



HAL
open science

Les effets des voyelles finales et des sous-systèmes lexicaux sur l'accentuation en anglais

Quentin Dabouis, Marie Gabillet

► **To cite this version:**

Quentin Dabouis, Marie Gabillet. Les effets des voyelles finales et des sous-systèmes lexicaux sur l'accentuation en anglais. RFP 2024 - Réseau Français de Phonologie, CNRS CERCLL, MODYCO, SFL, STL, Jun 2024, Amiens, France. hal-04834843

HAL Id: hal-04834843

<https://uca.hal.science/hal-04834843v1>

Submitted on 12 Dec 2024

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution 4.0 International License

The effects of final vowels and sublexica on stress in English

Quentin Dabouis & Marie Gabillet
Université Clermont Auvergne – LRL (UR 999)

Objectif

Focus : accentuation des noms longs sans morphologie

Réanalyse des données de [Moore-Cantwell \(2020\)](#) à l'aide du modèle de sous-systèmes lexicaux proposé par [Dabouis & Fournier \(2022\)](#)

Le rôle des voyelles finales

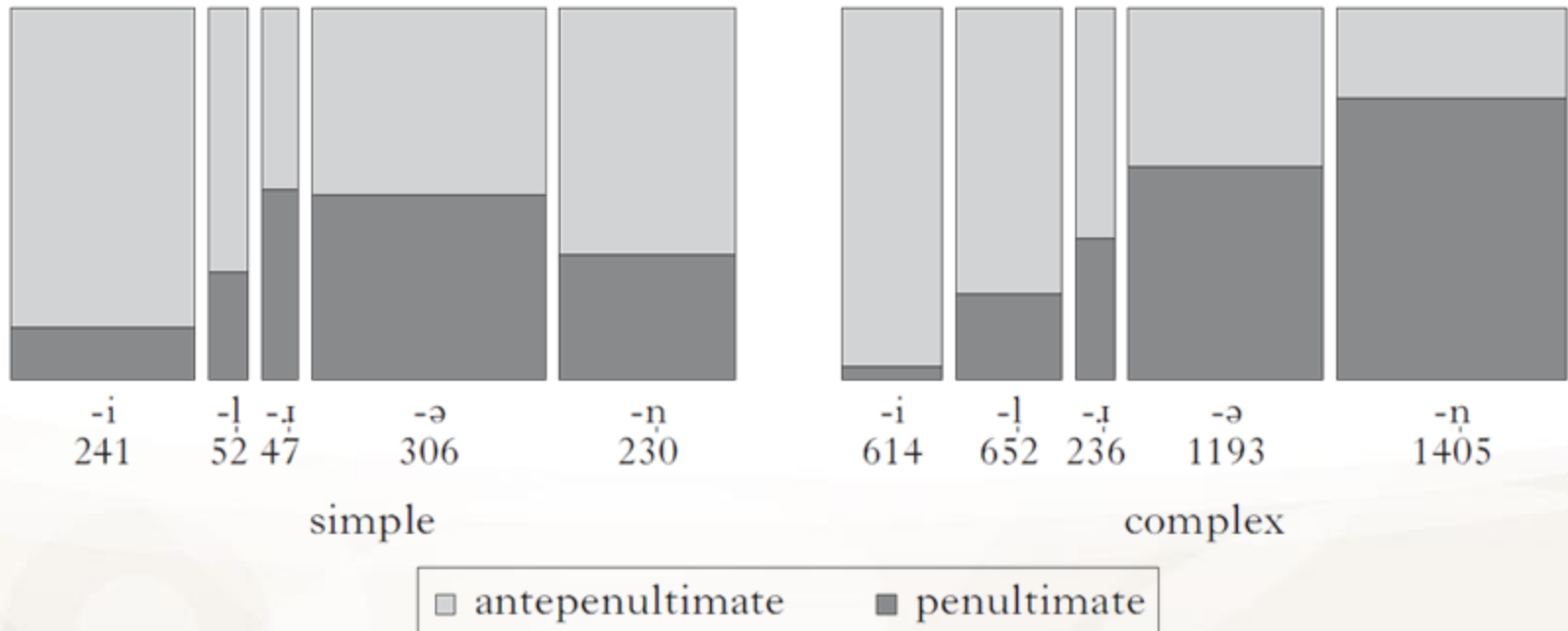
Moore-Cantwell (2020), deux études :

- Dictionnaire
- Expérimentale (production de non-mots)

Résultats :

/i/ dans la dernière syllable favorise l'accentuation antépénultième en comparaison des autres noyaux

Le rôle des voyelles finales



Cet effet est indépendant de celui du poids de la pénultième

Hypothèses

Hypothèse de Moore-Cantwell : le noyau final des mots longs joue un rôle dans l'attribution de l'accent principal en anglais

- Un noyau final /i/ repousse l'accent vers la gauche

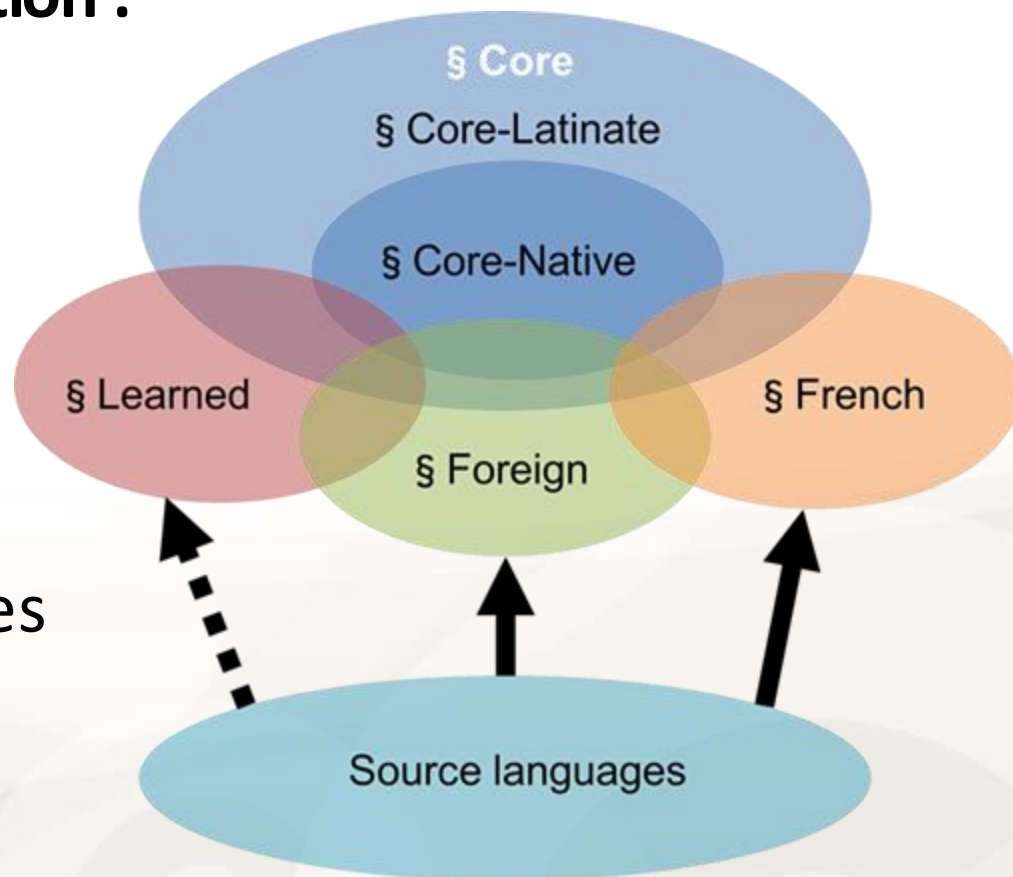
Notre hypothèse : les effets du noyau final sont largement dus au fait que différents noyaux se trouvent dans différentes strates du lexique.

Le modèle de Dabouis et Fournier (2022)

Proposition de formalisation :

Division du lexique de l'anglais en 4 sous-systèmes, possédant chacun des propriétés spécifiques

- phonologiques
- graphophonologiques
- morphologiques
- sémantiques



Le modèle de Dabouis et Fournier (2022)

- **§CORE** : la majeure partie du vocabulaire. Correspondances graphophonologiques régulières, sémantique non spécialisée et procédés morphologiques productifs
- **§FRENCH** : vocabulaire spécifique à certains domaines (e.g. alimentation), avec une notion de « prestige ». Proportion plus élevée d'accentuation finale (primaire ou secondaire), correspondances graphophonologiques spéciales (e.g. <i> - /i/)

Le modèle de Dabouis et Fournier (2022)

- **§FOREIGN** : mots empruntés, ou associés, aux langues romanes principalement (à l'exclusion du français). Massivement accentués sur la pénultième, réfèrent à des cultures étrangères (e.g. nourriture, vêtements, traditions), correspondances graphophonologiques spéciales (e.g. <i> - /i/)
- **§LEARNED** : vocabulaire technique ou scientifique, souvent directement emprunté, ou construit à partir d'éléments empruntés au Grec ou au Latin. Composés savants inclus

Le modèle de Dabouis et Fournier (2022)

En partie basé sur Fournier (2007, 2010), sur les mots à syllabe pénultième fermée :

- Accent pénultième quasi-systématique dans les emprunts aux langues romanes (e.g. *ànacónda*, *dilettánte*, *concérto*) et au latin moderne (e.g. *enígma*, *meménto*, *alúmnus*)

- Mots restants :
 - 2/3 accentués sur l'antépénultième (e.g. *díscipline*, *sácristan*, *cáalendar*, *lávender*, *mínister*)
 - 1/3 accentués sur la pénultième (e.g. *advántage*, *intéstine*, *advénture*, *epáulment*, *còriánder*, *seméster*, *elíxir*)

Le modèle de Dabouis et Fournier (2022)

Présupposés importants :

- L'accentuation pénultième n'a pas la même prédominance selon le sous-système : §FOREIGN > §LEARNED > §CORE
- distribution asymétrique des noyaux finaux selon les sous-systèmes

Implication :

- la prédominance de l'accentuation antépénultième dans les mots en [i] est une conséquence du fait que de nombreux mots en [i] appartiennent à §CORE

Réanalyse des données dictionnairiques

Récupération des données de [Moore-Cantwell \(2020\)](#), issues du dictionnaire de prononciation de l'Université de Carnegie Mellon (CMU, [Weide 1994](#)).

Première étape :

Recodage nécessaire, notamment morphologique

Réanalyse des données dictionnairiques

Analyse morphologique

Dans l'article : codage automatique basé sur l'orthographe à l'aide d'une liste d'affixes de [Teschner & Whitley \(2004\)](#), ainsi qu'une vérification d'échantillons aléatoires

Vérification manuelle de chaque mot : 564 mots (7%) à reclassifier.

- Comparable à la mesure de qualité de la catégorisation (*goodness of categorisation*) présentée dans l'article.

Réanalyse des données dictionnairiques

103 mots ont été reclassifiés de complexes en simples :

e.g. *acacia, alligator, coriander, coward, Indonesia, janitor, possum, traffic*

461 ont été reclassifiés de simples en complexes :

e.g.	Composés	<i>yesterday, nobody, Charleston, Newfoundland</i>
	Savants	<i>triangle, nitrogen, hydrocarbon, analogy, philosophy</i>
	Préfixés	<i>belittle, require, refuel, imperil, tomorrow, uncanny</i>
	Suffixés	<i>avoidance, nauseous, digestive, fiftieth, remembrance</i>

Réanalyse des données dictionnairiques

Accent secondaire final

26 mots avec un « ih » accentué (= /ɪ/) ont été exclus des données utilisées dans l'article.

Par cohérence : exclusion des autres noyaux finaux avec un accent secondaire dans le CMU :

1 /ɪ/ *catapult*

5 /r/ *taciturn, introvert, amateur, extrovert, universe*

45 /i/ *quarantine, manatee, Portuguese, magazine...*

Réanalyse des données dictionnairiques

Erreurs dans les transcriptions du CMU

Corundum, conundrum, addendum, Burundi étaient tous transcrits avec une pénultième courte : soit /n/ syllabique, soit une voyelle courte.

- Tous recodés en Vn.C

Réanalyse des données dictionnairiques

Codage du poids syllabique

Nous n'avons pas eu d'informations sur les classifications syllabiques utilisées dans l'article, ainsi nous avons codé :

- V: syllabe ouverte, voyelle courte
- VV: voyelle longue, syllabe éventuellement fermée
- TC: V + (L) + coda obstruante
- LC: V + sonante (nasale ou liquide)
- SC : V + sC (attaque légale en début de mot)

Nous avons aussi corrigé quelques erreurs de classification H/L.

Réanalyse des données dictionnairiques

Codage des sous-systèmes lexicaux

Les mots morphologiquement simples ont été classés comme appartenant à §FOREIGN, §LEARNED ou §FRENCH s'ils présentaient des caractéristiques spécifiques à ces sous-systèmes.

En l'absence de ces caractéristiques : §CORE

Réanalyse des données dictionnairiques

Codage des sous-systèmes lexicaux

19 mots ont été classés dans §FRENCH à cause de leurs terminaisons *-ie*, *-esque* et *-ette* ou bien car ils présentaient l'une des correspondances graphophonologiques suivantes :

<qu> - /k/ <ou> - /u/

<ch> - /ʃ/ <ie> - /i/

<gn> - /nj/

e.g. *brasserie, camaraderie, chicanery, coterie, couturier, etiquette, insouciant, mannequin, menagerie, picturesque, reverie, rotisserie...*

Réanalyse des données dictionnairiques

	§FOREIGN	§LEARNED	§CORE
Terminaisons	-Ca, -Va (V≠<e>), -Vh, -i, -o	-um, -is, -im, -in (éléments chimiques), -ea	Tous les mots restants
Grapho- phonologie	<e#> non-muet	<ch> - /k/, <ph> - /f/ <rh> - /ʀ/	
Exemples	<i>ballerina, chianti, hyena, nirvana, pakistani, ukulele</i>	<i>addendum, calculus, epiphany, metropolis</i>	<i>agony, dynasty, horizon, mackerel, similar</i>

Classification finale

§CORE: 236

§LEARNED: 82

§FOREIGN: 212

(§FRENCH (exclus): 19)

Analyse de régression

Prédicteurs testés :

PENULTWEIGHT : V, VV, TC, LC, SC

FINALV : α , i , l , η , μ

FINALC : 0, 1, 2 (cf. [Domahs et al 2014](#))

SUBSYSTEM : CORE, FOREIGN, LEARNED

Variable dépendante :

STRESS : PENULT, ANTEPENULT

Analyse de régression

Nous avons implémenté un modèle de régression logistique binaire aux données, avec les variables PENULTWEIGHT, FINALV et FINALC et SUBSYSTEM

<i>Predictors</i>	<i>Odds Ratios</i>	<i>CI</i>	<i>Statistic</i>	<i>p</i>
(Intercept)	5.91	1.65 – 37.72	2.35	0.019
SUBSYSTEM-FOREIGN	13.01	7.80 – 22.38	9.57	<0.001
SUBSYSTEM-LEARNED	2.65	1.33 – 5.22	2.81	0.005
PENULTWEIGHT-SC	0.04	0.00 – 0.27	-3.11	0.002
PENULTWEIGHT-TC	0.20	0.02 – 2.10	-1.40	0.162
PENULTWEIGHT-V	0.02	0.00 – 0.08	-5.06	<0.001
PENULTWEIGHT-VV	1.95	0.17 – 44.36	0.53	0.599
Observations	529			
R ² Tjur	0.423			
AIC	474.354			

Pas d'effet de FINALV et FINALC

Arbre de régression

Autre analyse : arbre de régression

- alternative à la regression logistique
- permet une meilleure analyse des interactions complexes entre prédicteurs ([Levshina 2021](#))

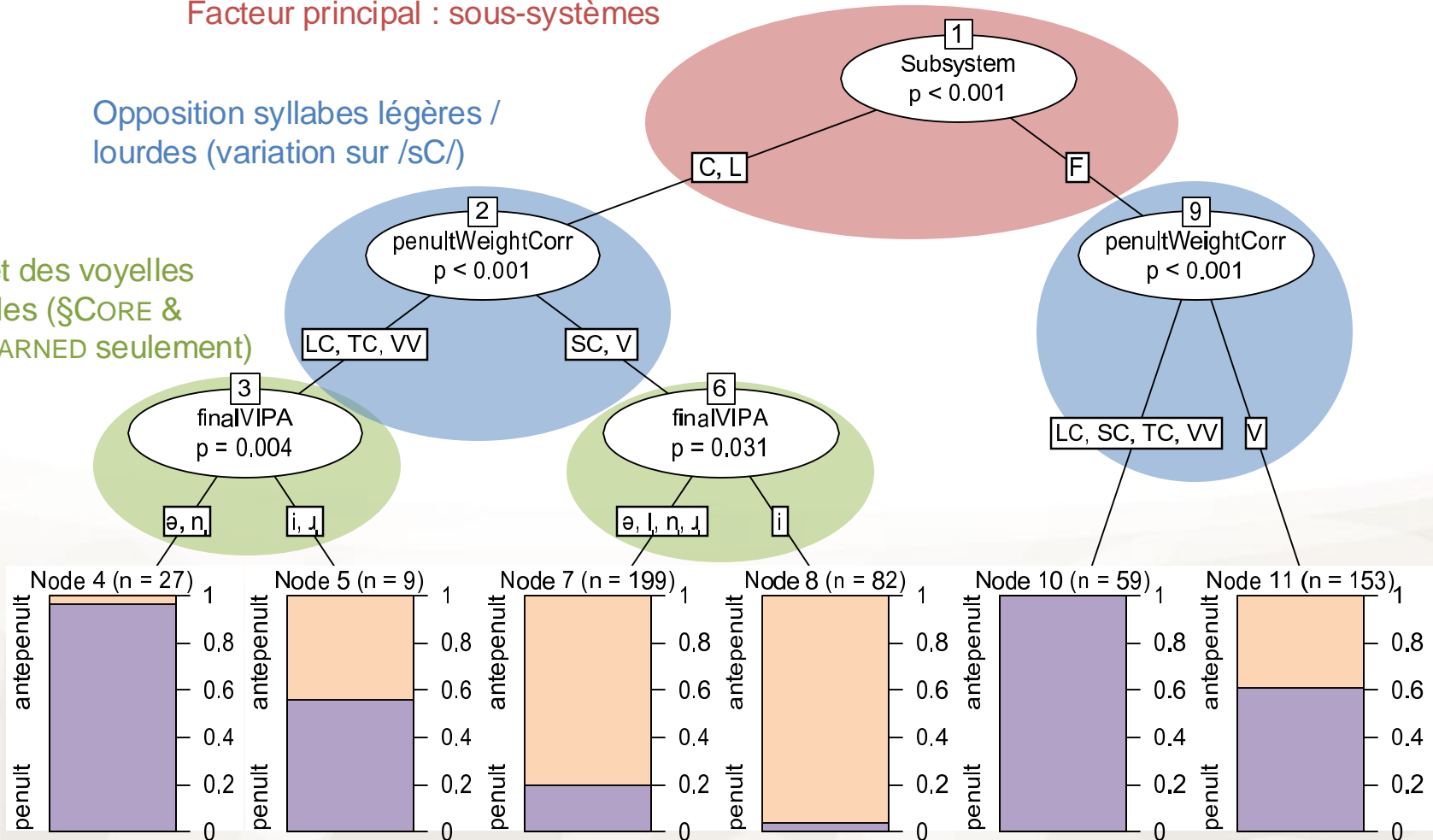
Fonctionne comme un classifieur : chaque noeud indique une division significative dans la distribution des données

Arbre de régression

Facteur principal : sous-systèmes

Opposition syllabes légères /
lourdes (variation sur /sC/)

Effet des voyelles
finales (§CORE &
§LEARNED seulement)

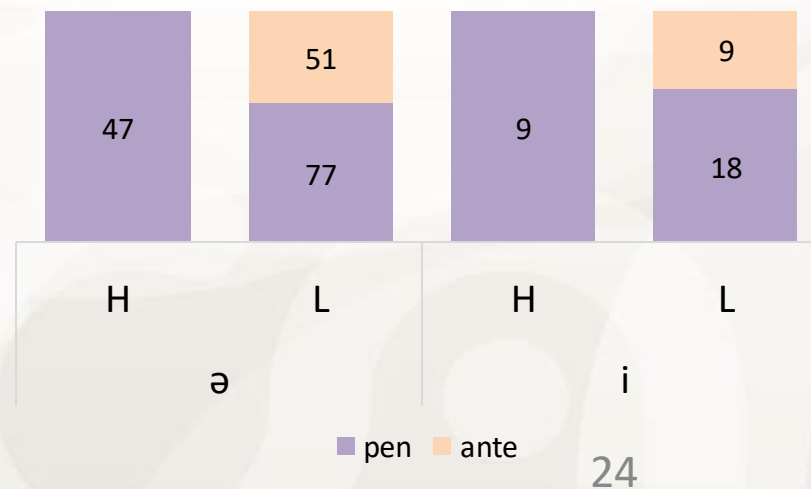


Discussion

Opposition §FOREIGN vs. §CORE & §LEARNED.

Effets du poids syllabique sont observables dans tous les sous-systèmes, mais §FOREIGN a le plus haut taux d'accentuation pénultième.

- Presqu'exclusivement des mots en /ə/ ou /i/, mais pas de différence entre eux :



Discussion

Effets du noyau final visibles dans §CORE et §LEARNED

- Pénultième lourde : /i/ = /ɨ/
- Pénultième légère : /i/ ≠ autres noyaux

Dans les deux cas, /i/ est associé à plus d'accentuation antépénultième

- Les sous-systèmes lexicaux ont un effet plus significatif, mais un petit effet du noyau final subsiste.

Discussion

Il serait intéressant d'explorer la question des sous-systèmes lexicaux directement dans des expériences utilisant des non-mots.

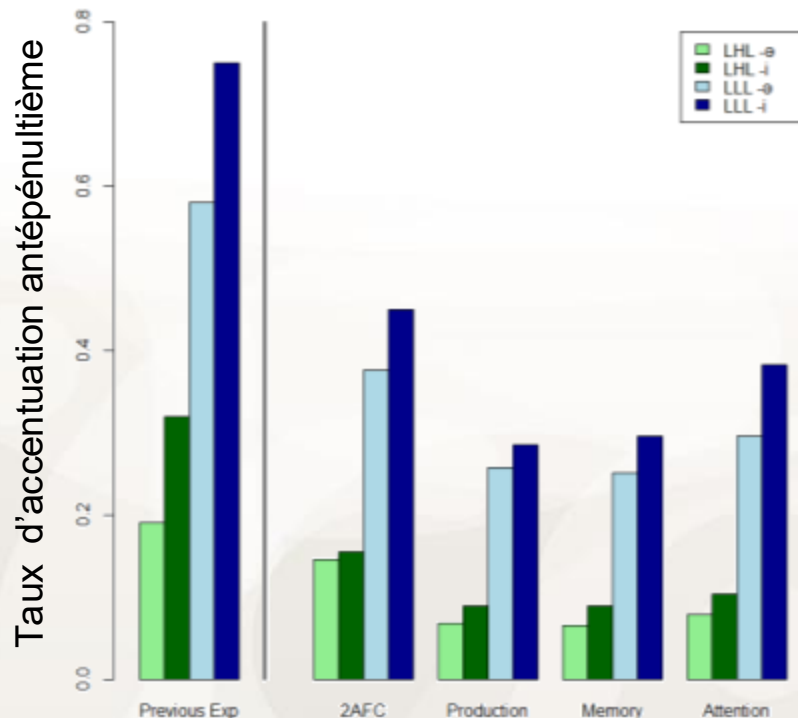
Une idée serait de proposer des définitions qui activeraient les phonologies des différents sous-systèmes :

- e.g. définitions techniques / scientifiques pour §LEARNED
- définitions référant à d'autres cultures (nourriture, monnaie, vêtements) pour §FOREIGN

Discussion

Des premiers éléments en ce sens :

Moore-Cantwell (2023) a répliqué son étude en présentant les non-mots comme étant des noms de villes américaines



Résultats :

Beaucoup plus
d'accentuation
pénultième !

Merci de votre attention !

Projet financé par le DFG et l'ANR dans le cadre du projet ERSaF ("English Root Stress across Frameworks", Clermont-Ferrand & Trier, Novembre 2022 - présent), ANR-21-FRAL-0001-01.

Références

- CARNEY, E. (1994). *A survey of English spelling*. London: Routledge.
- CHOMSKY, N. & HALLE, M. (1968). *The Sound Pattern of English*. New York: Harper & Row.
- DABOUI, Q. & FOURNIER, P. (2022). English PhonologiES? In V. Arigne & C. Rocq-Migette (Eds.), *Modèles et modélisation en linguistique / Models and Modelisation in Linguistics* (pp. 215–258). Bruxelles, Belgique: Peter Lang.
- DOMAHS, U. PLAG, I. & CARROLL, R. (2014). Word Stress Assignment in German, English and Dutch: Quantity-Sensitivity and Extrametricality Revisited. *The Journal of Comparative Germanic Linguistics*, 17(1), 59–96.
- FOURNIER, J.-M. (2007). From a Latin syllable-driven stress system to a Romance versus Germanic morphology-driven dynamics: in honour of Lionel Guierre. *Language Sciences*, 29, 218–236.
- FOURNIER, J.-M. (2010). Accentuation lexicale et poids syllabique en anglais : l'analyse erronée de Chomsky et Halle. *Paper Presented at the 8th Meeting of the French Phonology Network Held at the Université d'Orléans on 1–3 July 2010*.
- LEVSHINA, N. (2021). Conditional Inference Trees and Random Forests. In *A Practical Handbook of Corpus Linguistics*.
- LINDSEY, G. (2019). *English after RP: Standard British Pronunciation Today*. Cham: Palgrave Macmillan.
- MINKOVA, D. (2006). Old and Middle English Prosody. In A. van Kemenade & B. Los (Eds.), *The Handbook of the History of English* (pp. 95–125). Oxford: Blackwell.
- MOORE-CANTWELL, C. (2020). Weight and final vowels in the English stress system. *Phonology*, 37(4), 657–695.
- MOORE-CANTWELL, C. (2023). Rule stability in wug tests. *Manuscript*.
- PATER, J. (1994). Against the underlying specification of an exceptional English stress pattern. *Toronto Working Papers in Linguistics*, (13), 95–121.
- SZIGETVÁRI, P. (2016). No diphthong, no problem. In J. Szpyra-Kozłowska & E. Cyran (Eds.), *Phonology, its faces and interfaces* (pp. 123–141). Budapest: Department of English Linguistics, Eötvös Loránd University.
- TESCHNER, R. V. & WHITLEY, M. S. (2004). *Pronouncing English: a stress-based approach with CD-ROM*. Washington, D.C.: Georgetown University Press.
- WEIDE, R. L. (1994). CMU pronouncing dictionary.

Les caractéristiques des sous-systèmes

		Stress		Graphophonology		Morphology	Semantics	Segmental structure
§Core	§Native	Normal Stress Rule	Neutral suffixes?	Basic correspondences	No Velar Softening	Affixation by juxtaposition Compounding Conversion Some selectional restrictions for affixes	Basic vocabulary	Short words
	§Latinate	Germanic Law Normal Compound Stress	Stress-affecting or mixed suffixes?		Velar Softening	Base truncation in <i>-able</i> derivatives ← <i>-ate</i> bases	Common vocabulary	Polysyllabic words
§French		/-1/ Some stress-affecting suffixes (e.g. <i>-ette, -aire, -eur</i>)	Regular secondary stress rules	<a> - /ɑ:/ <e> - /eɪ/ <i> - /i:/	Specific correspondences e.g. <oi> - /wa:/, <ch> - /ʃ/, <g> - /ʒ/ Silent final consonants	Specific suffixes (e.g. <i>-ette, -aire</i>)	« Prestige »	Distribution of /ʒ/
§Foreign		/-10/ (especially "Italian" words)		<c(c)> - /tʃ/ <sch> - /ʃ/ <z> - /ts/	-	Foreign culture (food, clothing, etc.)	Polysyllabic words Non-silent final vowels ≠ <y> <-VCV#>	
§Learned		/-10/ if closed penult Otherwise /10/ or /-100/? Specific stress rules for neoclassical compounds	Specificities for secondary stress placement in suffixed neoclassical compounds?	<a> - /ɑ:/ (marginal) <ae, oe> - /i:/ <eu> - /ju:/ <ph> - /f/ <ch> - /k/	<ae, oe> - /i:/ <eu> - /ju:/ <ph> - /f/ <ch> - /k/	Specific suffixes (e.g. <i>-itis, -osis, -us, -on</i>) Base truncation more common Neoclassical compounding	Technical, scientific or learned vocabulary	Polysyllabic words CVC Structure for B roots in neoclassical compounds