Derivational Paradigms and The Frequency Factor : The French *-ion* Nouns Allomorphy Problem

Gauvain Schalchli

Université Bordeaux Montaigne/ CLLE (CNRS)

ParadigMo2 workshop

3-4 June 2021

Outline

1. one short introduction
2. three little experiments
3. one brief conclusion

Background - Derivational paradigms and the frequency factor

- classical method of derivational analysis (Jackendoff 1975, Aronoff 1976, Booij 1977, Dell 1979):
  - lexeme-based approach (Aronoff 1976) -> binary families (Van Marle 1985, Becker 1990)
  - 1 series of word pairs with same form/semantic relation -> 1 rule
    - no frequency selection
    - no chronological selection
    - no inflectional selection
- other sources of information
  - productivity studies
    - diachronic productivity (Aronoff & Lindsay 2014, Berg 2020) -> first attestation
    - synchronic productivity (Schultink 1961) -> neologisms
    - corpus productivity (Baayen 1992, Hay 2001) -> token frequency
  - word-based approach (Blevins 2006, Boyé & Schalchli 2019) -> word forms data

Background - A well known school case : the French –*ion* nouns

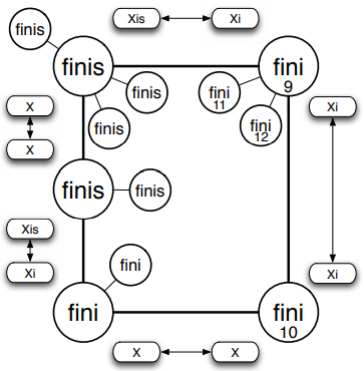
- many works
  - base allomorphy (Di-Lillo 1983)
  - low mean productivity (Dal & al 2008)
  - massive latin inheritance (Lignon & Namer 2010)
  - high productive niche with –*iser* bases (Dal & Namer 2015)
- one famous theory (Bonami, Boyé & Kerleroux 2009)
  - many-indexed stems approach (Aronoff 1994, Brown 1998, Pirelli & Baptista 2000, Stump 2001)
  - stem space model for french (Boyé 2000, 2011; Bonami & Boyé 2003, 2007)
  - lexeme-based approach for derivational morphology (Aronoff 1976, Corbin 1987, Fradin 2003)

Some examples:

- propulser > propulsion = -ion /j§/
- polluter > pollution = -tion /sj§/
- abdiquer > abdication = -ation /asj§/
- composer > composition = -ition /isj§/
- codifier > codification = -cation /kasj§/
- germer > germination = -ination /inasj§/
- abstraire > abstraction = -action /aksj§/

Background - The Hidden Stem Theory

Indice	Formes
1	IMPARFAIT, PRÉSENT 1PL et 2PL
2	PRÉSENT 3PL
3	PRÉSENT SG
4	PARTICIPE PRÉSENT
5	IMPÉRATIF 2SG
6	IMPÉRATIF 1PL et 2PL
7	SUBJONCTIF SG et 3PL
8	SUBJONCTIF 1PL et 2PL
9	INFINITIF
10	FUTUR, CONDITIONNEL
11	PASSÉ SIMPLE, SUBJ IMPARFAIT
12	PARTICIPE PASSÉ



12 theoretical abstract stems explaining verbal allomorphy distribution (Bonami & Boyé 2007)

Classe	Description	Exemple		Effectif
1	**Rad3** ⊕ asjõ	*vexation*	/vɛksasjõ/	1093
2	**Rad3** ⊕ kasjõ	*modification*	/modifikasjõ/	95
3	**Rad3** ⊕ jõ	*dispersion*	/dispɛrsjõ/	86
4	**Rad3** ⊕ isjõ	*composition*	/kõposisjõ/	33
5	**Rad3** ⊕ sjõ	*pollution*	/polysjõ/	50
6	**X** ⊕ jõ	*abstraction*	/abstraksjõ/	277
7	**Pas de base autonome**	*compétition*	/kõpetisjõ/	474

6 surface classes describing –ion polymorphy (Bonami, Boyé & Kerleroux 2009)

- 12 stems (or less!)
- 5 allomorphs
- 1 unpredictible class

(stupid) hypothesis 1:
1 visible stem (n°3)

(clever) hypothesis 2:
1 hidden stem (n°13)

- 13 verbal stems (or more!)
- 1 suffixal form
- 0 unpredictable class

Methodology (counting !)

Experiment 1: **counting the verbal surface stems**

- surface allomorphy in two representative lexicons
- inflection factor

Experiment 2: **counting the stem-allomorphic *–ion* nouns**

- derivational impact of the base inflectional allomorphy
- diachronic factor

Experiment 3: **counting the *–ion* suffix allomorphs**

- ranking data by token-frequency
- frequency factor

Experiment 1: Two databases

Flexique

- made by linguists
- 5200 verbal lexemes
- 265 000 word forms
- automatic paradigm completion + human revision

Lexique3

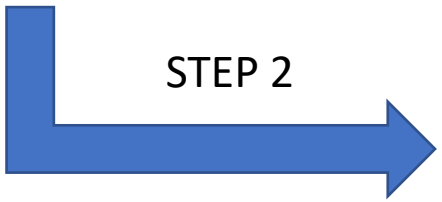
- made by psycholinguists
- 6400 verbal lexemes
- 81 000 word forms
- automatic corpus extracting + human revision

Experiment 1: Evidence of surface stem allomorphy for verbs from Flexique

Method

1. troncating the forms from uniform desinences

2. deleting thematic vowels and phonological alternances



Results

1.

Nb de radicaux	Nb de verbes	Proportion
3	353	6,8%
4	70	1,4%
5	2897	56,0%
6	1760	34,0%
7	70	1,4%
8	22	0,4%

2. 319 (6%) verbs with more than one surface stem

Experiment 1: Evidence of surface stem allomorphy from Lexique3

Word form attestation

The five highest cells			The five lowest cells		
CELL	FREQ	% of 6399 lexemes	CELL	FREQ	% of 6399 lexemes
inf	5294	83%	ind:pas:2p	14	0,2%
par:pas	5139	80%	sub:imp:2s	14	0,2%
ind:pre:3s	4265	67%	sub:imp:1p	12	0,2%
ind:imp:3s	3783	59%	sub:imp:2p	7	0,1%
par:pre	3370	53%	imp:pre:3s	1	0,02%

=> global coverage: 25% (vs 100% for Flexique)

Surface stem allomorphy

nb_rad	freq (nb_lemme)	%
1	5042	**98,1%**
2	45	0,9%
3	33	0,6%
4	16	0,3%
5	0	0,00%
6	2	0,04%
7	0	0,00%
8	1	0,02%
TOTAL	5139	
TOTAL > 1	97	**1,89%**

Experiment 2: Stem allomorphy criteria for *–ion* nouns

- *non-allomorphy criteria:*

1. V stem == N stem
2. N ending == (*-ion* | *-tion* | *-ation* | *-ition* | *-cation*)

Classe	Description	Exemple		Effectif
1	**Rad3** ⊕ asjɔ̃	*vexation*	/vɛksasjɔ̃/	1093
2	**Rad3** ⊕ kasjɔ̃	*modification*	/modifikasjɔ̃/	95
3	**Rad3** ⊕ jɔ̃	*dispersion*	/dispɛrsjɔ̃/	86
4	**Rad3** ⊕ isjɔ̃	*composition*	/kɔ̃posisjɔ̃/	33
5	**Rad3** ⊕ sjɔ̃	*pollution*	/polysjɔ̃/	50
6	**X** ⊕ jɔ̃	*abstraction*	/abstraksjɔ̃/	277
7	**Pas de base autonome**	*compétition*	/kɔ̃petisjɔ̃/	474

BASE	STEM
vexer	vex-
modifier	modifi-
disperser	dispers-
composer	compos-
polluer	pollu-
abstraire	abstr-

Experiment 2: Stem allomorphy counts for *-ion* nouns

- 1739 items (Lexique3)
  - 446 without corresponding verbs
  - 987 without allomorphy
  - 258 early directly borrowed from Latin
  - 48 native formations
    - 44 before 1901 => 4 prefixed (cf. 1a)
    - 41 family stems (30 prefixed) => 4 before XVIe + 1 scholar XVIe + 1 english borrowing + 1 *–faction* (cf. 1b)

1. (eventually) to-be-explained allomorphic data

a. décongélation, déconnection, déconnexion, déconstruction

b. comparution, contravention, électrocution, fustigation, glaciation, obtention, torréfaction

Experiment 3: Frequency ranking of the *–ion* allomorphs and the median threshold hypothesis

ENDINGS	SUFFIX HEADCOUNTS		RATIOS
	high frequency range	low frequency range	
/asj§/ (*-ation*)	297	386	1,2996633
/j§/ (*-ion*)	26	10	0,38461538
/sj§/ (*-tion*)	26	8	0,30769231
/kasj§/ (*-cation*)	18	39	2,16666667
/isj§/ (*-ition*)	8	6	0,75
	types over the threshold	*types under the threshold*	

threshold = median rank of the entire lexicon

Headcounts for the *-isation* subclass

ENDING	HIGH	LOW	Ratio
/izasj§/ (*-isation*)	17	134	**6,4**

Core references:
Baayen (1992)
Gaeta & Ricca (2006)
Fernandez-Dominguez (2010)

Discussion

- results of the experiments
  - experiment 1 -> very few stem allomorphy in the verbal category
  - experiment 2 -> very few stem allomorphy in –ion nouns
  - experiment 3 -> lots of informations about *–ion* suffixal allomorphy in the high frequency part of the attested lexicon
- interpretation of the results
  - experiment 1 -> unpredictible verbal stem allomorphy is exceptional in french
  - experiment 2 -> no productive stem allomorphy in french –ion formation
  - experiment 3 -> morphological knowledge available in the mental lexicon of a french standard speaker is sufficient to explain productive suffixal allomorphy of the *–ion* formation

Conclusion

- case study
  - no need to a hidden stem hypothesis (*Spencer's rasor?)*
  - no productive derivational impact of verbal allomorphy (*sweet mother nature!*)
  - plausibility of a suffixal approach *–ion* allomorphy (*sweet mother Bybee!*)
- theoretical consequences and paradigmatic modeling
  - high interaction between frequency and productivity (Baayen 1992, Hay 2001, Fernandez Dominguez 2010)
  - frequency factor corroborate exponant-based approaches of derivational allomorphy (Lignon & Roché 2011, Montermini 2003, 2018)
  - classical derivational paradigms are heterogeneous data (Blevins et al 2017, Boyé & Schalchli 2019)
- philosophy
  - sometimes stupid hypotheses are more robust than clever ones

To-do list

- completing the analysis
  - finishing diachronic checking *(-ition*, *no base*)
  - measuring relative frequency
  - applying the method to the entire deverbal family
- answering reviewers' remarks
  - -if/-eur correponding deverbal nouns
    - exécuter → exécution/éxécuteur/exécutif
    - many triplet without verbal base
    - no taking account of frequency, neither diachrony

Thank you for your attention!

References

- Aronoff, M. (1976). Word formation in generative grammar. Linguistic Inquiry Monographs Cambridge, Mass, (1), 1-134.
- Bonami, Olivier, et Gilles Boyé. 2003. « Supplétion et classes flexionnelles ». *Langages* 37(152): 102-26.
- ———. 2007. « Remarques sur les bases de la conjugaison ». In *Des sons et des sens (données et modèles en phonologie et en morphologie)*, Langues et syntaxe, éd. Elisabeth Delais Roussarie et Laurence Labrune. Hermès Sciences, 77-90.
- Bonami, Olivier, Gilles Boyé, et Françoise Kerleroux. 2009. « L'allomorphie radicale et la relation flexion-construction ». In *Aperçus de morphologie*, éd. Fradin, Bernard; Kerleroux, Françoise; Plénat, et Marc; Presses de l'Université de Vincennes, 103-25.
- Boyé, G., & Schalchli, G. (2019). Realistic data and paradigms: The paradigm cell finding problem. Morphology, 29(2), 199–248.
- Montermini, Fabio. 2018. Les affixes dérivationnels ont-ils des allomorphes ? Pour une modélisation de la variation des exposants dans une morphologie à contraintes. In The lexeme in descriptive and theoretical morphology. Ed. Bonami et al.
- Stump, Gregory T. 2001. *Inflectional Morphology: A Theory of Paradigm Structure.* Cambridge:Cambridge University Press.
- Tribout, Delphine. 2010. How Many Conversions from Verb to Noun Are There in French? In CSLI Publications.
- ———. 2012. Verbal Stem Space and Verb to Noun Conversion in French. *Word Structure* 5(1): 109-28.