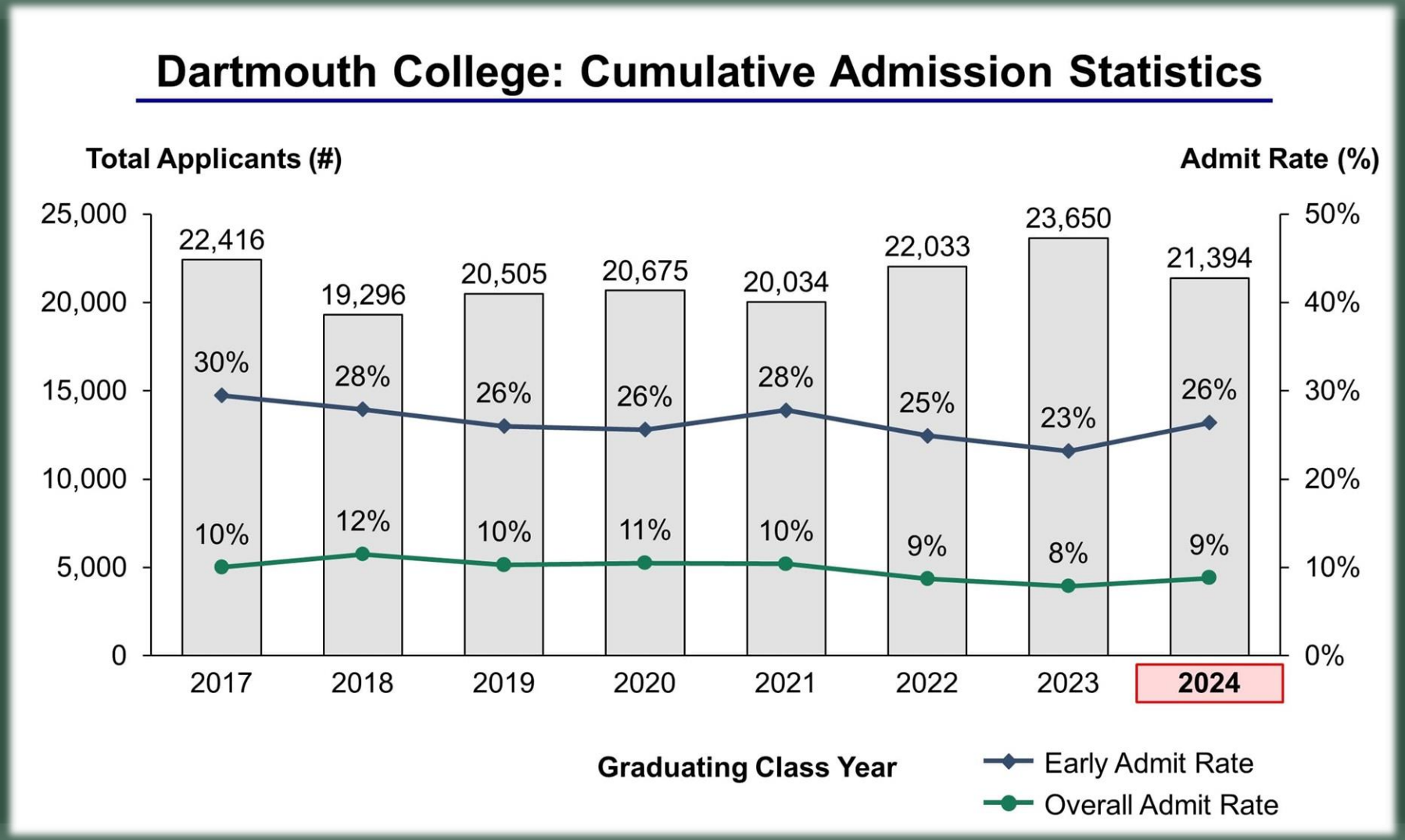
- What makes Time Series Data special?
- Speech signals as a special kind of Time Series
- German articulatory phonetics 101
- Undoing time: Dimensionality reduction
- Visualizing emerging patterns
- Take-aways and final words

What makes Time Series Data special?

- A **time series** is a series of **data points indexed by time.**

What

- A time



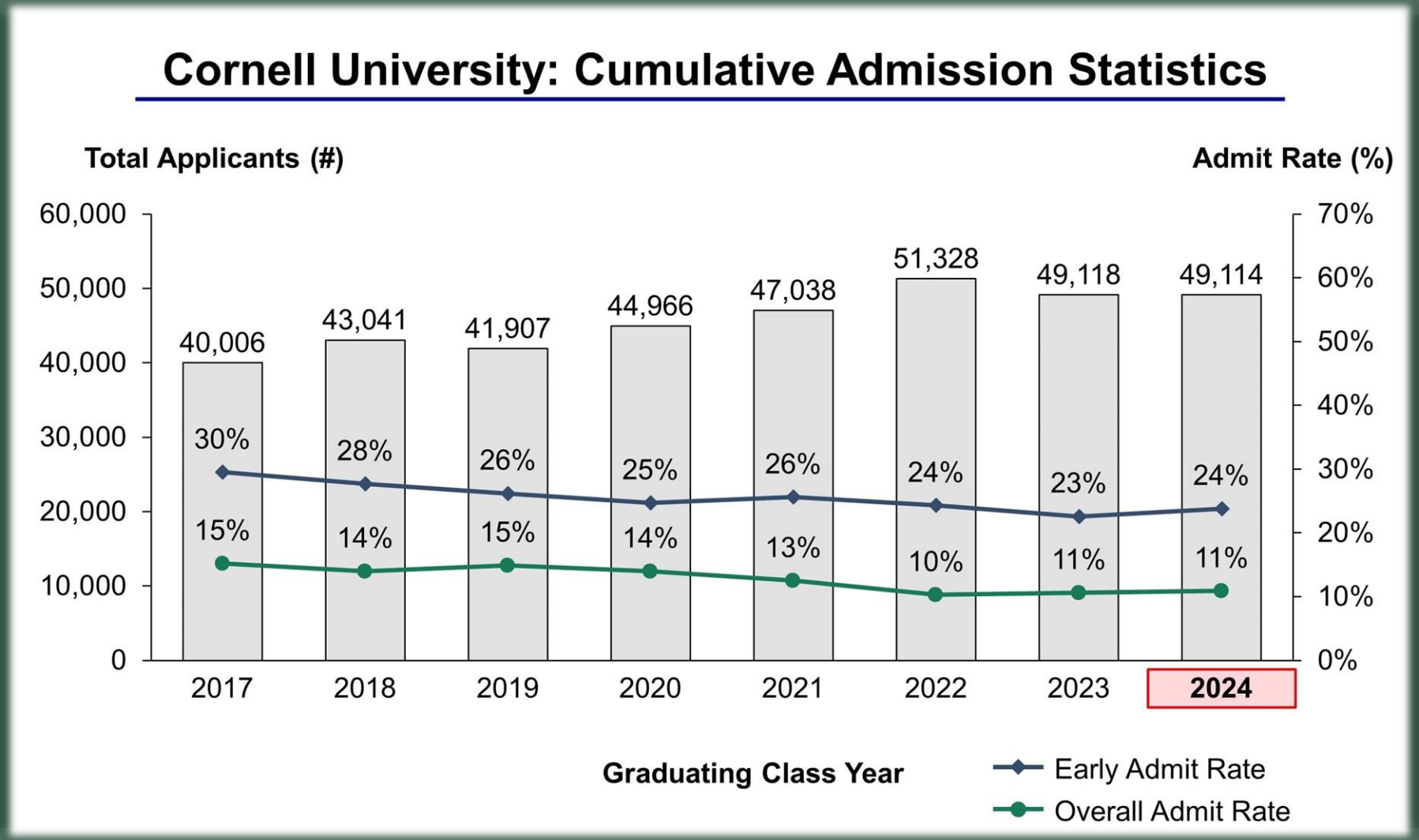
Dartmouth College: Cumulative Admission Statistics. (n.d.). [Graph]. Ivy League Prep. https://ivyleagueprep.com/dartmouth-college-class-of-2024/

What makes Time Series Data special?

- A **time series** is a series of **data points indexed by time**.
- **Information** may be encoded **in the temporal evolution**.
- **Summary statistics** often **fail** to capture this information.
- **Visualizing** several instances of the same time series is **hard**.

What

- A time
- Informa
- Summa
- Visualiz



Cornell University: Cumulative Admission Statistics. (n.d.). [Graph]. Ivy League Prep. https://ivyleagueprep.com/cornell-university-class-of-2024/

What makes Time Series Data special?

A **time series** is a series of **data points indexed by time**.

Information may be encoded **in the temporal evolution**.

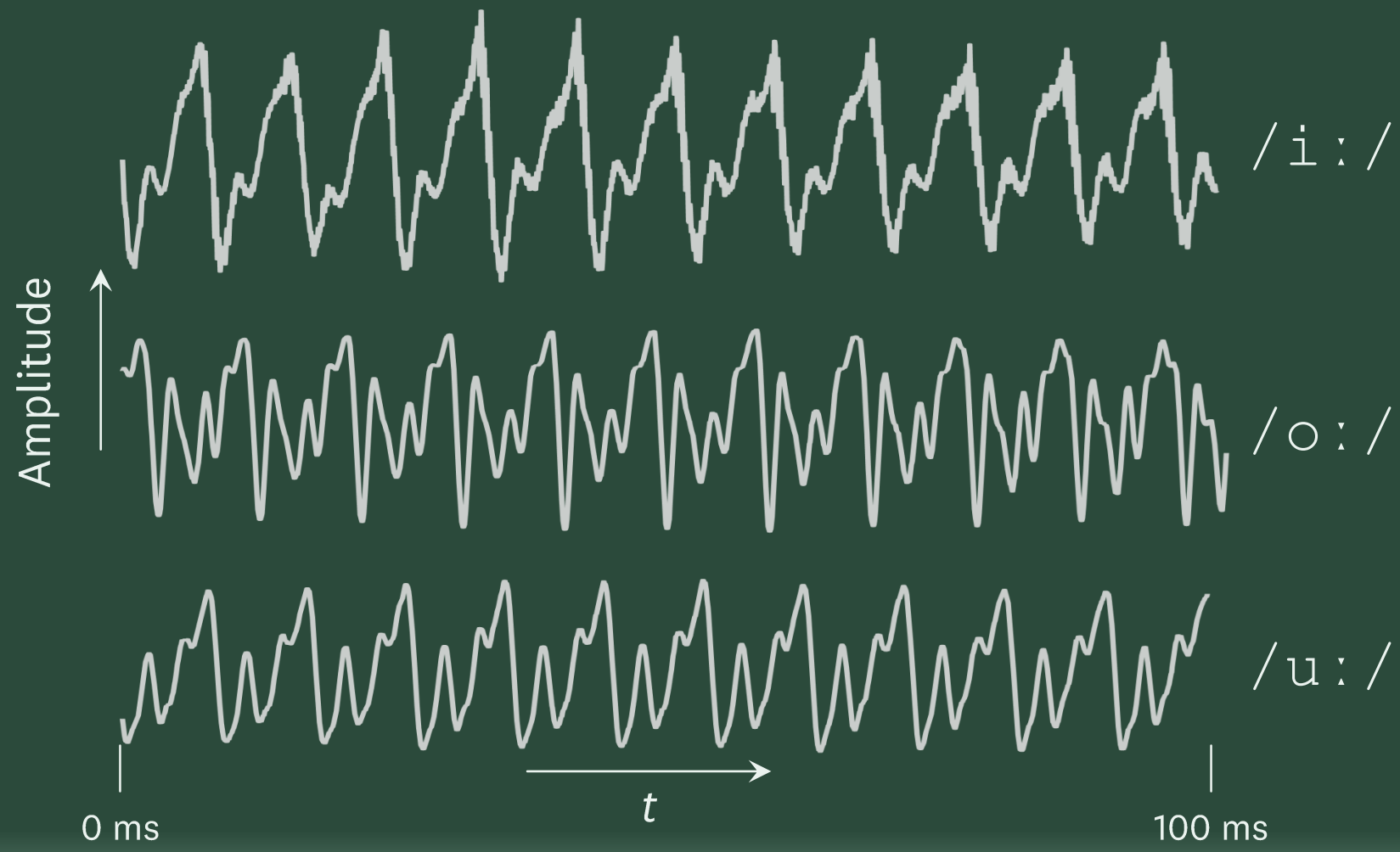
Summary statistics often **fail** to capture this information.

Visualizing several instances of the same time series is **hard**.

→ **Time-related processing required**

Speech signals as a special kind of Time Series

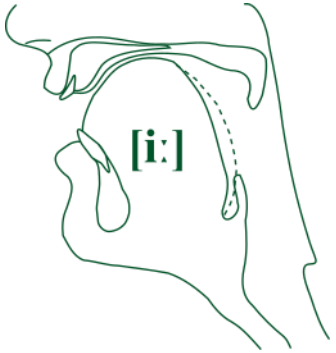
- Acoustic **speech** signals are **pressure waves.**
- **Amplitude** and **frequency** composition **vary over time.**
- **Information** is encoded in these **variations.**
- Our **ear detects** the signal and our **brain decodes** them into meaning.

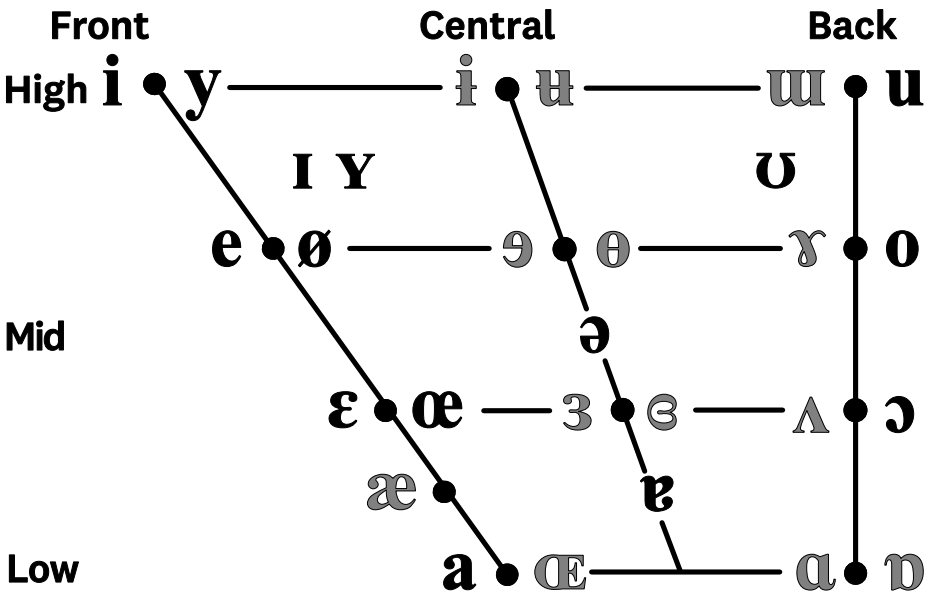

/i:/
/o:/
/u:/
Amplitude
t
0 ms
100 ms

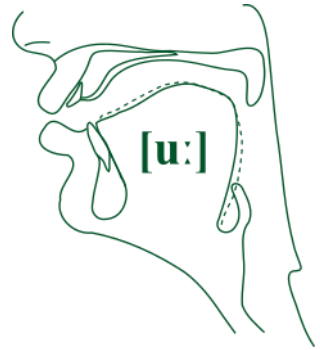
Speech signals as a special kind of Time Series

- Acoustic speech signals are **pressure waves**.
- **Amplitude** and **frequency** composition **vary over time**.
- **Information** is encoded in these **variations**.
- Our **ear detects** the signal and our **brain decodes** them into meaning.
- We can easily and intuitively **compare instances** of speech sounds.
- How to **visualize perceptual similarity** for a large number of sounds?

German articulatory phonetics 101







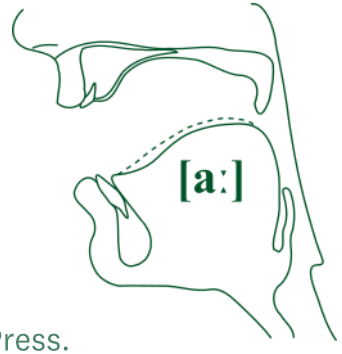


Chart adapted from Figure 4, p. 12 in:
Handbook of the International Phonetic Association: A Guide to the Use of the International Phonetic Alphabet. (1999). United Kingdom: Cambridge University Press.

Undoing time: Dimensionality reduction

Goal:

Remove the dimension „**time**“ without removing the information encoded in this dimension.

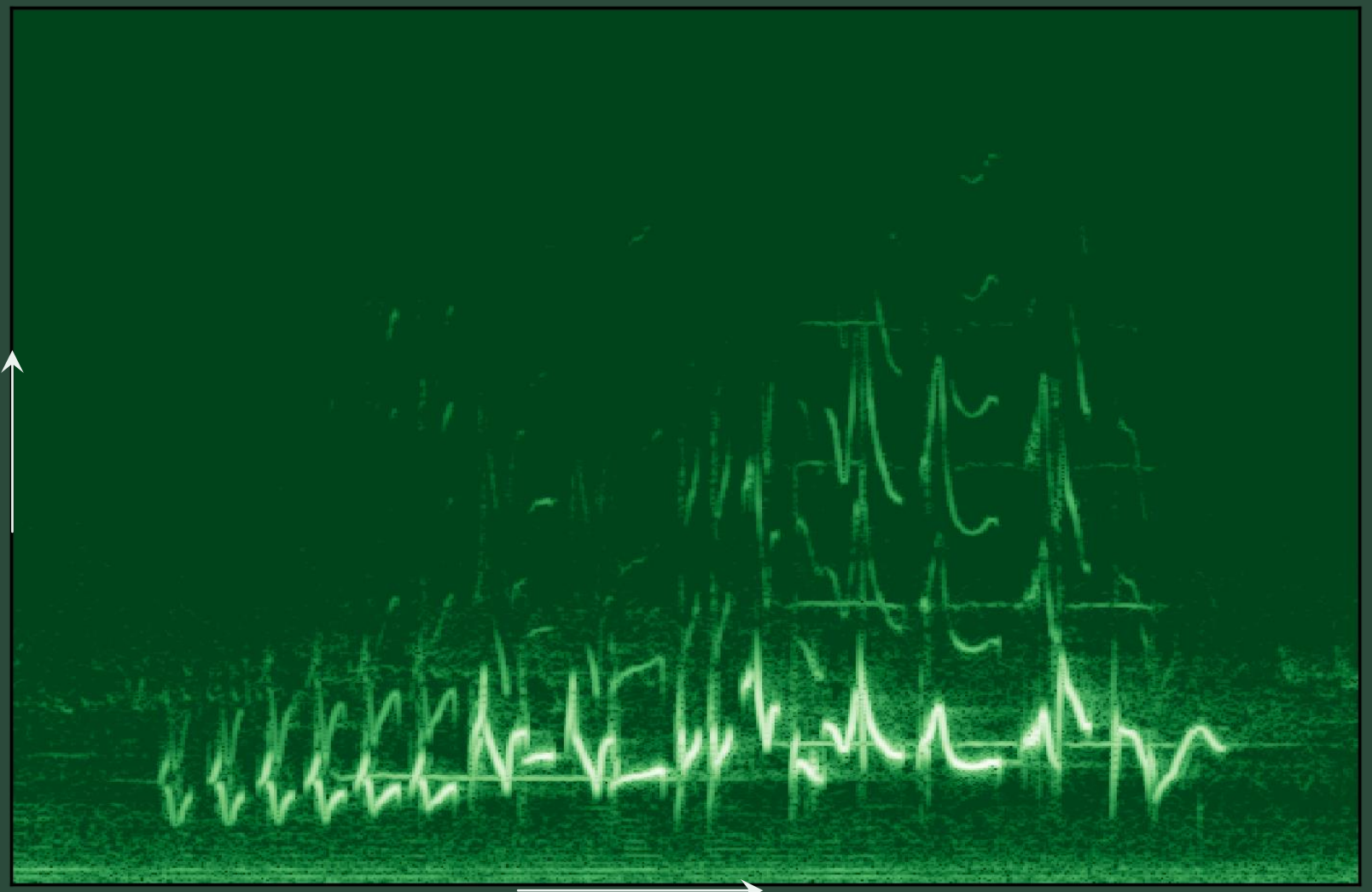
Methods:

- **Statistical** methods (LDA, PCA, *t*-SNE)
- Using **domain knowledge**

Undoing time: Dimensionality reduction

Domain knowledge:

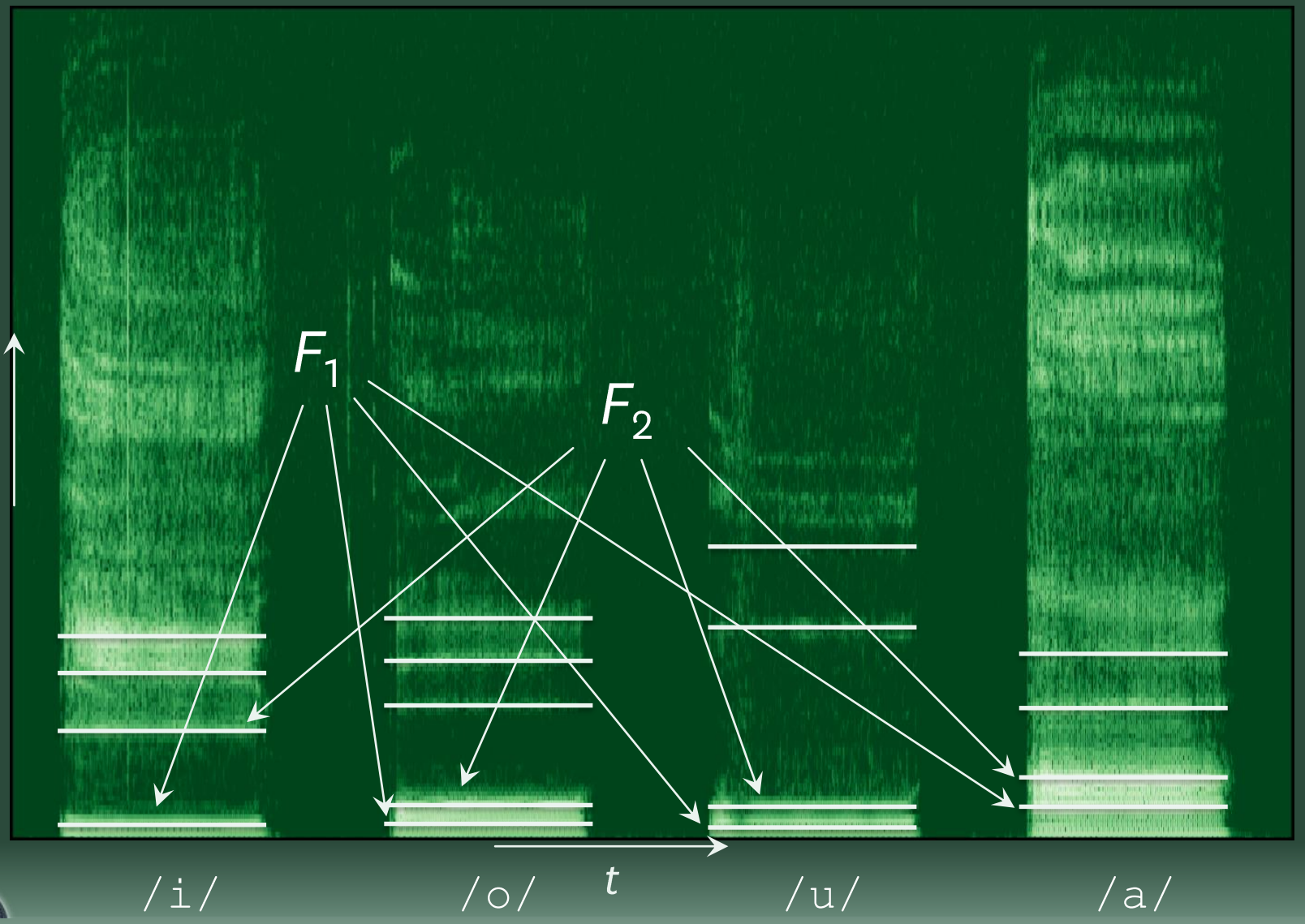
Sounds can be **described by the frequencies** they contain.



Sound excerpt from:
Matthew D. Medler, "*Purple Finch (Eastern)*", 2019, https://macaulaylibrary.org/asset/172805541

Bird photo by Cephas - Own work, CC BY-SA 3.0,
https://commons.wikimedia.org/w/index.php?curid=15362339

—— : Formants



Undoing time: Dimensionality reduction

Domain knowledge:

Sounds can be **described by the frequencies** they contain.

In vowel sounds, a **pattern of specific resonant frequencies** can be observed.

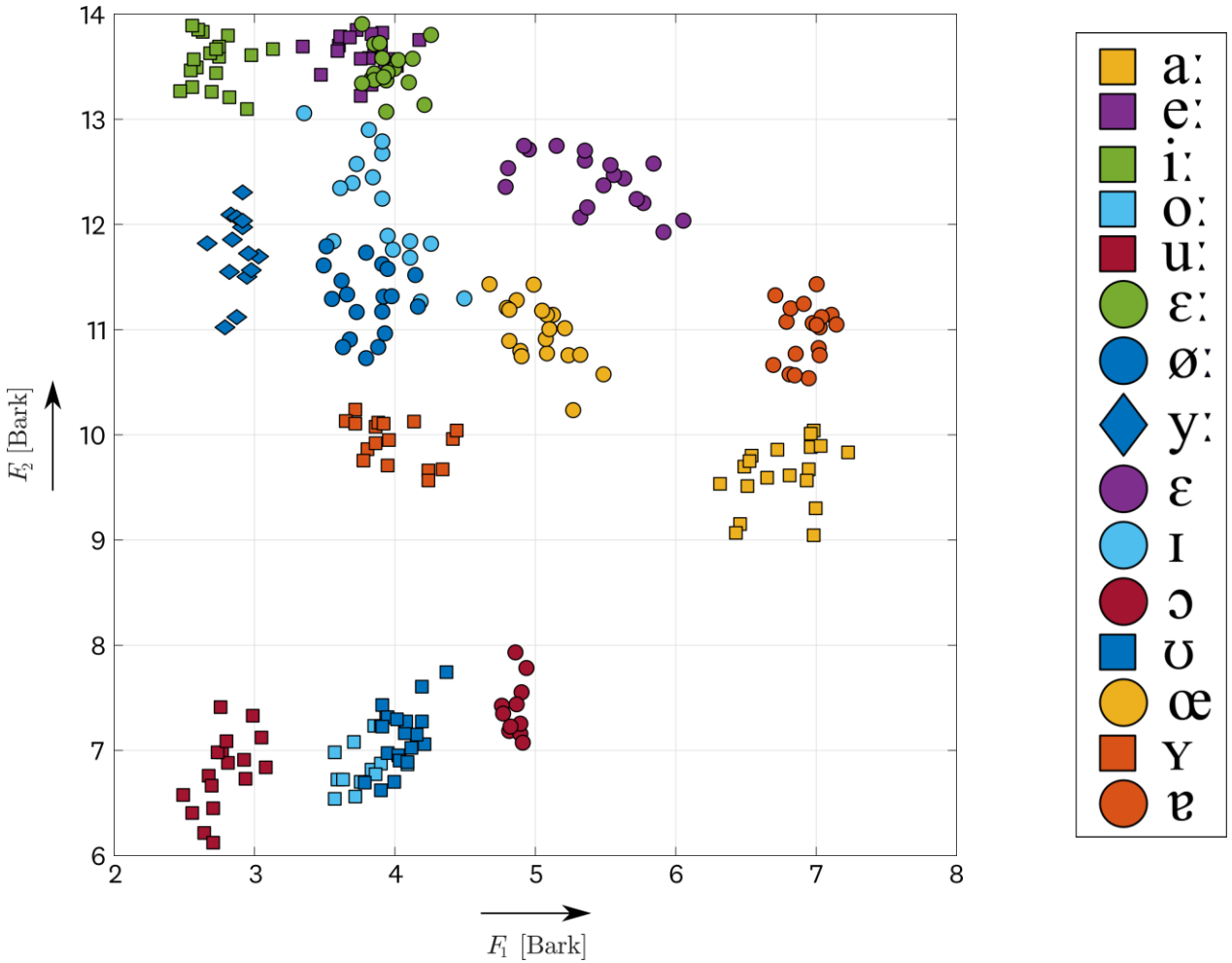
These resonant frequencies are called **formants**.

Amplitude changes over time, but the **formants are constant** (more or less).

Visualizing emerging patterns

Approach:

- **Measure F_1 and F_2** for each instance of a vowel sound
- **Plot** each sound in the **formant space**
- **Similar** sounds are **close, dissimilar** sounds are **not as close**

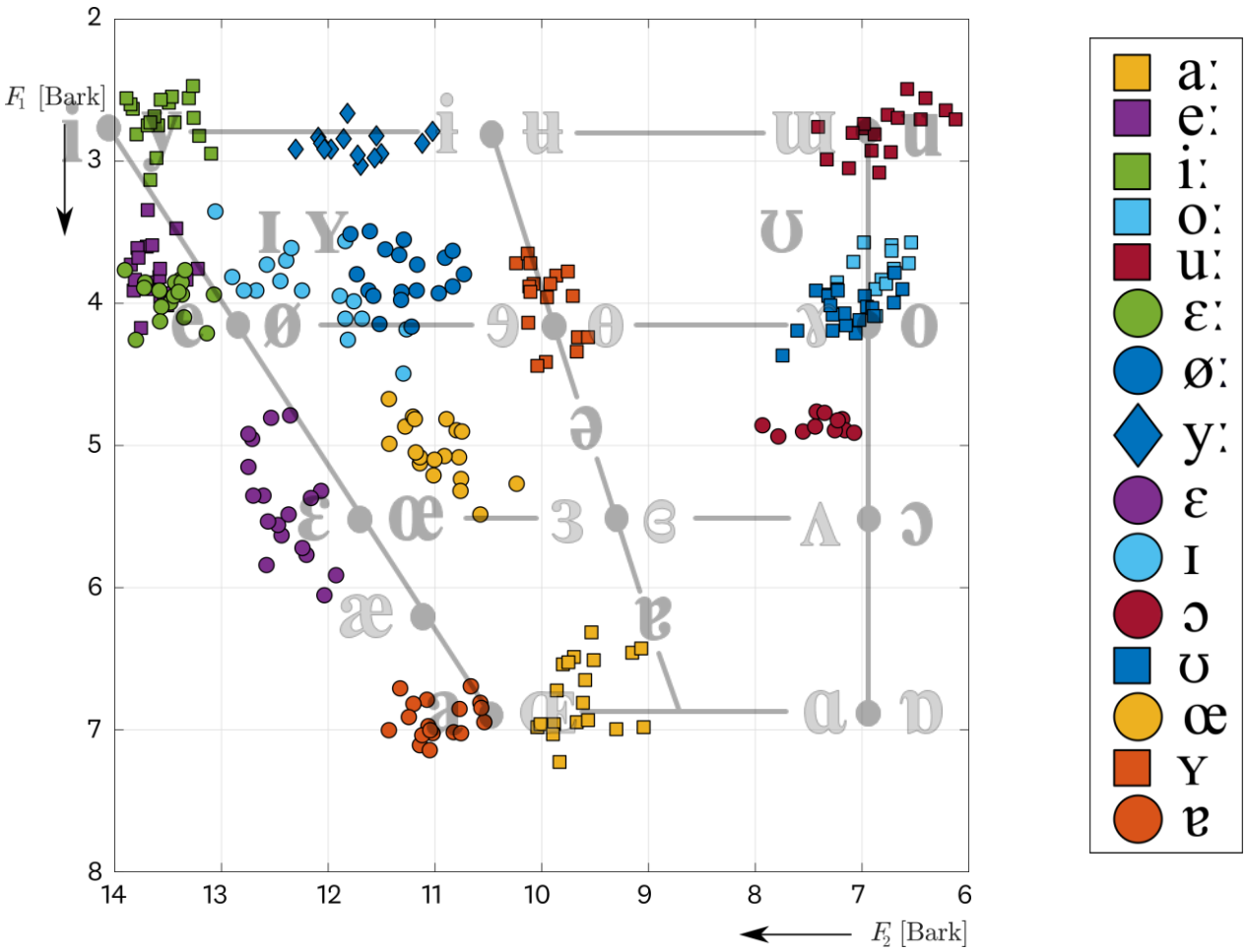


But wait, there's more...

Visualizing emerging patterns

Information contained in this visualization:

- **Perceptual similarity** between instances of the same vowel
- **Articulatory relationships** (tongue height and position) between different vowels

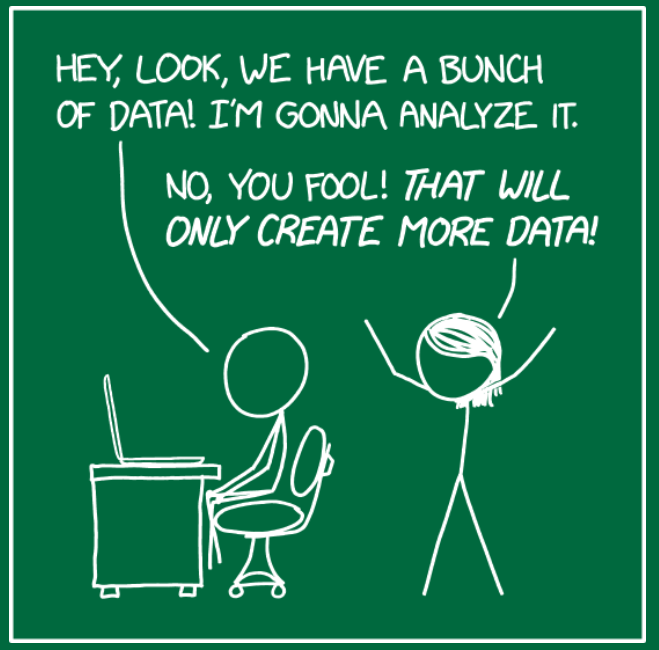


Take-aways and final words

- **Time series are tricky.**
- **Get rid of time but keep its information.**
- **Simple is not wrong and not easy.**
- **Stick to conventions.**
- **Don‘t be afraid to break with conventions.**

- Efficient **data visualization** requires **subject matter expertise** and **domain knowledge**.
- Data **processing** and **visualization** are **not easily outsourced**.
- As **your Research Data Specialist**, I would **work with you** to implement solutions that truly **support your research** and effectively **promote your findings**.

Thank you.



Munroe, Randall. "Data Trap." *xkcd.com*, CC BY-NC 2.5

Slides, code and data used in this presentation:
git clone https://github.com/Simon-Stone/DartmouthInterviewDay.git

A lot of this was based on:
S. Stone, Y. Gao and P. Birkholz, "Articulatory Synthesis of Vocalized /r/ Allophones in German, " in *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, vol. 30, pp. 879-889, 2022, doi: 10.1109/TASLP.2021.3130969.