



# Gamification in Phonetic Perception Data Collection

Accent Identification in North-East England

Yanyu Li, Damar Hoodland,  
Daniel Duncan, Cong Zhang

## **Demographic questionnaires**

Age

Gender

Place of origin

...

Which accent do you think it's from?



Tyne and Wear

County Durham

Northumberland



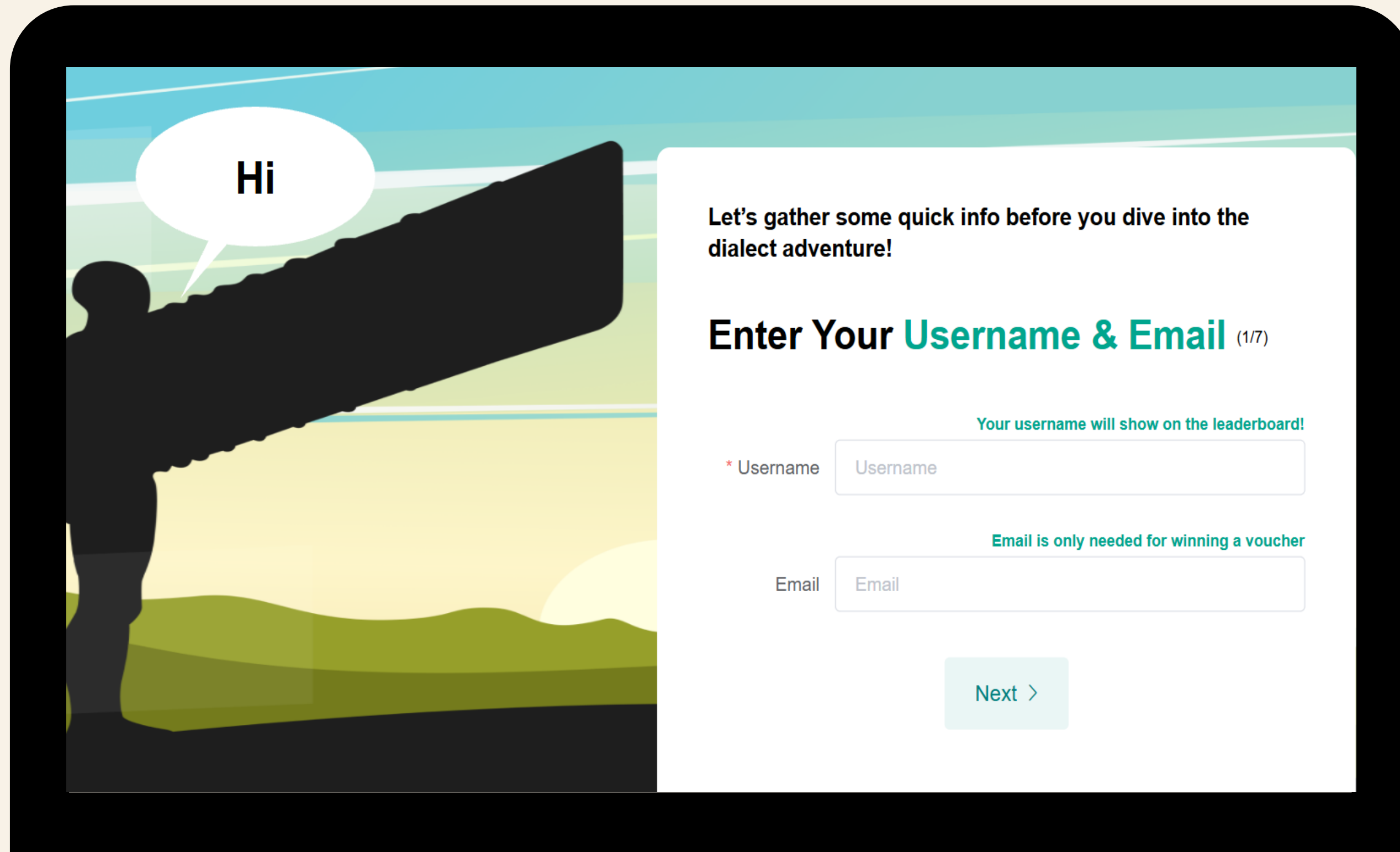
Debriefing

Acknowledgement

# North - East Accent Game

**START!**

<https://research.ncl.ac.uk/ne-accent-games/game/>



<https://research.ncl.ac.uk/ne-accent-games/game/>

## 🎮 How to Play 🎮

📢 In the game, click the 'Wey Aye' button to listen to the sound clip.

📢 Now, let's make sure your audio is working—click the button to hear a long beep!

Wey Aye



📖 Then, identify where the accent in the audio is from and **click** on it on the map.

🏆 Finally, complete the **questionnaire** to qualify for winning the vouchers.

💰 The top three players who complete the entire game (including questionnaire) will win. Ties will be decided by a draw 🎲.

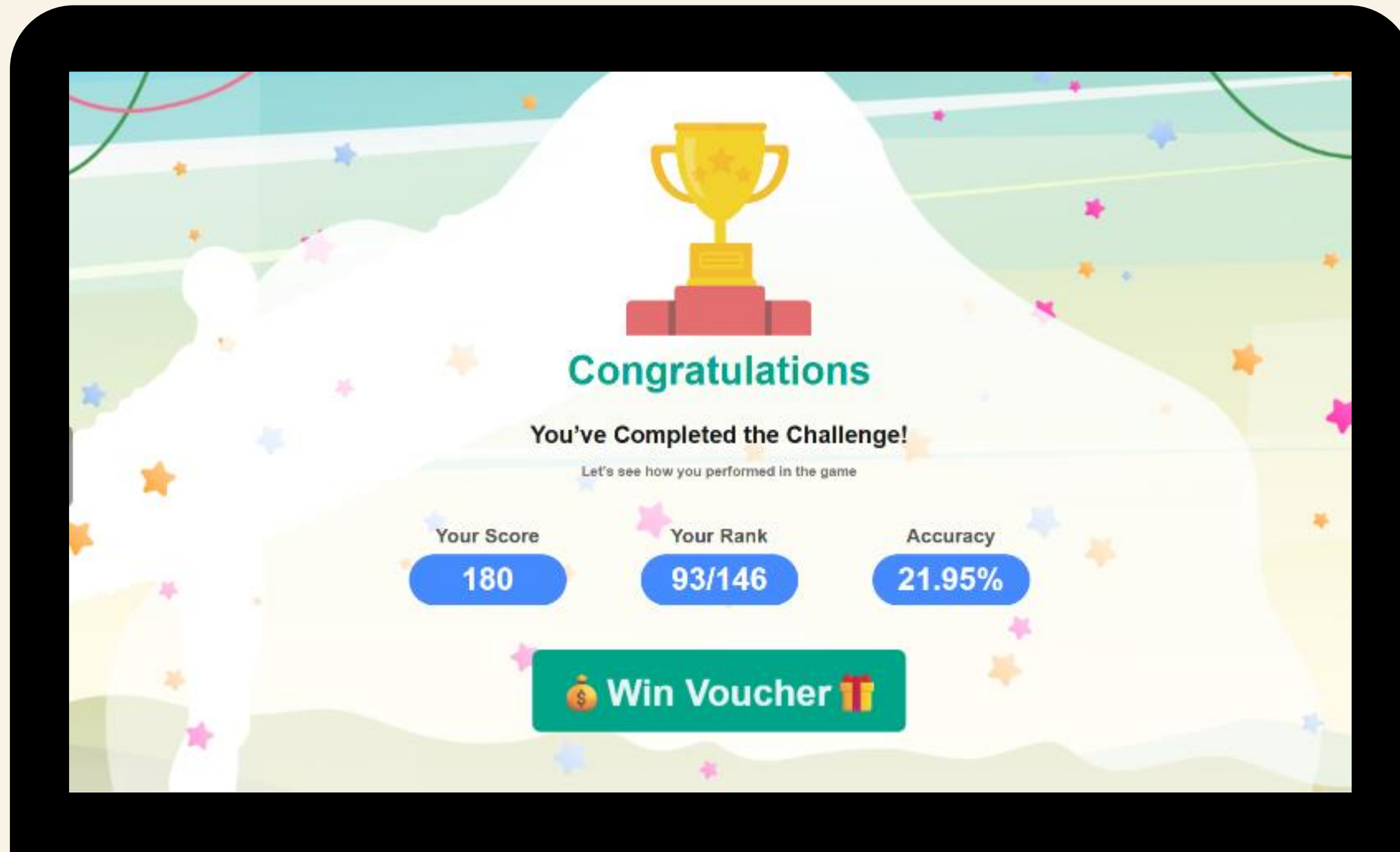
- 🏆 £30 for 1st place
- 🏆 £20 for 2nd place
- 🏆 £10 for 3rd place

Play!

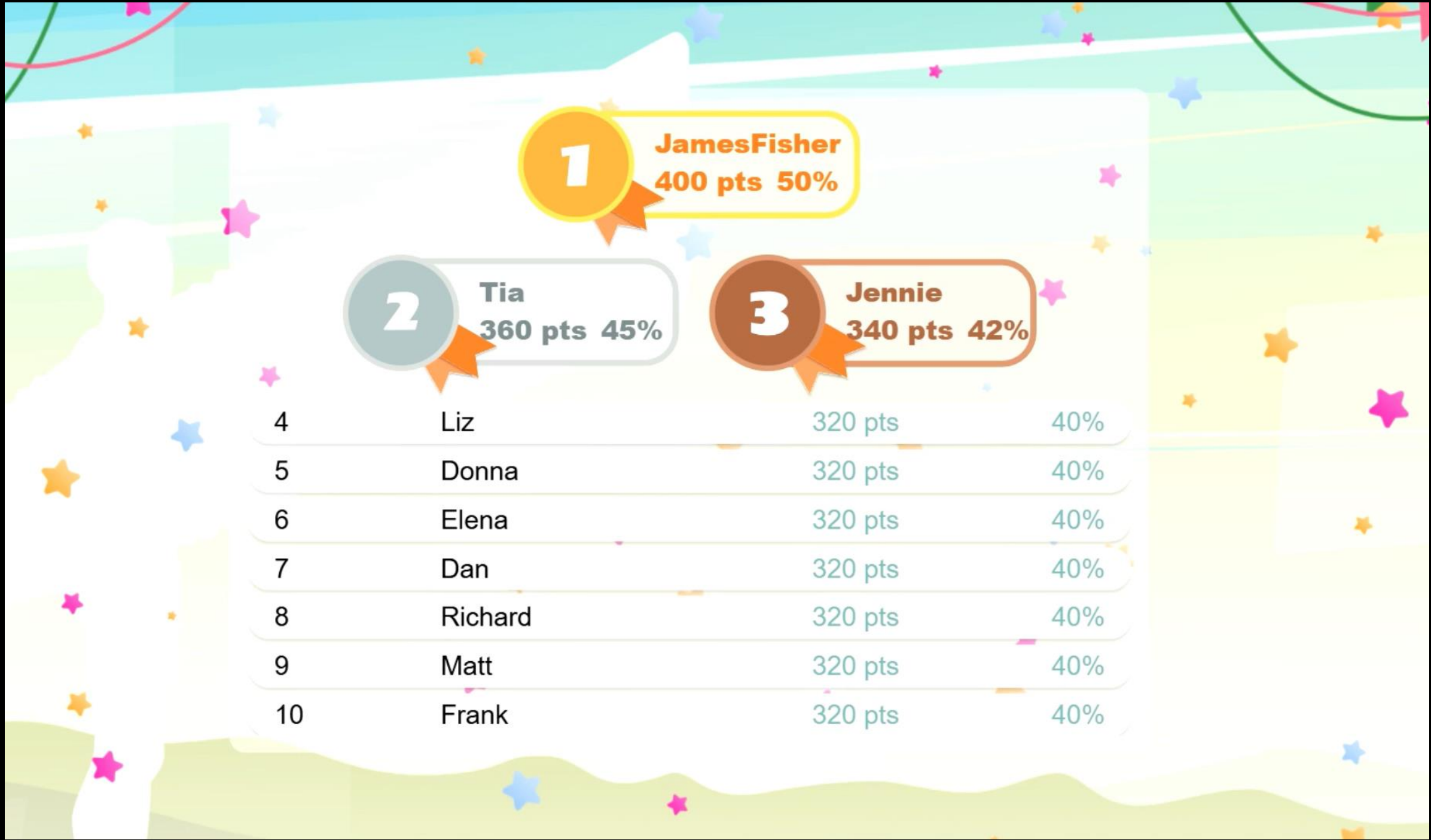
<https://research.ncl.ac.uk/ne-accent-games/game/>



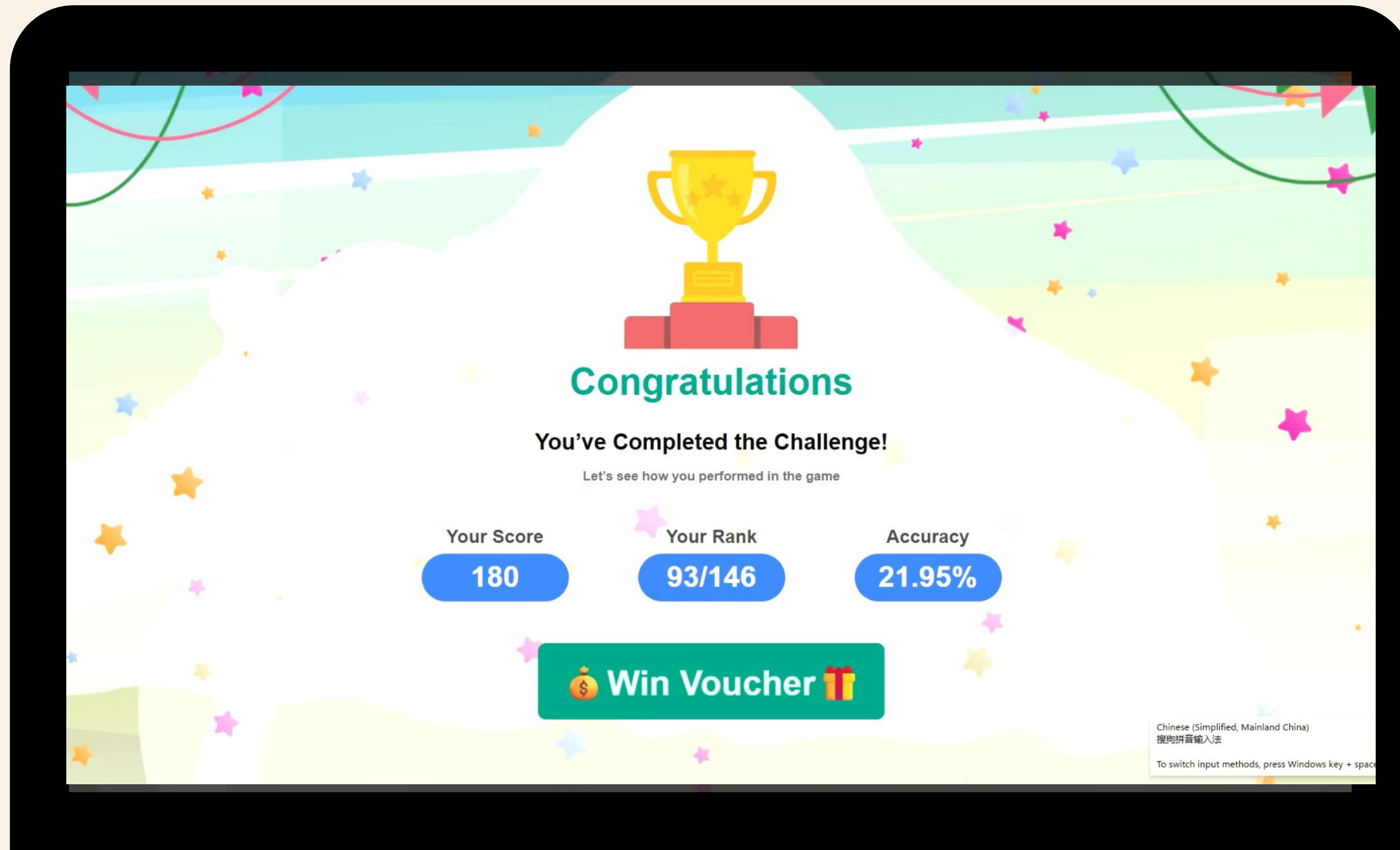
<https://research.ncl.ac.uk/ne-accent-games/game/>



<https://research.ncl.ac.uk/ne-accent-games/game/>



<https://research.ncl.ac.uk/ne-accent-games/game/>



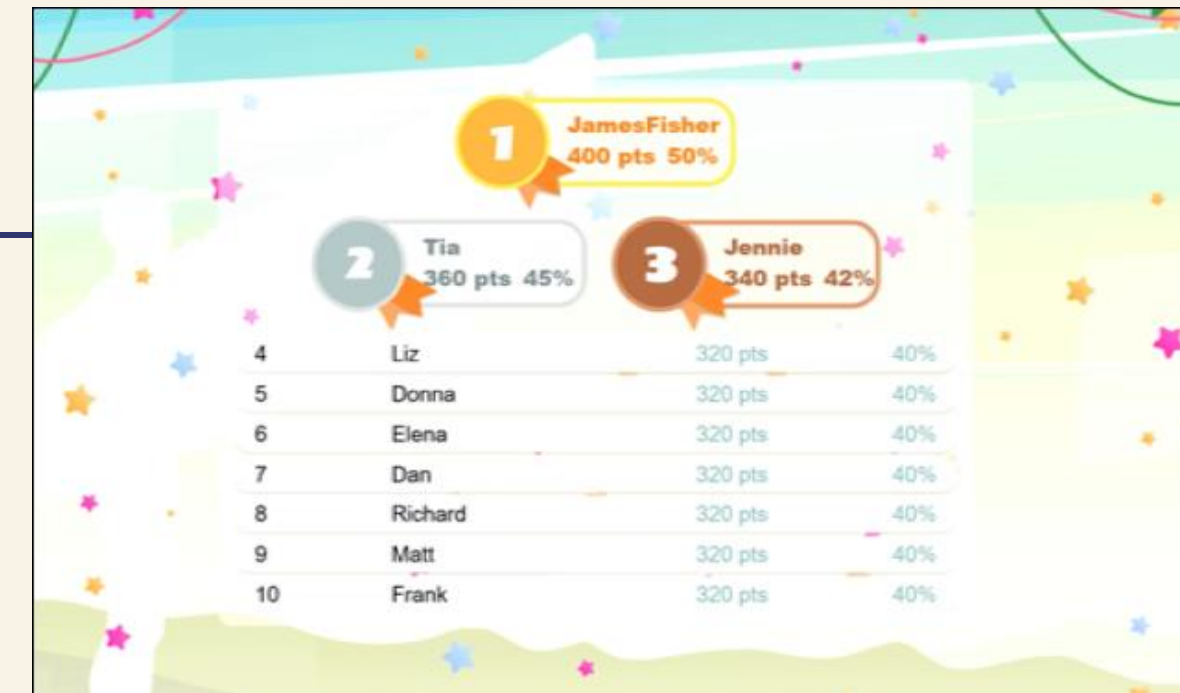
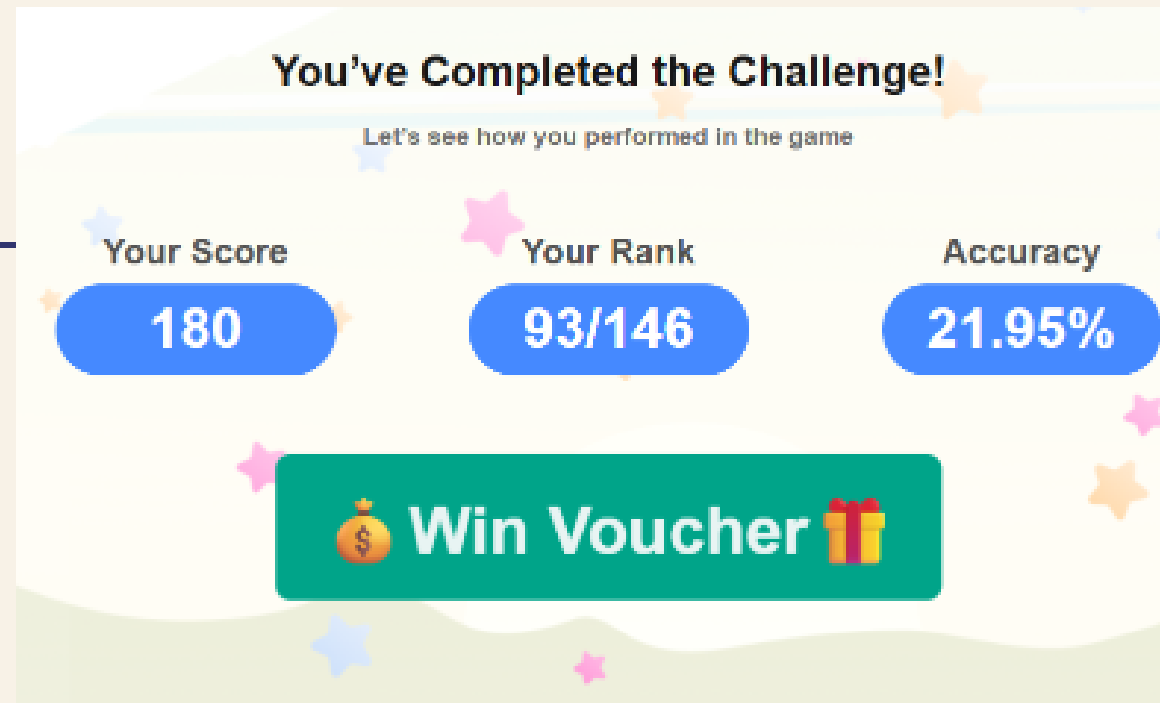
<https://research.ncl.ac.uk/ne-accent-games/game/>

# Elements of gamification is based on Self-Determination Theory (SDT).

progress bar

points

leaderboard



Autonomy  
(feeling in control)

Competence  
(feeling capable)

Relatedness  
(feeling connected)

**The points-badges-leaderboard design (PBL)** (Escher 2016, Seaborn & Fels 2015)

# The PBL elements have been used in various research domains...

progress bar

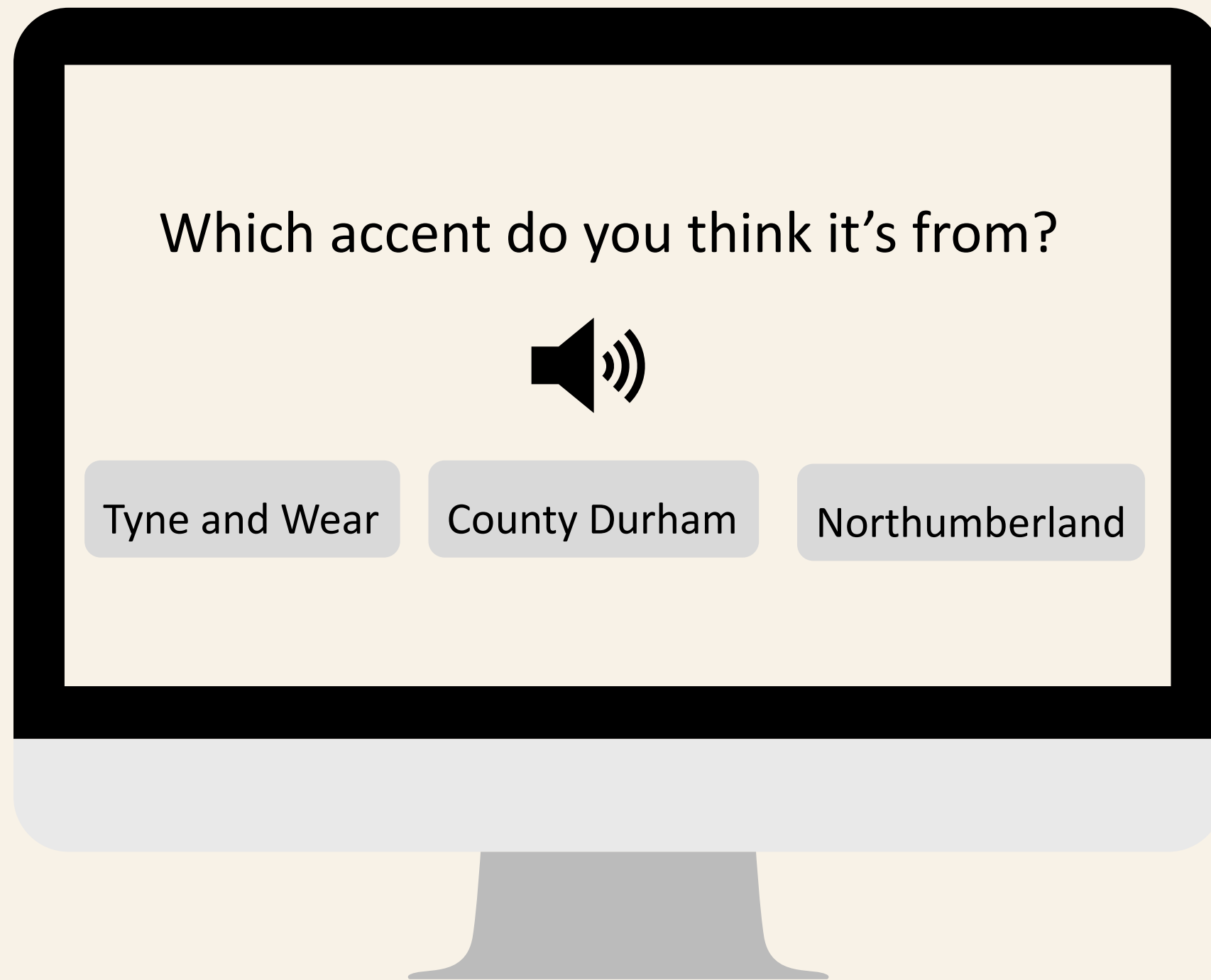
points

leaderboard

Research areas: economics, health, education, and marketing  
(Landers et al. 2018, Bozkurt & Durak 2018, Long et al. 2023, Shortt et al. 2023, Seaborn & Fels 2015).

Non-research areas: Duolingo, NIKE fitness app etc

# Gamification test-case in phonetic research

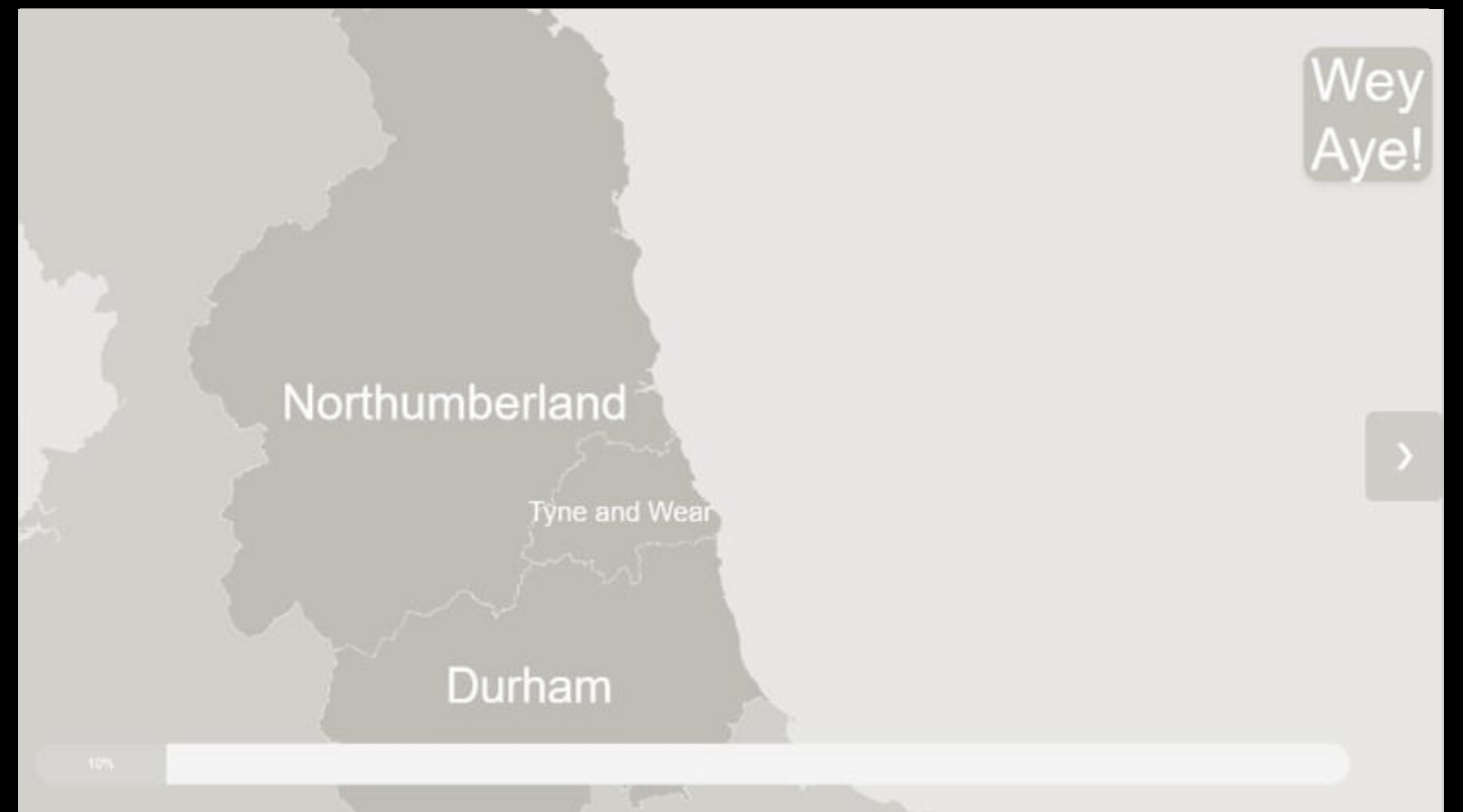


# RQ1: If gamification methods can be used in phonetic data collection?



Which accent do you think it's from?

- A. Tyne and Wear
- B. County Durham
- C. Northumberland



# Why accent identification?

- Speakers draw inferences about their interlocutor(s) based on accents. (Eckert 2008)
- This intersection produces perceptions of prestige vs. non-standard dialects, affecting social solidarity and identity. (Coupland & Bishop 2007)



# New Yorkers' belief

- categorical differences
  - E.g., Queens is “nasal” to the exclusion of other boroughs (Becker & Newlin-Łukowicz 2018)
- However, research suggests a “uniformity of abstract patterns of variation” (Labov 1972: 121)

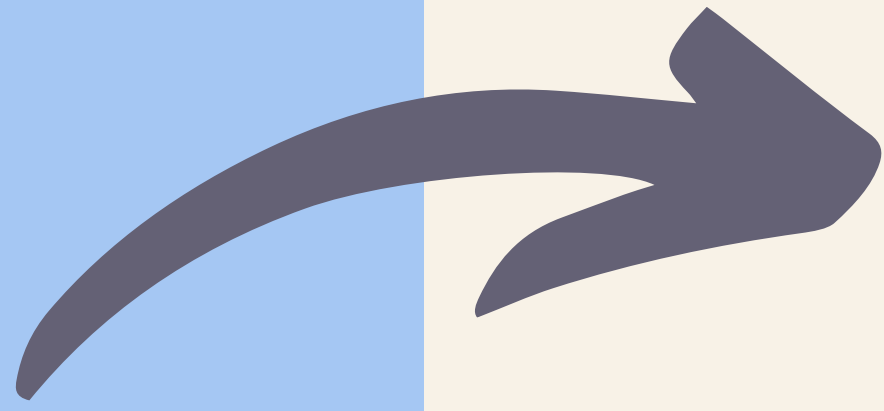
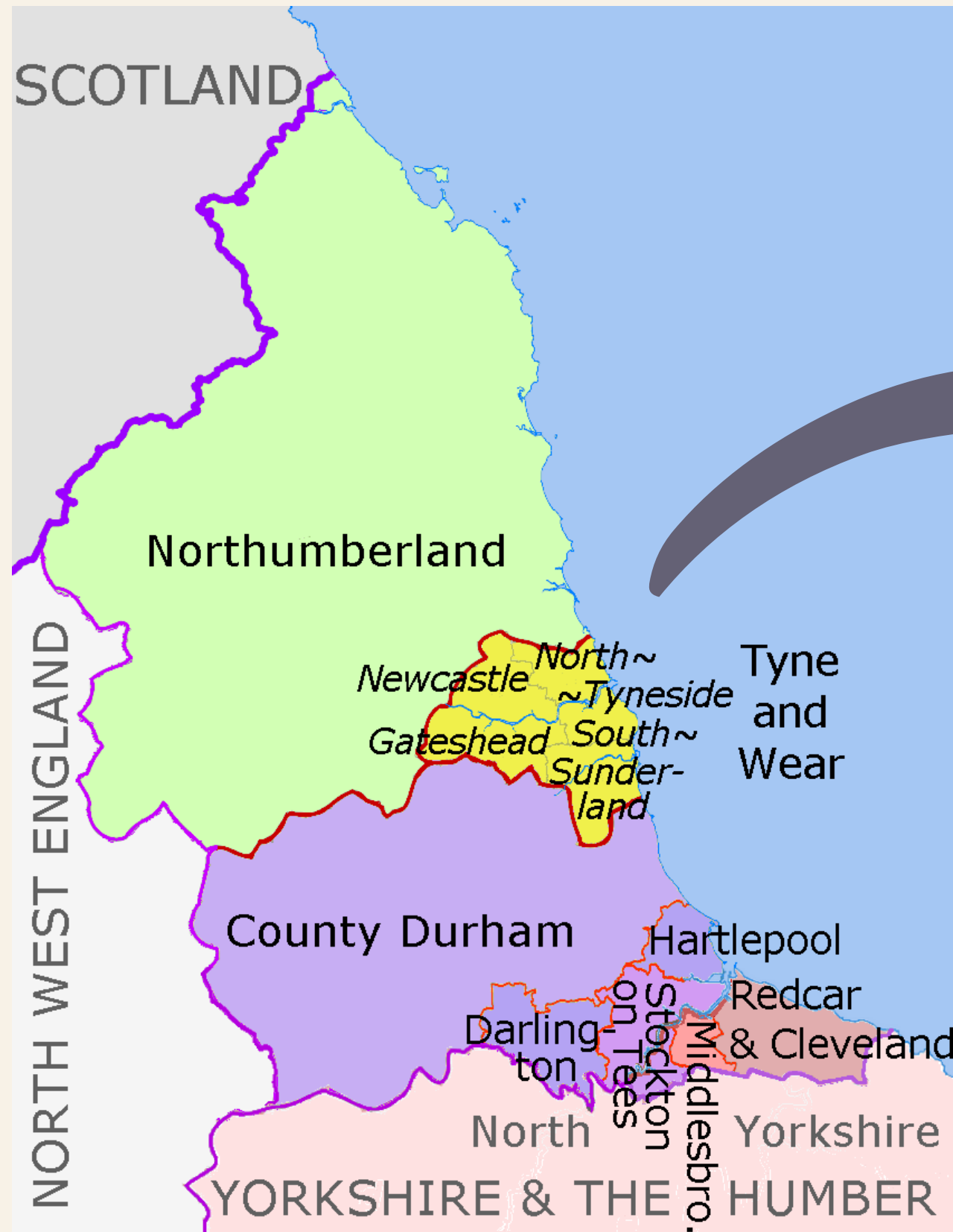


# Local belief: North East of England

- E.g., Tyne and Wear (T&W): Newcastle ("Geordie") and Sunderland ("Mackem").
- Locals' claims on more fine-grained identifiable regional variation.

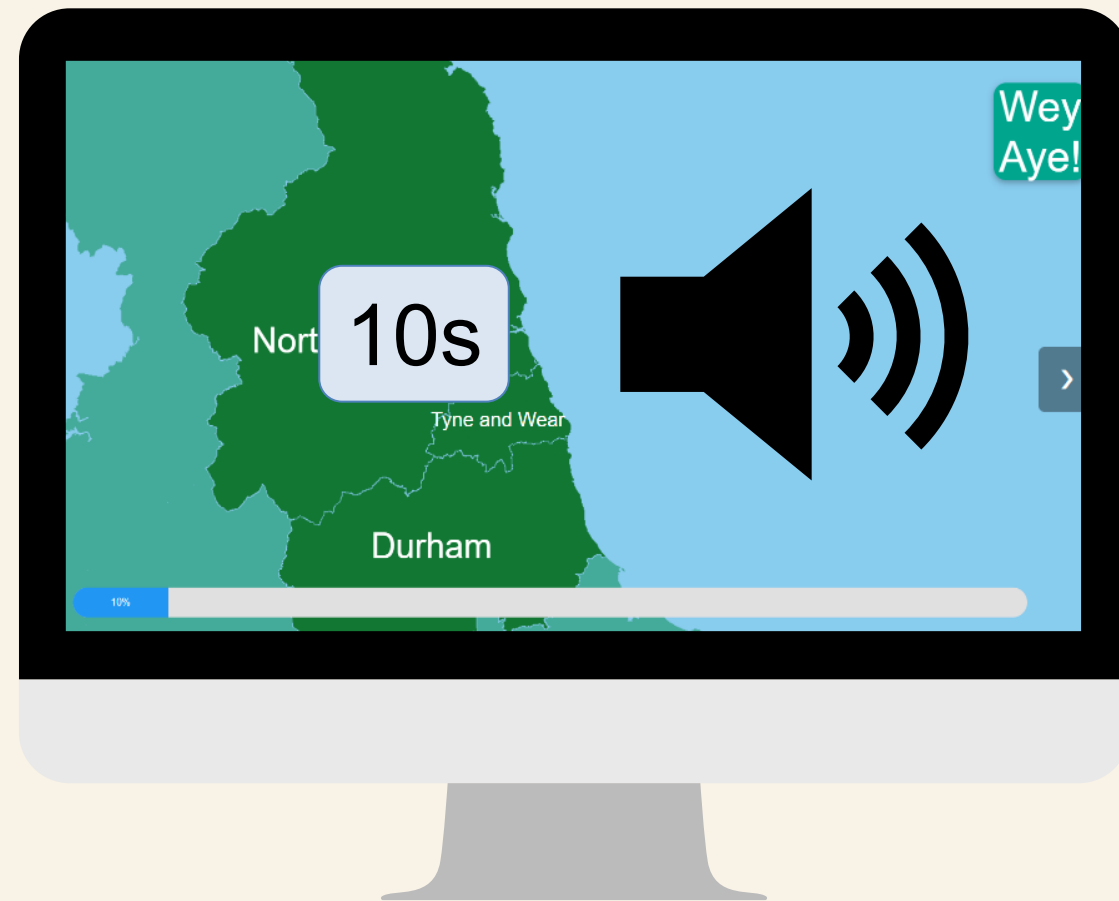


# RQ2: whether listeners can distinguish fine-grained North-East English accents



# Phonetic features

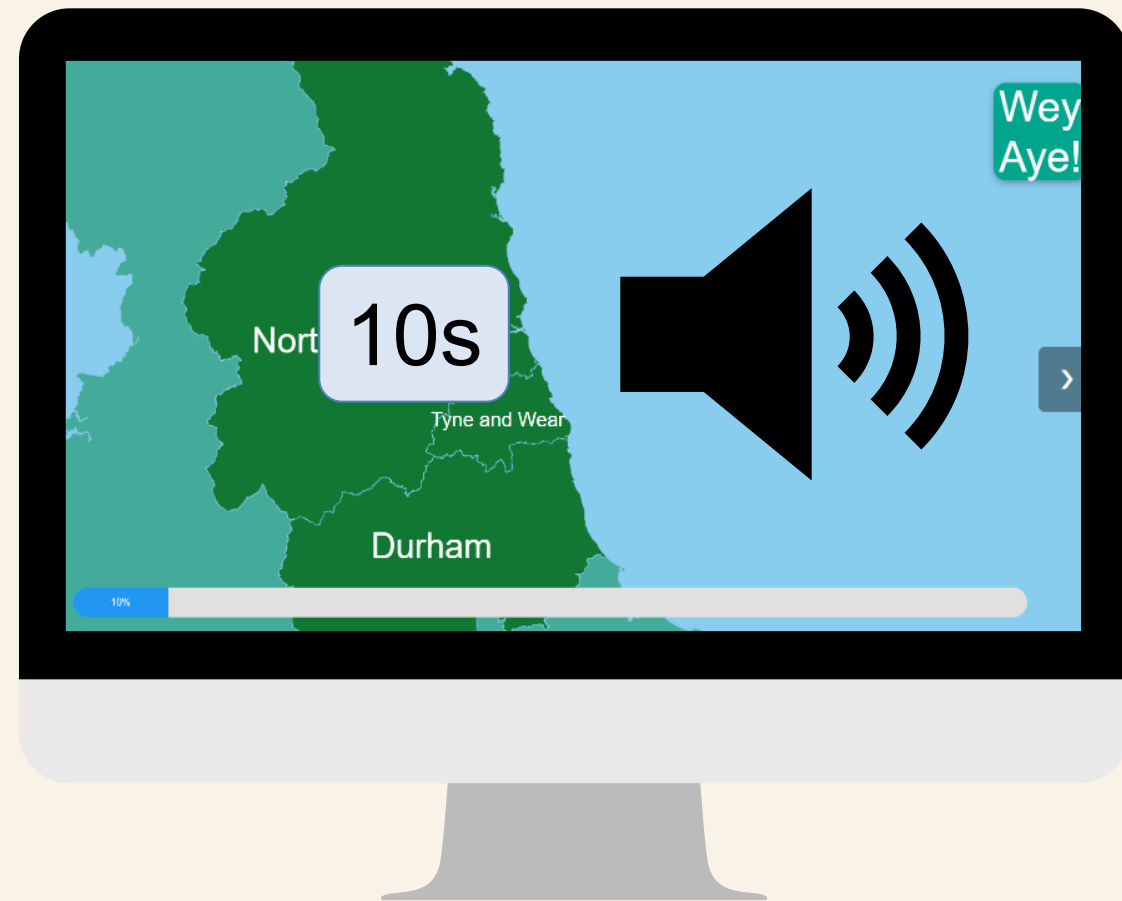
(Pearce 2009, Burbano-Elizondo 2008b)



County level:  
H-dropping and **STAR**

Borough level:  
**NURSE & GOAT**

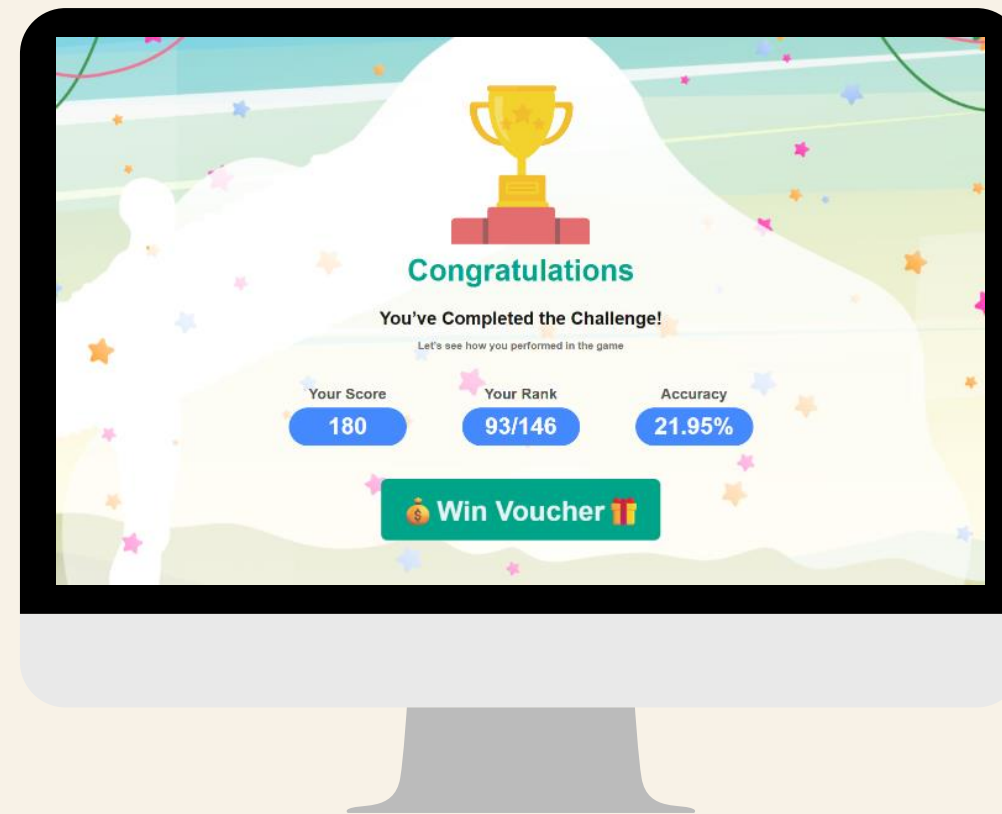
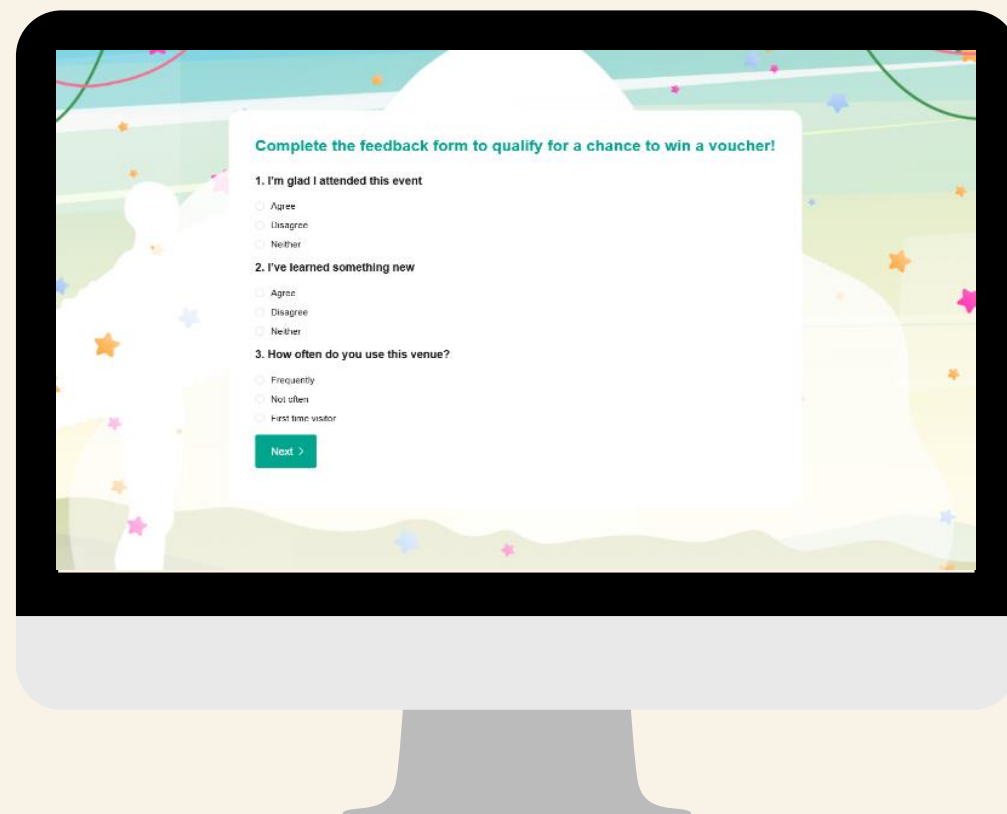
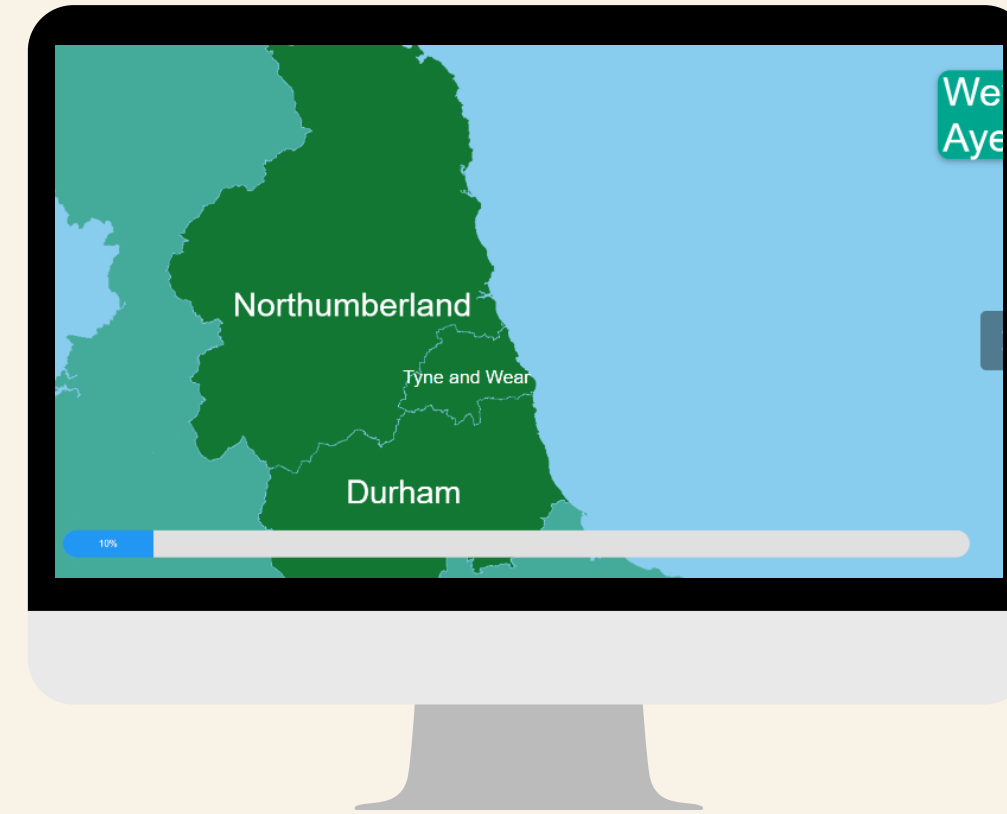
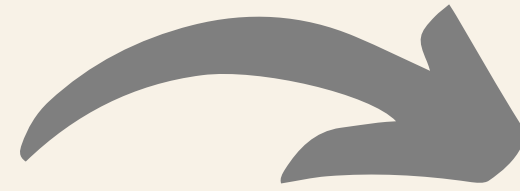
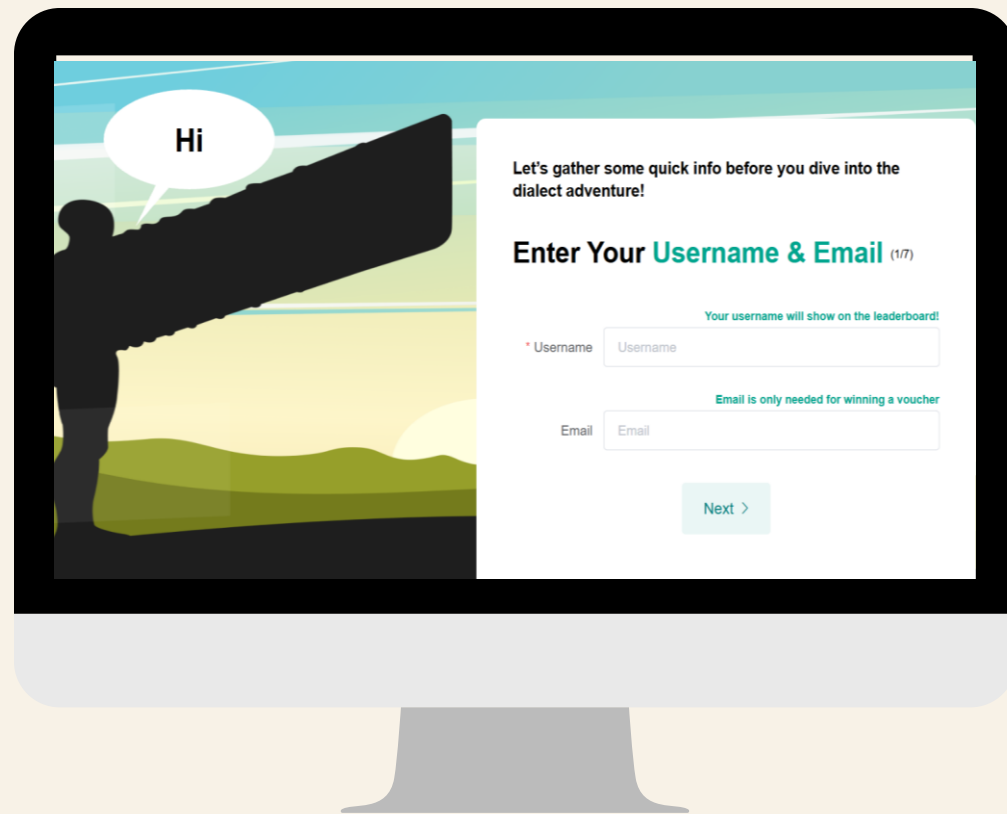
# RQ3: if the selected phonetic features can predict the identification accuracy



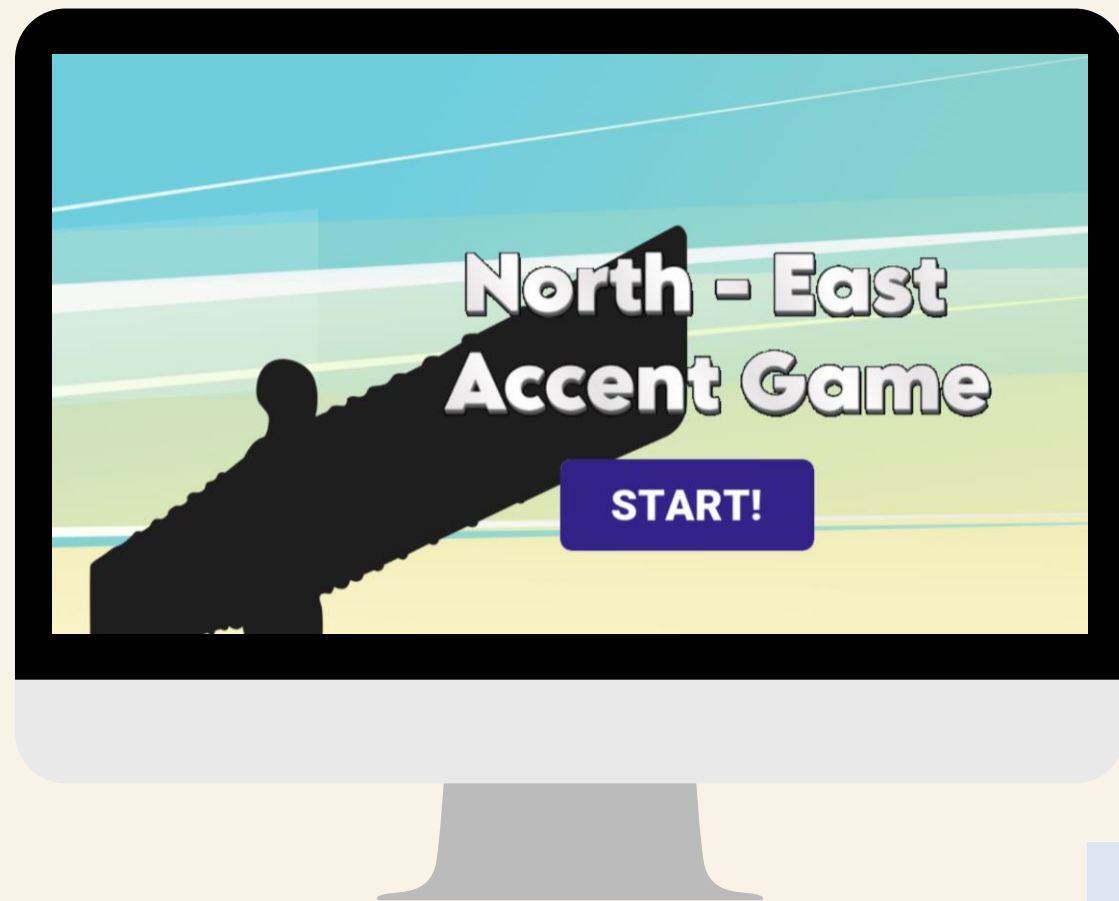
County level:  
H-dropping and STAR

Borough level:  
NURSE & GOAT

# The 'game' procedure

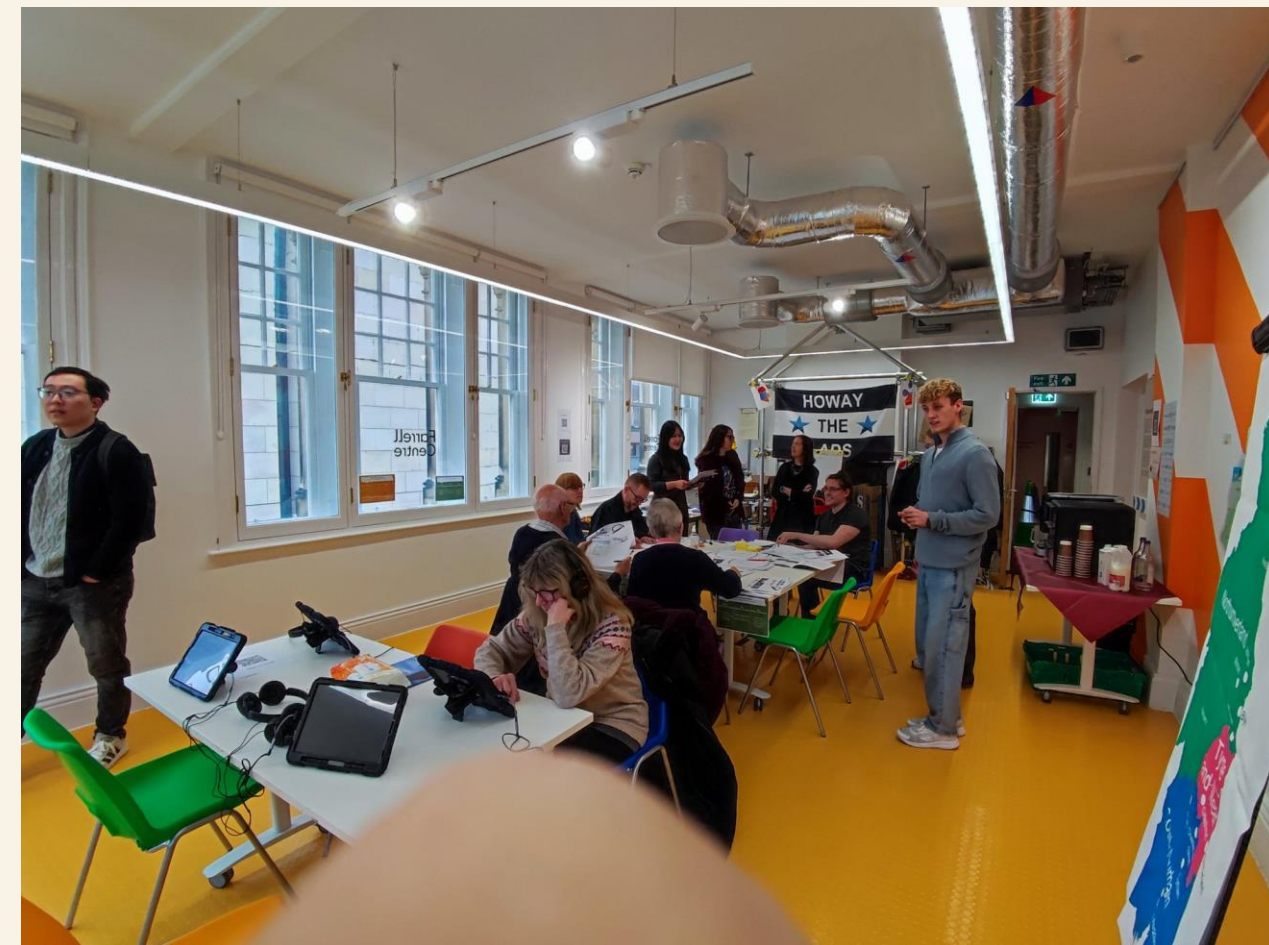


# Data collection

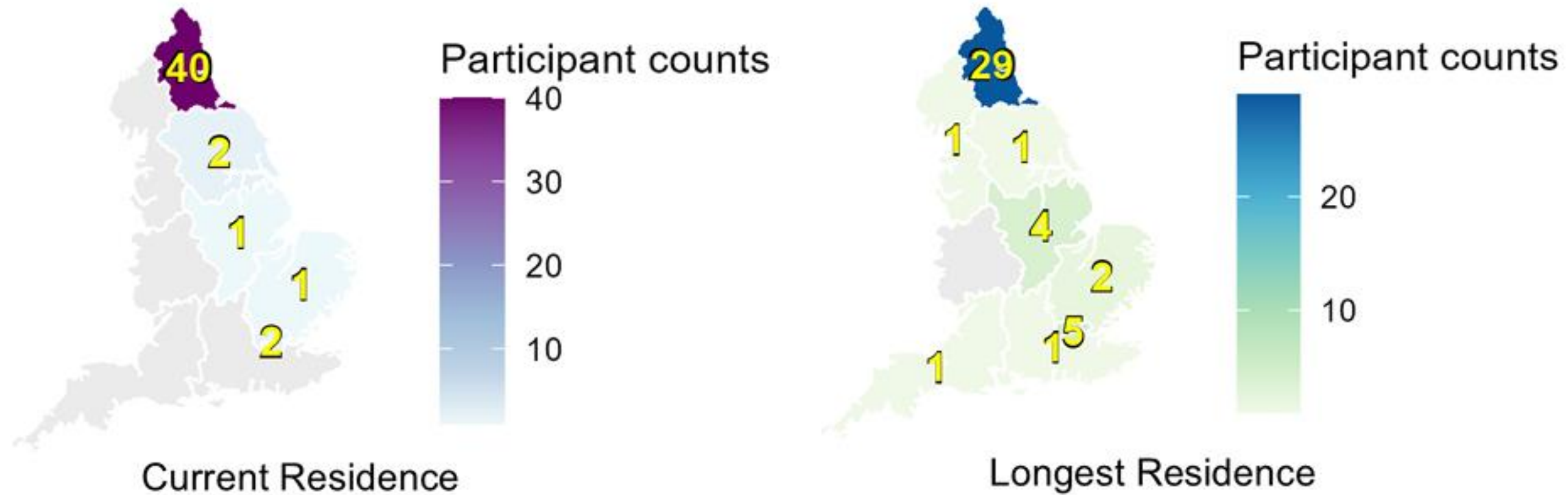


Online N = 47  
(mean age = 31, SD = 9)

Museum outreach  
N = 64



# Most participants are based in Northeast regions



**Fig. 3: Left:** Participant (N=46) distribution by **current residence** in England. *Note: One participant self-reported living in France.* **Right:** Participant (N=44) distribution by **longest residence** in England. *Note: One participant did not report valid information. Two participants lived in Wales and Scotland.*

# Data structure and analysis



Accuracy data

Logistic mixed-effect regression

```
accuracy ~ feature * accent_origin  
+ (1|item) + (1|participant)
```



Random forest

Ranking factors' contribution

```
R package:  
Boruta
```

# Accuracy at county level: T&W most accurate

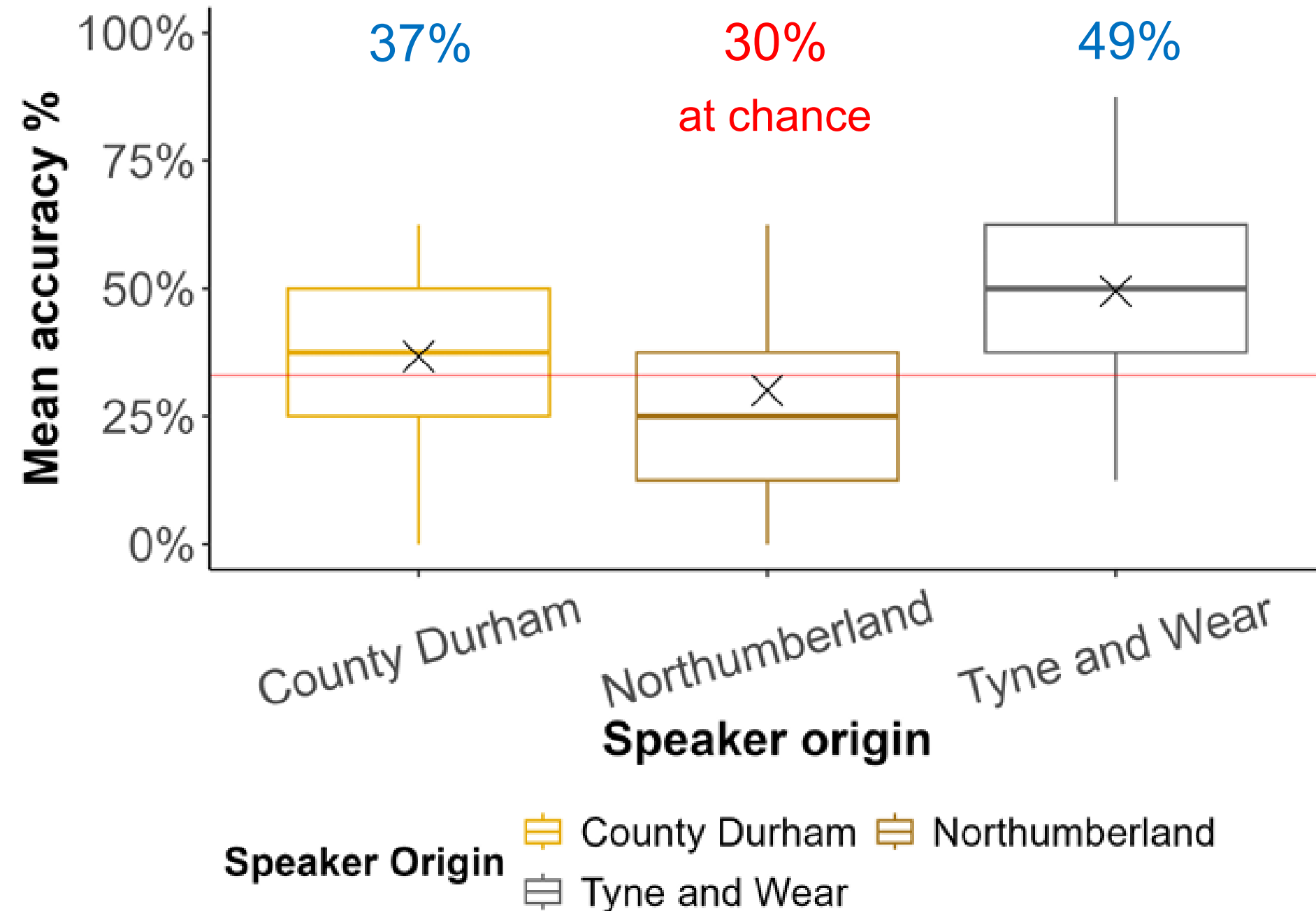
~ accent\_origin

T&W > Northumberland  
 $\beta = 0.88, p < .001$

T&W > County Durham  
 $\beta = 0.55, p = .059$

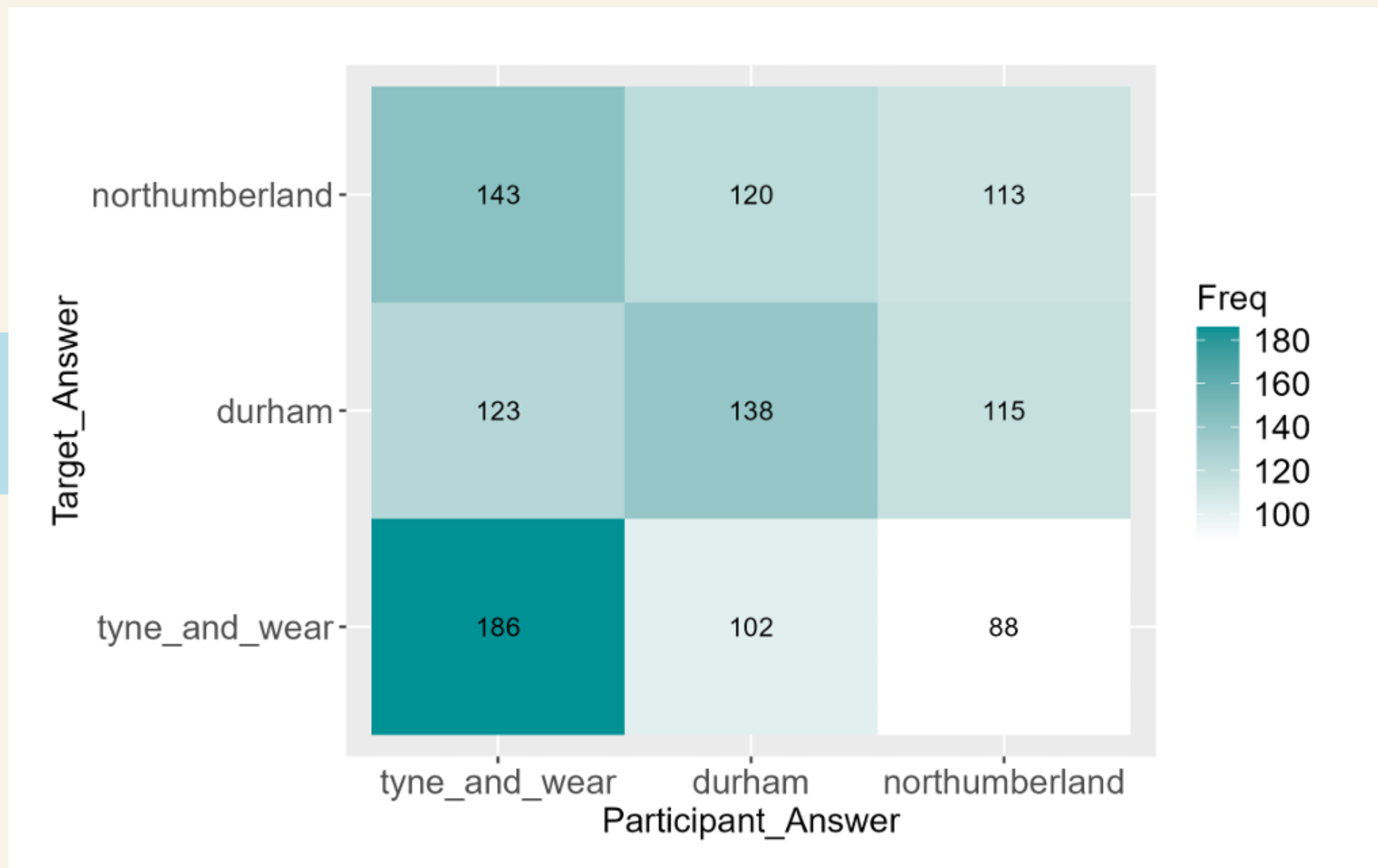
Phonetic features:

H-dropping and STAR



# But people tend to select T&W more often...

Phonetic features:  
H-dropping and STAR

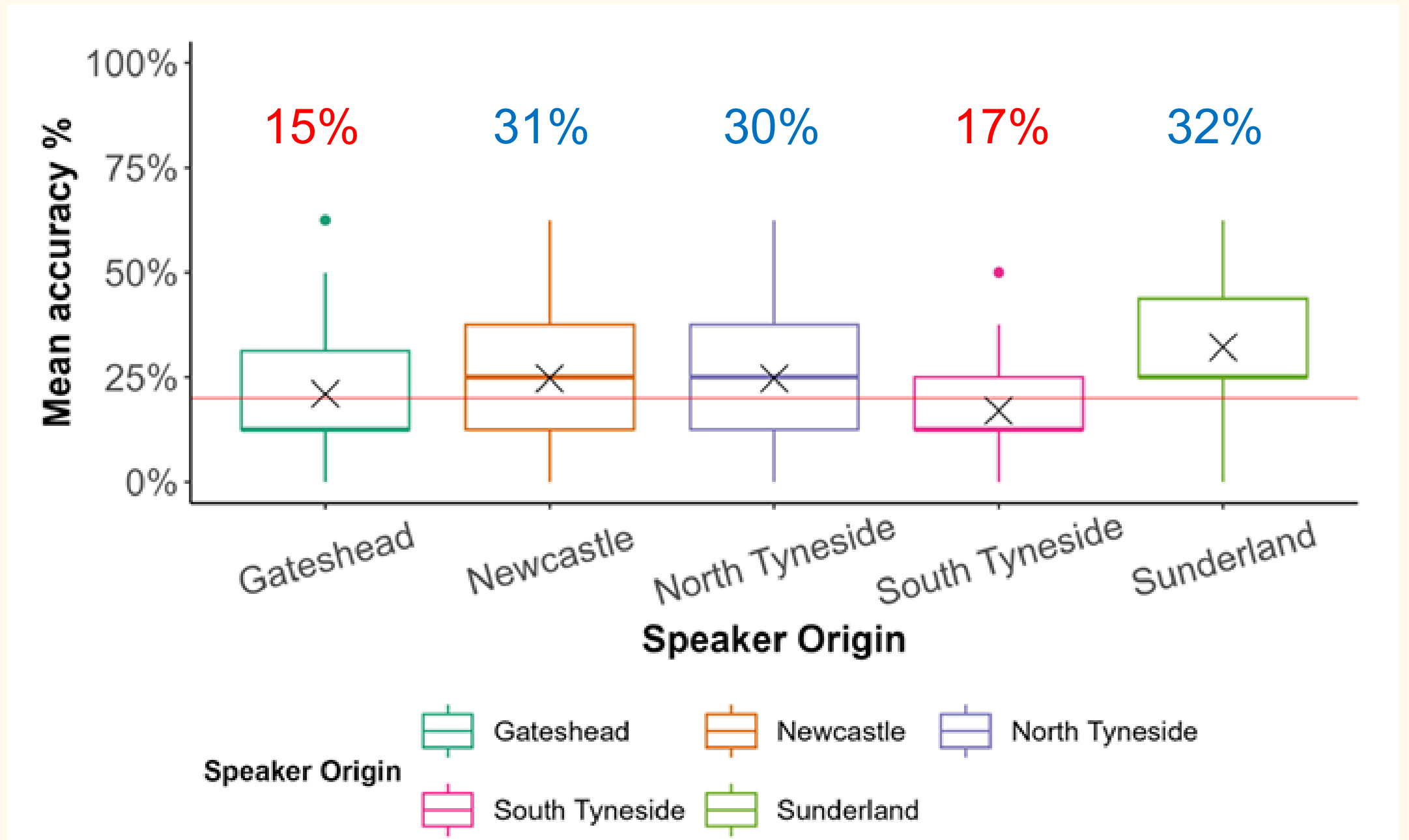


# At borough level: Sunderland most accurate

~ accent\_origin

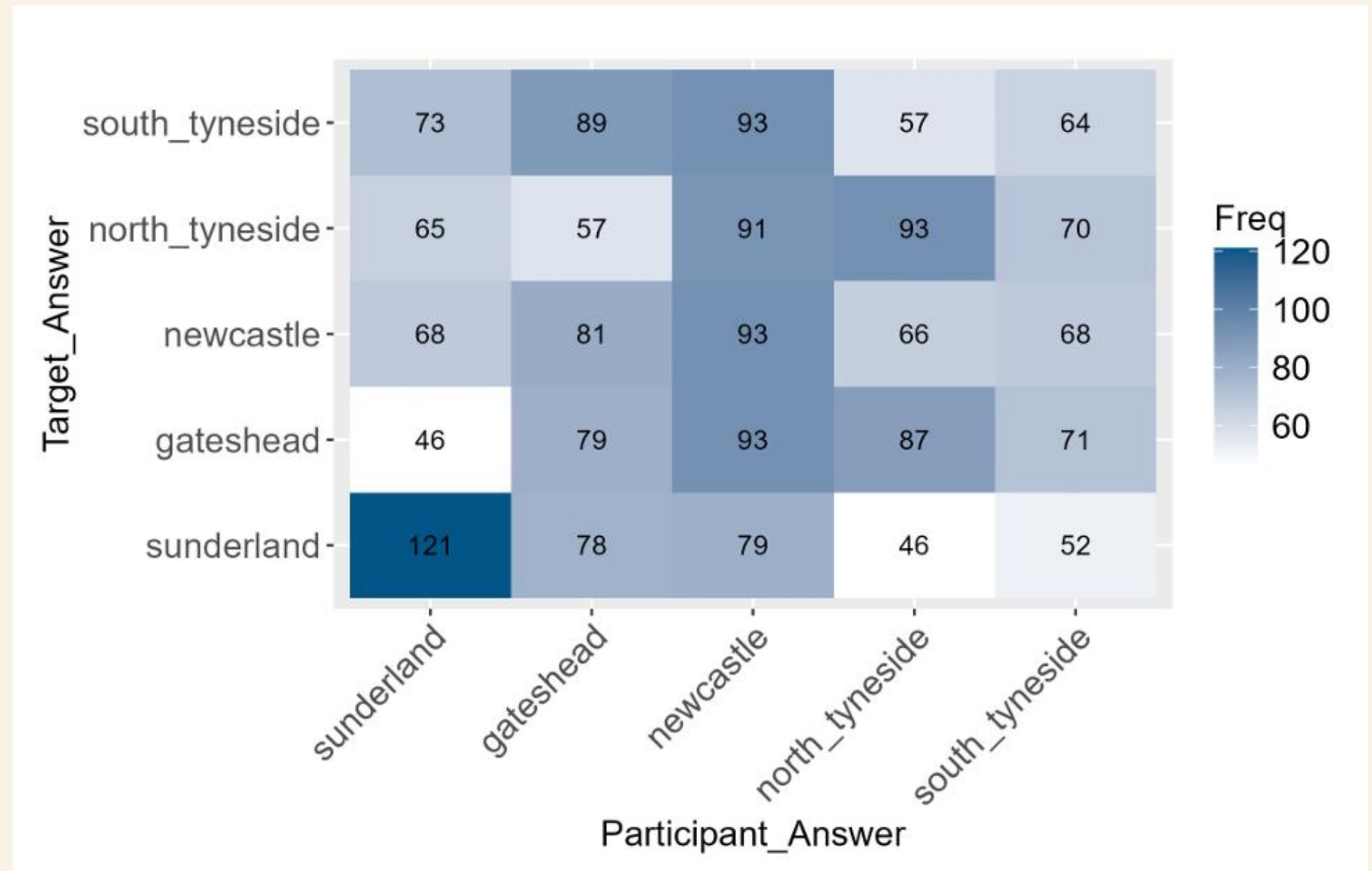
Sunderland > South Tyneside:  
 $\beta = 0.85, p = 0.004$

Phonetic features:  
**NURSE & GOAT**



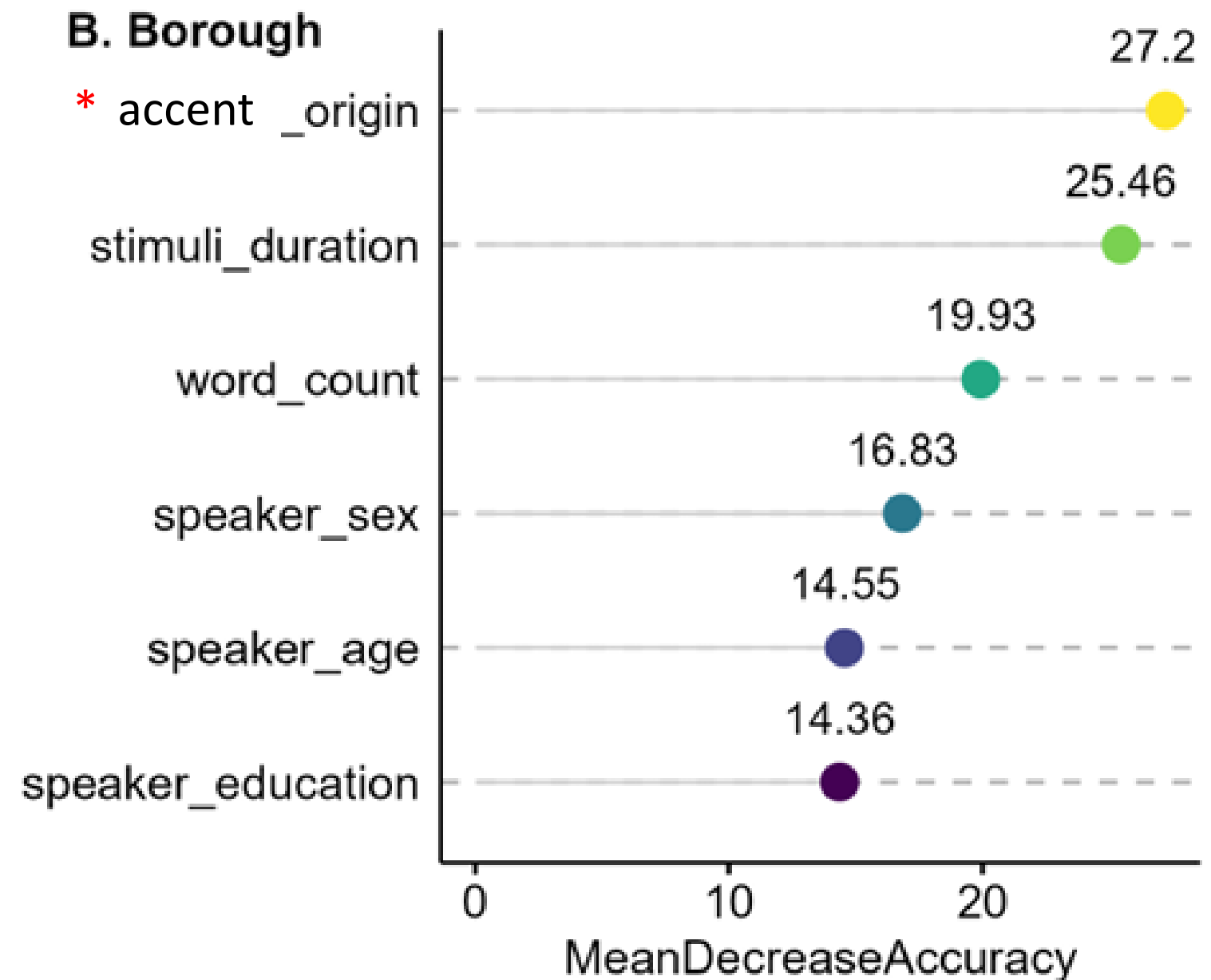
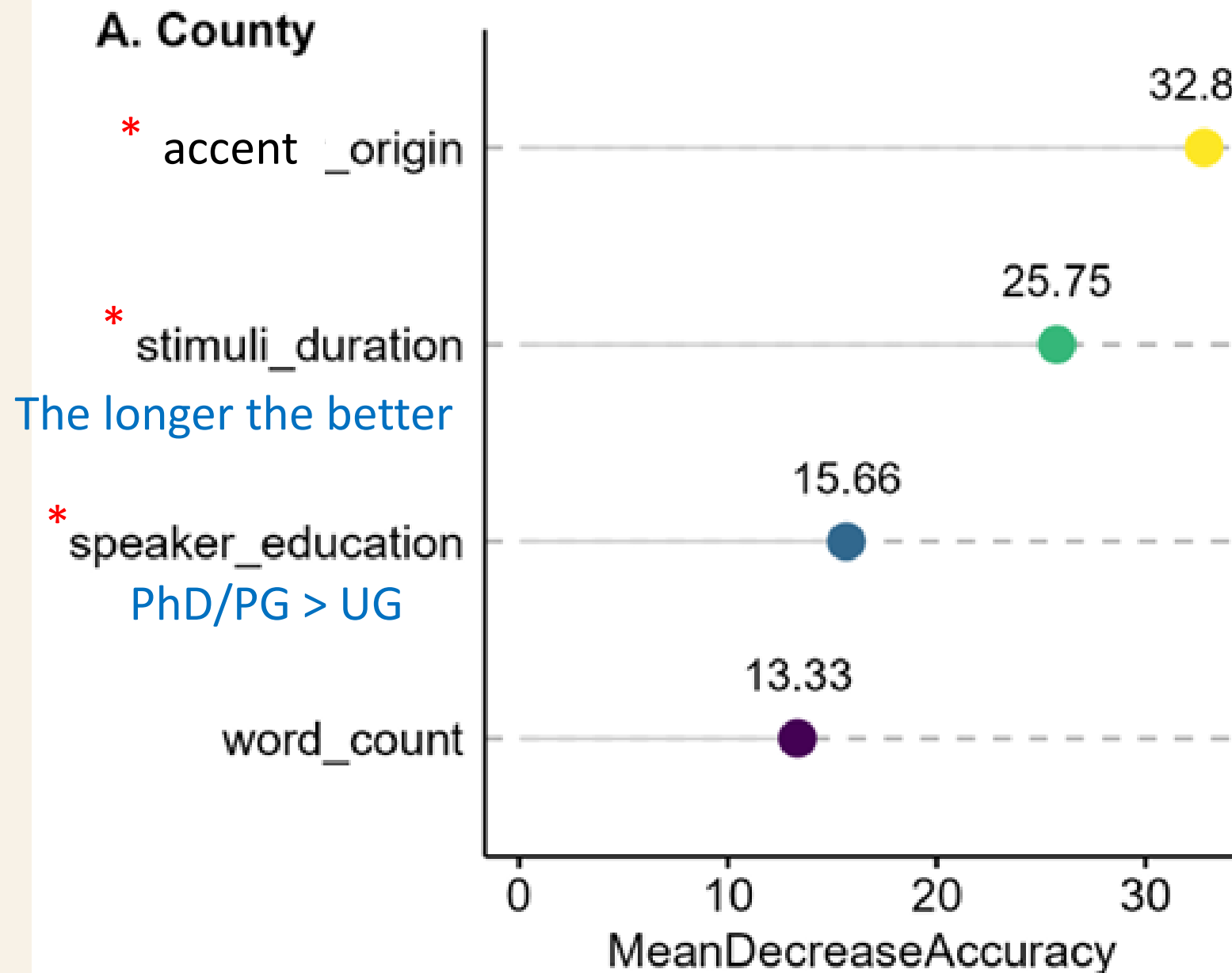
# People misattribute speakers between neighbouring regions

Phonetic features:  
**NURSE & GOAT**



# Feature ranking

## random forest analysis



# To answer the RQs...

RQ1: If gamification methods can be used in phonetic data collection?

1

Yes

To replace traditional per-participant incentive method:  
**£100 vs £300-600**

RQ2: whether listeners can distinguish North-East English accents

2

Not really...

But this can be due to stimuli design, e.g., longer duration might be better

RQ3: if the selected phonetic features can predict the identification accuracy

3

Not really...

Requires more controlled feature input in stimuli design

# Future potentials

Game-building platforms are not accessible

- Building & maintenance
- Server issues...

Go beyond individual game elements (e.g., points-badges-leaderboard design (PBL))

# THANK YOU!



Paper

Link to paper: [https://yanyu-li.github.io/assets/paper/LiEtAl\\_2026\\_Gamification\\_accepted.pdf](https://yanyu-li.github.io/assets/paper/LiEtAl_2026_Gamification_accepted.pdf)

Game website: <https://research.ncl.ac.uk/ne-accent-games/game/>

My website: <https://yanyu-li.github.io/>



Game



Yanyu Li  
PhD Speech and Language Sciences



Damar Hoogland  
Postdoc Jožef Stefan Institute, Slovenia



Daniel Duncan  
Senior Lecturer in Sociolinguistics



Cong Zhang  
Lecturer in Phonetics and Phonology