Vowel Tuner

Soklong HIM Nora LINDVALL Maxime MÉLOUX
Jorge VASQUEZ-MERCADO

NLP M2

Software Project
Nov. 18, 2022

Outline

1. Introduction
2. Rule-based approach
3. Deep learning
4. Timeline

Introduction

Our idea

Main goal

Help language learners improve their pronunciation of French oral vowels

Two approaches

- Rule-based approach
- Deep learning approach

Rule-based approach

Updated results

	i	y	e	ø	ɛ	œ	a	u	o	ɔ
F1	**273** (25, 9.2%)	**273** (25, 9.2%)	400 (40, 10%)	400 (35, 8.6%)	624 (50, 8.1%)	585 (72, 12.3%)	**805** (130, 16.1%)	288 (28, 9.2%)	417 (36, 8.3%)	577 (80, 13.9%)
F2	**2524** (271, 10.7%)	2037 (125, 6.1%)	2504 (205, 8.2%)	1571 (127, 8.1%)	2244 (136, 6.1%)	1579 (154, 9.8%)	1301 (178, 13.7%)	**770** (70, 9.1%)	791 (80, 10.1%)	1063 (89, 8.4%)
F3	**3787** (227, 6%)	**2460** (190, 7.7%)	3364 (165, 4.9%)	2645 (173, 6.5%)	3066 (165, 5.4%)	2751 (229, 8.3%)	2832 (167, 5.9%)			
F4	4428 (195, 4.4%)									

Figure: Formant data from *Do Isolated Vowels Represent Vowel Targets in French? An Acoustic Study on Coarticulation*

Results (corpus 1)

Subset	2 formants	3 formants	4 formants
Native	0.193 (+0.073)	0.289 (+0.156)	**0.398** (+0.241)
Non-native	0.114 (-0.066)	0.170 (-0.035)	**0.273** (+0.159)
Female	0.168 (-0.010)	0.198 (-0.010)	**0.297** (+0.129)
Male	0.129 (+0.029)	0.271 (+0.127)	**0.386** (+0.300)
Overall	0.152 (+0.006)	0.228 (+0.058)	**0.333** (+0.198)

Table: Accuracy between the detected and perceived vowels in the InterFra sub-corpus

Corpus 2

- +20 points overall!
- More balanced across categories
- Performance now scales with formants
- Can be improved further using co-articulation

→ What about the experimental corpus?

Results (corpus 2)

Subset	2 formants	3 formants	4 formants
Native	0.312 ($\approx$ 0)	0.203 (-0.156)	0.281 (-0.063)
Non-native	0.487 (+0.128)	0.308 (+0.042)	0.410 (+0.128)
Female	0.462 (+0.136)	0.308 (+0.116)	0.308 (+0.270)
Male	0.351 (+0.026)	0.221 (-0.143)	0.338 (-0.078)
Overall	0.379 (+0.049)	0.243 (-0.077)	0.330 (+0.010)

Table: Accuracy between the detected and perceived vowels in the experimental corpus

Analysis

- Significantly better!
- 2 or 4 formants?
- Performance for 3 formants?
- Reference vowels (gender)
- Still a single annotator, highly subjective

Is the linguistic approach still viable?

Analysis

	F1				F2				F3			
	p	t	k	R	p	t	k	R	p	t	k	R
i	0.1	0.2	0.1	1.2	- 0.3	- 0.2	0.1	0	- 0.4	- 0.5	- 0.2	- 0.6
y	0	0.1	0.1	0.9	0	0.3	0.3	- 1.1	0.2	0.4	0	0.6
e	0.3	0.3	0.2	0.8	- 0.5	- 0.5	- 0.1	- 0.5	- 0.5	- 0.5	- 0.3	- 0.6
ø	0.2	0.2	0.2	0.6	0.2	0.6	0.6	-1.1	0	0.2	0.3	0.3
ɛ	0.1	-0.2	-0.5	0.5	- 0.4	- 0.3	0.2	- 0.7	- 0.2	- 0.1	- 0.1	- 0.1
œ	0.2	0	-0.2	0.6	0.2	0.7	0.9	- 0.7	0	0.2	0.4	0.1
a	0.1	-0.3	-0.6	0.2	1.2	2.1	2.7	0.6	0	0.2	- 0.2	0
u	0.1	0.1	0.1	0.4	0.6	2.7	0.5	- 0.4				
o	0.1	0.1	0.2	0.2	0.7	2.4	1	0.1				
ɔ	0.5	0.3	0.3	0.7	1	2	1.1	0.3				
Average	0.2	0.1	0	0.6	0.3	1	0.7	-0.4	-0,1	0	0	0

Figure: Co-articulation formant data from [1]

Analysis

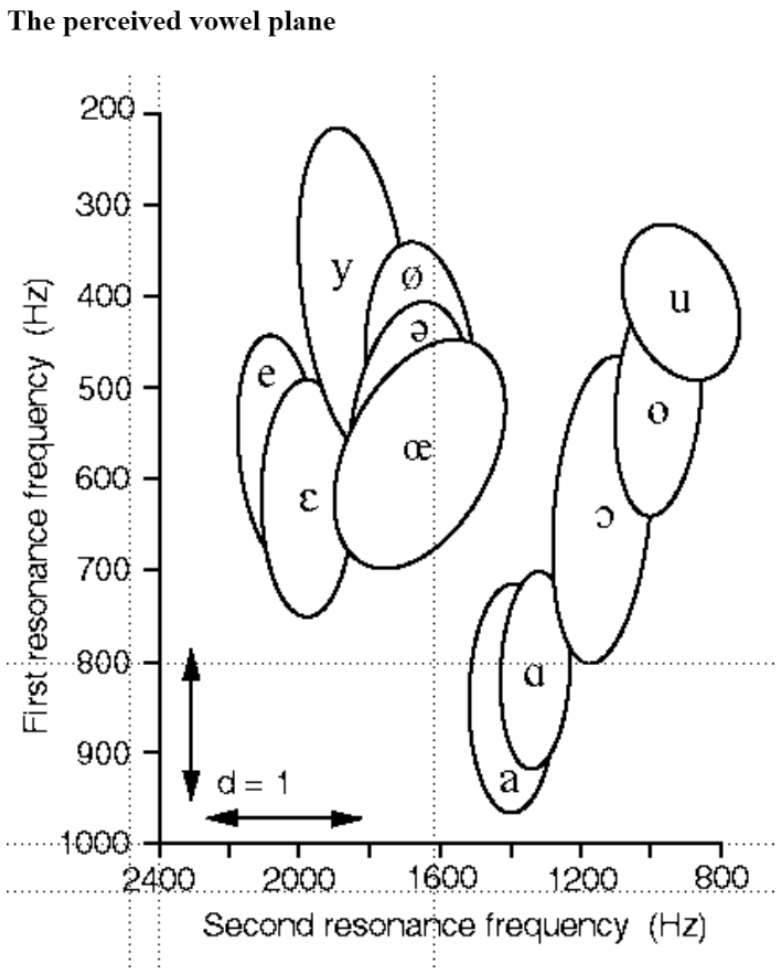


Figure: Data from *Annette Dowd, John Smith and Joe Wolfe*

Deep learning

Native dataset

Good news!

Native dataset

We found a great corpus!

- 8 female + 12 male French native speakers
- 14 vowel phonemes
- 63 min of recordings
- audio and video recordings

Corpus example:
pi, lit, mi, si, ti, ti
pu, lu, mu, su, tu, tu

Data collection

Data collection

- Date:TBA (current plan: Dec 5th)
- Environment: LORIA Recording room (awaiting details)
- Recording time: 2-3 minutes (reading 84 French syllables - short!)
- You will receive: our eternal gratitude and a little snack
- Your French level doesn't matter (beginners welcome)
- All native and non-native speakers can join!

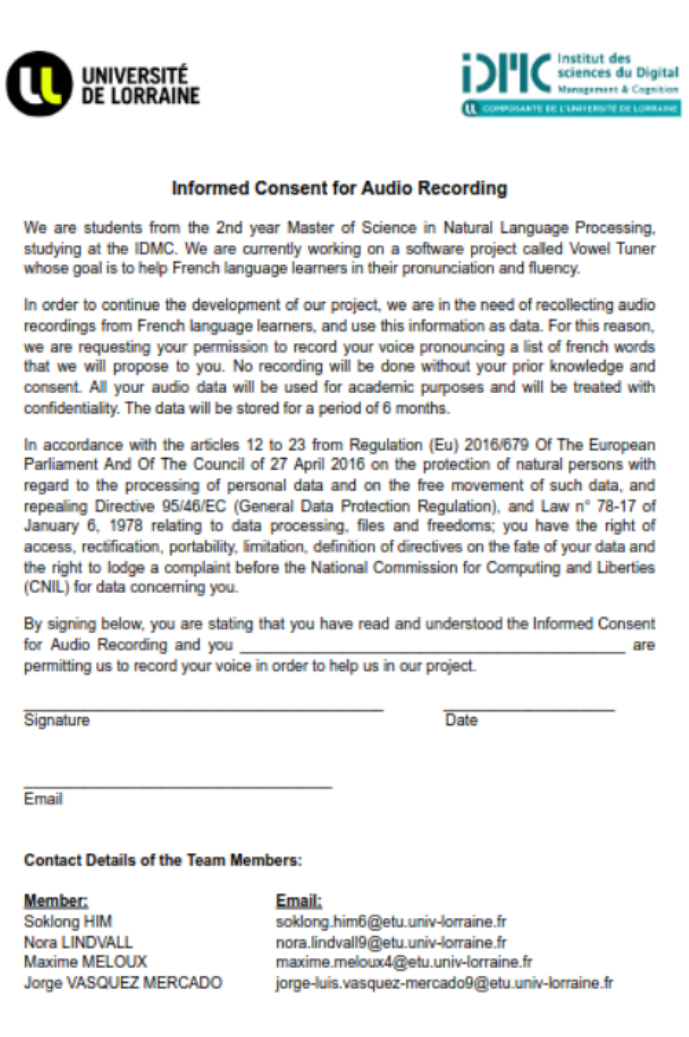
Consent form

Our consent form is based in the General Data Protection Regulation (GDPR), article 13. The information provided to the data subject through the consent form are:

- Giving the contact details of the controllers of the data.
- Giving the purpose of the data collection
- Period for which the personal data will be stored
- Explaining the rights the data subject has according to the Regulation (Eu) 2016/679 Of The European Parliament And Of The Council of 27 April 2016 and Law n° 78-17 of January 6, 1978 relating to data processing, files and freedoms

Consent form

A preview of our consent form



UNIVERSITÉ DE LORRAINE

IDMC Institut des sciences du Digital Management & Cognition

COMPOSANTE DE L'UNIVERSITÉ DE LORRAINE

Informed Consent for Audio Recording

We are students from the 2nd year Master of Science in Natural Language Processing, studying at the IDMC. We are currently working on a software project called Vowel Tuner whose goal is to help French language learners in their pronunciation and fluency.

In order to continue the development of our project, we are in the need of recollecting audio recordings from French language learners, and use this information as data. For this reason, we are requesting your permission to record your voice pronouncing a list of french words that we will propose to you. No recording will be done without your prior knowledge and consent. All your audio data will be used for academic purposes and will be treated with confidentiality. The data will be stored for a period of 6 months.

In accordance with the articles 12 to 23 from Regulation (Eu) 2016/679 Of The European Parliament And Of The Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation), and Law n° 78-17 of January 6, 1978 relating to data processing, files and freedoms; you have the right of access, rectification, portability, limitation, definition of directives on the fate of your data and the right to lodge a complaint before the National Commission for Computing and Liberties (CNIL) for data concerning you.

By signing below, you are stating that you have read and understood the Informed Consent for Audio Recording and you _____ are permitting us to record your voice in order to help us in our project.

_____ _____

Signature Date

_____

Email

Contact Details of the Team Members:

Member:	**Email:**
Soklong HIM	soklong.him6@etu.univ-lorraine.fr
Nora LINDVALL	nora.lindvall9@etu.univ-lorraine.fr
Maxime MELOUX	maxime.meloux4@etu.univ-lorraine.fr
Jorge VASQUEZ MERCADO	jorge-luis.vasquez-mercado9@etu.univ-lorraine.fr

Annotation questionnaire

1) Is French one of your native languages?
☐ Yes ☐ No Stop here

2) What French-speaking country and region would you say your French was most influenced by? This can be the place where you or your parents grew up, for example.
[Free input] Try to keep only northern/Parisian French variants

3) Do you pronounce these words the same?
il **fait** - une **fée**
☐ Yes ☐ No
deux (2) - **de**
☐ Yes ☐ No
brun (couleur) - **brin** (de paille)
☐ Yes ☐ No
pâte (à pizza) - **patte** (de chien)
☐ Yes ☐ No
Hope for: N/A - yes - yes - yes. possibly adapt answers to questions otherwise

---------- For each annotation ----------
4) What sound do you hear in this recording [after the d, before the r]?
☐ The **o** in "un **mot**"
☐ The **o** in "un **mort**"
☐ The **eu** in "la **peur**"
☐ The **eu** in "un **jeu**"
☐ ...
☐ I'm not sure
☐ Other/several of these are the same (please explain)

5) Do you think French is a native language of this person?
☐ Yes skip next question
☐ No

6) How natural does this person's pronunciation sound? Please only focus on the sound you selected in question 4 and not the sounds before/after.
☐ Unrecognizable
☐ Unnatural
☐ Almost natural

Timeline

Updated timeline

Date	Task
~~15 Nov~~ <23 Nov	Corpus collection meeting
~~25 Nov~~ 5 Dec	Finish data collection
~~9 Dec~~ 12 Dec	Finish annotation
16 Dec	Model trained, tested and evaluated
20 Jan	Interface final version
26 Jan	Finish report
27 Jan	Project deadline

Mitigation plans

Plan B: Use only native data

Plan C: Abandon deep learning and perfect rule-based approach

Thank you!

Questions? Feedback?