

De formes en thèmes

Olivier Bonami
 Université Paris-Sorbonne
 Institut Universitaire de France
 Laboratoire de Linguistique Formelle

Gilles Boyé
 Université Bordeaux 3
 CLLE-ERSS

Depuis Bonami & Boyé (2002, 2003), nous avons défendu une analyse de la morphologie du français basée sur la notion d'ESPACE THEMATIQUE : l'espace thématique d'un lexème est une famille indexée de radicaux, qui prend la place du radical unique supposé par la plupart des analyses, au moins pour les lexèmes réguliers. Initialement proposée pour rendre compte du rôle structurant de l'allomorphie radicale dans la conjugaison, l'hypothèse des espaces thématiques s'est avérée utile pour rendre compte de nombreux phénomènes d'allomorphie, en particulier à la frontière entre flexion et morphologie constructionnelle (Bonami, Boyé & Kerleroux (2009) ; Bonami & Boyé, 2005 ; Kerleroux, 2007 ; Boyé & Plénat, sous presse ; Plénat, 2008 ; Roché, 2010 ; Tribout, 2010, sous presse). L'ensemble de ces travaux présupposent, sans la remettre en question, la proposition initiale de Bonami & Boyé (2003) sur la structure de l'espace thématique des verbes du français.

Dans ce travail nous revenons sur cette proposition initiale à la lumière de deux stratégies de recherche en morphologie qui n'étaient pas disponibles, ou du moins pas utilisées par nous au moment de nos premiers travaux.

La première est l'utilisation de méthodes quantitatives instrumentées pour guider l'analyse de grands ensembles de données. De fait, depuis Boyé (2000) nous travaillons sur des données semi-exhaustives, mais l'analyse de ces données a longtemps été faite à la main et de manière impressionniste. Dans la deuxième section de ce chapitre, nous utilisons une variante de la méthodologie définie par Ackerman *et al.* (2009) pour quantifier les relations d'interprédictibilité entre cases du paradigme, et montrons comment l'application d'une telle méthodologie conduit à l'observation d'une partition du paradigme verbal du français proche mais subtilement différente de celle de Bonami & Boyé (2003).

La deuxième est la distinction, due à Blevins (2006), entre approches CONSTRUCTIVES et approches ABSTRACTIVES de la morphologie. Dans une approche constructive, la flexion est modélisée comme une procédure permettant de construire les formes fléchies d'un lexème à partir d'une représentation abstraite de ce lexème, typiquement un ou plusieurs radicaux assorti(s) d'une classe flexionnelle. Dans une approche abstractive, la flexion est modélisée comme une procédure permettant de déduire une forme d'un lexème à partir d'une autre de ces formes. Radicaux et affixes émergent alors comme des abstractions observables à partir des relations de mot-forme à mot-forme, mais l'analyse des relations entre mots-formes ne présuppose pas la possibilité d'une segmentation en radical et affixe.

Dans la troisième section de cet article, nous proposons une reconceptualisation des espaces thématiques sous la forme d'une analyse abstractive des verbes du français. Nous montrons comment cette approche abstractive permet sur une base conceptuelle bien définie de converger précisément sur la partition du paradigme des verbes observée grâce à l'analyse quantitative. Le résultat est cependant une analyse qui n'est pas directement exploitable en termes constructifs.

1 Des forts en thèmes

1.1 La renaissance de l'allomorphie radicale

Au cours des vingt dernières années, la réflexion sur l'allomorphie radicale a fondamentalement modifié la conception de la morphophonologie, sur la base des travaux fondateurs d'Aronoff (1994 : chap. 2) et Maiden (1992). Le travail d'Aronoff sur la conjugaison du latin, reprenant des observations remontant à Priscien et déjà discutées par Matthews (1972), a donné aux alternances radicales du latin le statut de parangon d'un phénomène purement morphologique, ou MORPHOMIQUE : la distribution des trois radicaux latins ne peut s'expliquer par le fait qu'ils encodent chacun un ensemble cohérent de propriétés morphosyntaxiques. Le travail de Maiden montrait simultanément que l'existence de patrons d'alternance radicale à distribution morphomique était une propriété stable des langues romanes, ayant une influence notable sur l'évolution de leur lexique—en conséquence, l'existence de ces patrons ne peut pas être considérée comme un pur accident synchronique, mais s'impose comme un élément systémique de la grammaire de certaines langues. L'intérêt subséquent pour les patrons d'alternance radicale a amené un certain nombre de chercheurs, parmi lesquels Brown (1998), Pirelli & Battista (2000), Stump (2001 : chap. 6), Bonami & Boyé (2003) ou Blevins (2003), à donner une substance empirique et formelle à une nouvelle forme d'analyse constructive des systèmes flexionnels, qui se caractérise par l'ensemble d'hypothèses en (1).

- (1) a. Les lexèmes ne sont pas nécessairement associés à un radical unique ou privilégié, mais à une collection indexée de radicaux¹. Dans la tradition française, à la suite de Bonami & Boyé (2003), cette collection a été appelée ESPACE THÉMATIQUE².
- b. La partition du paradigme induite par l'espace thématique d'un lexème ne recouvre pas nécessairement des zones morphosyntaxiquement cohérentes du paradigme.
- c. La même partition est mise en jeu pour rendre compte aussi bien d'alternances régulières concernant des classes ouvertes de lexèmes que de supplétions du radical concernant des lexèmes isolés (Boyé, 2000).
- d. Les règles régissant les relations entre les cases de l'espace thématique (REGLES DE FORMATION DE RADICAL au sens de Stump 2001) sont indépendantes des règles régissant le

¹ La notion d'ENSEMBLE DE RADICAUX d'un lexème d'Anderson (1992) est un précurseur évident de cette vision, à ceci près qu'Anderson ne concevait pas les ensembles de radicaux comme entretenant des relations purement morphomiques.

² On utilise ici THÉMATIQUE comme l'adjectif dérivé sur RADICAL, faute d'un candidat plus euphonique.

choix d'une case pour former une certaine forme du lexème (REGLES DE CHOIX DU RADICAL au sens de Stump 2001).

1.2 Les espaces thématiques à la française

Dans ce paragraphe nous précisons l'ensemble d'hypothèses qui, ajoutées à (1), sous-tendent nos travaux sur les espaces thématiques (Boyé 2000, 2011 ; Bonami & Boyé 2002, 2003, 2005, 2006 ; Bonami, Boyé & Kerleroux 2009 ; Boyé & Cabredo, 2005, 2006 ; Montermini & Boyé, sous presse), travaux dont nous désignerons collectivement les auteurs comme Boyé *et al.*

L'approche des espaces thématiques de Boyé *et al.* se caractérise par l'ensemble d'hypothèses suivant :

- (2) a. La flexion à proprement parler est invariable : en dehors de cas rares de supplétion d'une forme fléchie, tous les lexèmes d'une même catégorie portent les mêmes marques de flexion. En conséquence, toute variation apparente des marques de flexion est analysée comme une allomorphie radicale.
- b. La taille des espaces thématiques est fixe : tous les lexèmes appartenant à une même catégorie syntaxique ont un espace thématique de même structure.
- c. Les seules règles de formation de radical pertinentes sont des RELATIONS IMPLICATIVES PAR DEFAUT. Ces relations lient deux cases *i* et *j* de l'espace thématique et indiquent, sur la base des propriétés phonologiques du radical, quelle fonction morphologique lie le contenu de la case *i* au contenu de la case *j* pour les lexèmes d'une même classe flexionnelle.

Cet ensemble d'hypothèses a pour but de contraindre fortement les analyses possibles pour un système flexionnel donné³, et produit par conséquent des analyses qui vont souvent à l'encontre de la tradition. Prenons quelques exemples issus de la conjugaison du français. En vertu de (2a), il ne peut y avoir d'exposant affixal du participe passé en français. Les affixes apparents *-é* /e/ (cf. *mangé*), *-i* /i/ (cf. *sorti*), *-u* /y/ (cf. *pendu*), voire *-ert* /ɛʁ/ (cf. *couvert*) sont nécessairement des portions de radical, puisqu'ils ne sont pas constants. En vertu de (2b), on est amené à dire que tous les verbes du français, si réguliers soient-ils, ont un espace thématique à au moins quatre cases, puisque le verbe ALLER a au moins quatre radicaux distincts : /v(a)/ (cf. *tu vas*), /al/ (cf. *allons*), /i/ (cf. *tu iras*), /aj/ (cf. *que tu ailles*). En vertu de (2c), la relation entre les radicaux /bwav/ et /bwa/ de BOIRE est tout aussi supplétive que celle entre les radicaux /va/ et /al/ de ALLER, parce que dans les deux cas, il n'y a pas de classe ouverte de lexèmes qui soit caractérisée par l'alternance en question.

Des résultats plus surprenants encore surgissent si on applique les mêmes critères à un système de déclinaison typique, comme celui du tchèque, partiellement illustré dans le tableau 1. (2a) a

³ Les travaux de Boyé *et al.* se distinguent sur ce point de ceux qui ont été menés dans le cadre de *Network Morphology* (voir en particulier Brown 1998, Hippiisley 1998, Bærman *et al.* 2005, Brown & Hippiisley 2012), et qui ne se donnent pas d'heuristique déterministe pour distinguer l'allomorphie radicale de l'exponence.

pour conséquence que les seuls exposants affixaux intervenant dans la déclinaison des noms masculins en tchèque sont les suffixes de génitif pluriel *-ů* et de datif pluriel *-ům*, parce que ce sont les seuls à être constants dans toutes les classes flexionnelles.

		POKOJ	MOST	PAN	MUZ	PREDESDA	SOUDECE
		‘pièce’	‘pont’	‘monsieur’	‘homme’	‘président’	‘juge’
SG	NOM	<i>pokoj</i>	<i>most</i>	<i>pán</i>	<i>muž</i>	<i>předseda</i>	<i>soudce</i>
	GEN	<i>pokoje</i>	<i>mostu</i>	<i>pána</i>	<i>muže</i>	<i>předsedy</i>	<i>soudce</i>
	DAT	<i>pokoji</i>	<i>mostu</i>	<i>pánovi</i>	<i>muži</i>	<i>předsedovi</i>	<i>soudci</i>
	ACC	<i>pokoj</i>	<i>most</i>	<i>pána</i>	<i>muže</i>	<i>předsedu</i>	<i>soudce</i>
PL	NOM	<i>pokoje</i>	<i>mosty</i>	<i>pánové</i>	<i>muži</i>	<i>předsedové</i>	<i>soudci</i>
	GEN	<i>pokojů</i>	<i>mostů</i>	<i>pánů</i>	<i>mužů</i>	<i>předsedů</i>	<i>soudců</i>
	DAT	<i>pokojům</i>	<i>mostům</i>	<i>pánům</i>	<i>mužům</i>	<i>předsedům</i>	<i>soudcům</i>
	ACC	<i>pokoje</i>	<i>mosty</i>	<i>pány</i>	<i>muže</i>	<i>předsedy</i>	<i>soudce</i>

Tableau 1. Paradigmes partiels des principales classes de noms masculins tchèques

La nature exacte des hypothèses en (2), et en particulier leurs conséquences pour l'analyse de données telles que celles de la déclinaison du tchèque, a souvent été mal comprise — ce qui est dû pour une large part à un emploi non contrôlé par Boyé *et al.* du vocabulaire descriptif traditionnel avec un sens qui n'est pas celui de la tradition. (2) encode une hypothèse particulière sur la relation entre morphologie implicative, règles d'exponence (Stump 2001) et classes flexionnelles : selon (2), tout ce qui dans la flexion n'est pas général à l'ensemble des lexèmes relève de la morphologie implicative ; tout ce qui est général relève de l'exponence. Le fait d'appeler « portion de radical » les morceaux récurrents (*recurrent partials*) gérés par la morphologie implicative, et « marques de flexion » ceux qui sont gérés par les règles d'exponence est une facilité de vocabulaire qui se trouve coïncider à peu près avec l'emploi usuel de ces termes quand on parle de la conjugaison des langues romanes. Le fait que la même coïncidence ne se rencontre pas dans le cas de la déclinaison des langues slaves ne disqualifie pas en tant que tel l'hypothèse que ces systèmes de déclinaison sont essentiellement implicatifs, ce qui dans le vocabulaire de Boyé et ses collègues se décrit en termes d'alternance radicale. De la même manière, la séparation entre flexion constante et flexion variable que présuppose (2) ne suppose en rien que la flexion constante soit concaténative et la flexion variable ne le soit pas : une marque de flexion peut être une opération non concaténative, une alternance radicale peut consister en l'ajout d'un augment marginal (contra Cameron & Carstairs-McCarthy 2000).

Dans la suite de cet article nous nous concentrons sur la conjugaison du français, si bien que ces flottements de vocabulaire devraient être sans conséquences majeures.

Dans Bonami & Boyé (2003), les hypothèses en (2), qui ont vocation à être de portée générale, sont assorties des deux hypothèses complémentaires en (3) spécifiques à la conjugaison du français :

- (3) a. L'association entre cases de l'espace thématique et cases du paradigme est invariable : si un lexème L utilise son radical i pour remplir la case σ de son paradigme, alors tous les lexèmes utilisent leur radical i pour cette case.
- b. Il n'y a qu'une seule classe flexionnelle de lexèmes réguliers. Donc, en vertu de (2c), les seules relations implicatives par défaut qui doivent être prises en compte dans la grammaire sont celles qui valent pour les lexèmes réguliers ; tout lexème irrégulier l'est en vertu de la violation d'une relation implicative par défaut.

Si on peut débattre du fait que les hypothèses (2) soient défendables pour la description de tous les systèmes flexionnels, ce n'est certainement pas le cas pour celles de (3), qui sont des hypothèses sur un système particulier⁴. Elles visent à modéliser des observations impressionnistes sur les propriétés macroscopiques de la conjugaison du français, et guident la forme de l'analyse.

1.3 La version officielle de l'espace thématique des verbes du français

En fonction de ces hypothèses, Bonami & Boyé (2003) produit une analyse constructive des verbes du français basée sur un espace thématique à 12 cases⁵, qui est synthétisée dans le tableau 2, qui indique quelle case du paradigme est basée sur quel radical. Le tableau 3 précise les marques de flexion qui s'appliquent au radical donné par le tableau 2. On notera que ces marques s'éloignent de manière caractéristique des inventaires traditionnels sous l'effet de (4a) : au passé simple, à l'infinitif, et au participe passé, une sous-séquence habituellement traitée comme marque de flexion est intégrée au radical⁶.

⁴ Bonami & Boyé (2005, 2006) proposent une analyse des adjectifs du français en termes d'espaces thématiques qui ne respecte ni (3a) ni (3b): elle utilise des classes flexionnelles substantielles, et comporte une règle de sélection du radical qui explore deux cases de l'espace thématique. Par ailleurs, Tribout (2010) montre, sur la base des propriétés morphologiques des verbes désadjectivaux, que (3b) ne peut être maintenue dans le cas général.

⁵ Kerleroux (2007) et Bonami, Boyé et Kerleroux (2009) proposent d'augmenter l'espace thématique d'un radical caché, numéroté 13, correspondant au radical du supin latin et qui a pour caractéristique de servir d'entrée à des règles de morphologie constructionnelle, mais jamais à des règles de flexion. Tribout (2010) à son tour a motivé l'ajout d'un radical supplémentaire, numéroté 0 et typiquement identique au radical 3 ou plus petit que celui-ci, et qui entre en jeu dans les opérations de conversion impliquant des verbes.

⁶ Deux autres hypothèses inhabituelles sont (i) la postulation d'un suffixe infinitif /v/ même pour les verbes du premier groupe (on suppose que ce suffixe est supprimé par une règle phonologique $v \rightarrow \emptyset / e_ \#$ sauf dans les contextes de liaison) ; (ii) la postulation d'une opération non concaténative H au passé simple 1sg et 3pl qui permet de gérer le changement de voyelle pour les verbes du premier groupe, (iii) la troncation des consonnes finales du radical du participe passé au masculin.

	1SG	2SG	3SG	1PL	2PL	3PL
PRESENT	3					2
IMPARFAIT						
IMPERATIF		5		6		
SUBJONCTIF	7			8		
FUTUR	10					
CONDITIONNEL	10					
PASSE SIMPLE	11					
SUBJ. PASSE	11					

INFINITIF	PART. PRESENT	M.SG	M.PL	F.SG	F.PL
9	4	PART. PASSE 12			

TABLEAU 2 — Partition du paradigme des verbes du français selon Bonami & Boyé (2003)

	1SG	2SG	3SG	1PL	2PL	3PL
PRESENT	id.	id.	id.	⊕ \tilde{o}	⊕e	id.
IMPERATIF		id.		⊕ \tilde{o}	⊕e	
IMPARFAIT	⊕ ϵ	⊕ ϵ	⊕ ϵ	⊕j \tilde{o}	⊕je	⊕ ϵ
SUBJONCTIF	id.	id.	id.	⊕j \tilde{o}	⊕je	id.
FUTUR	⊕ $\mathcal{B}\epsilon$	⊕ $\mathcal{B}a$	⊕ $\mathcal{B}a$	⊕ $\mathcal{B}\tilde{o}$	⊕ $\mathcal{B}e$	⊕ $\mathcal{B}\tilde{o}$
CONDITIONNEL	⊕ $\mathcal{B}\epsilon$	⊕ $\mathcal{B}\epsilon$	⊕ $\mathcal{B}\epsilon$	⊕ $\mathcal{B}j\tilde{o}$	⊕ $\mathcal{B}je$	⊕ $\mathcal{B}\epsilon$
PASSE SIMPLE	H	id.	id.	⊕m	⊕t	Ho(⊕ \mathcal{B})
SUBJ. PASSE	⊕s	⊕s	id.	⊕sj \tilde{o}	⊕sje	⊕s

INFINITIF	PART. PRESENT	M.SG	M.PL	F.SG	F.PL
⊕ \mathcal{B}	⊕ \tilde{a}	PART. PASSE tronc_C	tronc_C	id.	id.

Tableau 3 — Marques de flexion pour les verbes du français selon Bonami & Boyé (2003) ⁷

Ces tableaux appellent deux clarifications. D'une part, Bonami & Boyé (2003) suppose que quelques lexèmes maximale­ment irréguliers listent des FORMES FLECHIES SUPPLETIVES qui ne se segmentent pas en radical et marque de flexion. La liste des formes en question est donnée dans le tableau 4. Cette hypothèse permet de limiter la prolifération des cases de l'espace thématique, qui étaient au nombre de 16 dans Boyé (2000). Elle est en outre psychologiquement plausible, dans la mesure où il est établi que les formes fléchies les plus fréquentes sont mémorisées par les locuteurs, même quand elles sont formées régulièrement (voir p. ex. Baayen *et al.* 1997). Dans le

⁷ « H » est une fonction qui transforme un /a/ final de radical en / ϵ /, et ne change rien si le dernier segment n'est pas /a/. « Ho(⊕ \mathcal{B}) » est la composition de cette fonction avec la concaténation de / \mathcal{B} /. « tronc_C » est une fonction qui tronque la consonne finale du radical s'il y en a une, et ne change rien sinon. Les participes passés ne donnant pas lieu à liaison, cette analyse par troncation morphologique rend compte simplement de la distribution des consonnes finales dans les participes.

cas de lexèmes aussi fréquents que ceux-ci, mémoriser une forme supplétive ne constitue donc pas un coût supplémentaire pour le locuteur.

LEXEME	CASE	FORME ATTENDUE	FORME SUPPLEEE
ETRE	PRS.1.SG	<i>es</i>	<i>suis</i>
	PRS.1.PL	<i>étons</i>	<i>sommes</i>
	PRS.2.PL	<i>étez</i>	<i>êtes</i>
AVOIR	PRS.1.SG	<i>as</i>	<i>ai</i>
ALLER	PRS.1.SG	<i>vas</i>	<i>vais</i>
DIRE	PRS.2.PL	<i>disez</i>	<i>dites</i>
	IMP.2.PL	<i>disez</i>	<i>dites</i>
FAIRE	PRS.2.PL	<i>faites</i>	<i>faites</i>
	IMP.2.PL	<i>faites</i>	<i>faites</i>

TABLEAU 4 — LES FORMES FLECHIES SUPPLETIVES

D'autre part, l'hypothèse de Bonami & Boyé (2003) est que l'espace thématique est structuré par des relations implicatives par défaut ayant pour conséquence que, dans le cas régulier⁸, les radicaux numérotés 1 à 8 sont identiques, les radicaux 9 à 12 sont déductibles par des règles implicatives sans exceptions, listées sous une forme simplifiée en (4). La multiplication des cases de l'espace thématique ne conduit donc pas à multiplier mécaniquement le nombre de stipulations qui doivent être encodées dans le lexique : seuls les lexèmes qui utilisent des radicaux ne se conformant pas à (4) nécessitent de lister des radicaux supplémentaires.

- (4) a. Les radicaux 1 à 8 sont identiques.
 b. Le radical 9 est obtenu en ajoutant /e/ à la fin du radical 1.
 c. Le radical 10 est obtenu en ajoutant /ə/ à la fin du radical 3.
 d. Le radical 11 est obtenu en ajoutant /a/ à la fin du radical 1.
 e. Le radical 12 est obtenu en ajoutant /e/ au radical 1.

2 Déformants thèmes

L'analyse de Bonami & Boyé (2003) est fondée sur une exploration supposée semi-exhaustive des patrons de conjugaison du français : elle rend compte de l'ensemble des patrons de flexion recensés par les guides de conjugaison, tout en prenant en compte des données quantitatives sur l'effectif de chaque patron, compilées à partir de (Arrivé 1997). La méthodologie employée n'est donc pas sans manquer de rigueur : la manipulation de grands ensembles de données (88 tables de 51 formes

⁸ Pour les besoins de ce paragraphe, on suppose, comme Bonami & Boyé (2003) ou Kilani-Schoch & Dressler (2005), que seuls les verbes du premier groupe sont réguliers en français. Bonami *et al.* (2008) montre que cette hypothèse est erronée, et Bonami & Boyé (2007), Boyé (2011) construisent une analyse des relations implicatives au sein de l'espace thématique compatible avec ce résultat.

dans Arrivé 1997) interdit de s'assurer, sans instrumentation, que l'on a produit l'analyse la plus satisfaisante possible étant donné les hypothèses de départ.

Dans cette section nous pallions enfin cette limite du travail de Bonami & Boyé (2003) en utilisant une méthode d'analyse quantitative des patrons d'alternance entre les cases d'un paradigme basée sur la théorie de l'information. La méthode est un descendant direct de celle d'Ackerman *et al.* (2009), et a été élaborée pour les besoins de Bonami *et al.* (2011) ; nous renvoyons à ce texte pour une description et une motivation détaillée, et nous contentons ici d'une description rapide. Nous présentons ensuite les résultats de cette méthode appliquée aux paradigmes des 6440 verbes non défectifs présents dans la base de données lexicales BDLEX (Calmès & Pérennou 1998), puis nous analysons les contrastes entre les résultats de cette étude quantitative et ce qui est attendu sur la base de l'analyse de Bonami & Boyé (2003).

2.1 La méthode

L'objectif est de quantifier le niveau de certitude avec lequel on peut prédire la forme occupant une case du paradigme à partir de la forme occupant une autre case, étant donné une connaissance statistique de la distribution des différents patrons d'alternance disponibles dans le lexique existant.

2.1.1 Classification des paires alternantes

Pour ce faire, il est d'abord nécessaire de disposer d'une classification des lexèmes en fonction de la relation morphologique qui lie les deux cases en question de leurs paradigmes. Nous appellerons PAIRES ALTERNANTES ENTRE LA CASE σ ET LA CASE τ les paires de formes $\langle \varphi, \psi \rangle$ telles qu'un certain lexème utilise φ pour sa case σ et ψ pour sa case τ . Notre but est de classer les paires $\langle \varphi, \psi \rangle$ et de déterminer les effectifs de chaque classe. Nous utilisons à cette fin une série d'heuristiques inspirées de celle du *Minimal Generalization Learner* (Albright, 2002). Nous identifions d'abord pour chaque lexème L la règle phonologique au format SPE (Chomsky & Halle, 1968) la plus spécifique qu'il instancie (5). (6) donne des exemples de règles élémentaires reliant le présent singulier au participe passé.

(5) Etant donnée une paire alternante $\langle \varphi, \psi \rangle$, on identifie la séquence la plus longue X et la séquence la plus courte α telles qu'il existe une séquence Y et une séquence β telles que $\varphi = X + \alpha + Y$ et $\psi = X + \beta + Y$. La règle élémentaire instanciée par $\langle \varphi, \psi \rangle$ est la règle $\alpha \rightarrow \beta / \#X_Y\#$.

- (6) a. $\emptyset \rightarrow ty / \#ba_ \#$ (BATTRE)
 b. $\emptyset \rightarrow ty / \#v\epsilon_ \#$ (VÊTIR)
 c. $\emptyset \rightarrow ty / \#deba_ \#$ (DÉBATTRE)
 d. $\emptyset \rightarrow ty / \#\text{v}\partial\text{v}\epsilon_ \#$ (REJETIR)

On procède ensuite, pour chaque changement $\alpha \rightarrow \beta$, à la généralisation minimale au sens d'Albright, pour aboutir à la règle la plus spécifique qui rende compte de l'ensemble des paires $\langle \varphi, \psi \rangle$ observées. Pour prendre un exemple simple, quand on cherche à relier le présent singulier au

participe passé, on tombe sur 11 verbes instanciant le changement $\emptyset \rightarrow ty$, tous dérivés de BATTRE ou de VETIR. La généralisation minimale conduit à produire la règle en (7) à partir des règles élémentaires en (6).

$$(7) \quad \emptyset \rightarrow ty \quad \left[\begin{array}{l} +cons \\ -voc \\ +voisé \\ +labial \\ -cont \\ -son \end{array} \right] \left[\begin{array}{l} -cons \\ +voc \\ -haut \\ -arrière \\ -nasal \end{array} \right] \quad \#$$

Le résultat de cette première étape est que l'on a abouti à un nombre minimal de règles tels que (i) chaque paire $\langle \varphi, \psi \rangle$ observée instancie une et une seule règle, et (ii) chaque règle impose la restriction maximale (exprimable en termes de décomposition en traits) sur l'aspect phonologique des formes d'entrées auxquelles elle peut s'appliquer.

2.1.2 Classification des formes d'entrée

On procède alors à une classification des formes d'entrée en fonction des règles identifiées. Une règle est *APPLICABLE* à une paire $\langle \varphi, \psi \rangle$ si l'aspect phonologique de φ est tel que la règle pourrait lui être appliquée ; une règle est *APPLIQUEE* à la paire $\langle \varphi, \psi \rangle$ si en outre l'application de cette règle à φ donne effectivement ψ . Deux formes d'entrée φ et φ' appartiennent à la même classe si l'ensemble des règles applicables à φ est identique à l'ensemble des règles applicables à φ' . Cette méthode de classification aboutit à placer les formes d'entrées dans la même classe si elles partagent certaines propriétés phonologiques, mais sans faire de choix *a priori* quant aux propriétés phonologiques pertinentes : appartiennent à la même classe deux formes qui sont traitées de manière équivalente par les règles identifiées.

2.1.3 Evaluation de l'entropie conditionnelle sur la base des deux classifications

La dernière étape consiste à quantifier l'incertitude quant à la forme de sortie quand on connaît la forme d'entrée. L'idée est la suivante. L'effectif des classes de formes d'entrée nous donne une idée de la probabilité d'appartenir à chaque classe, alors que l'effectif des classes de paires alternantes nous donne une idée de la probabilité d'exemplifier chaque patron d'alternance. Par exemple, arrêtons-nous sur la relation entre l'imparfait 1.SG et l'imparfait 1.PL. La procédure indiquée en 2.1.1 aboutit à l'identification de 3 règles dans chaque direction, exemplifiées dans les tableaux 5 et 6, qui donnent également le nombre de verbes vérifiant chacune de ces règles. Sur la base de cette classification des paires alternantes on peut déduire la classification des formes d'entrée et l'effectif de chaque classe, comme l'indiquent les tableaux 7 et 8.

REGLE	EX. DE PAIRE	EX. DE LEXEME	EFFECTIF	PROBABILITE
I. $\varepsilon \rightarrow j\tilde{\omega} / _ \#$	$\langle kaze, kazj\tilde{\omega} \rangle$	CASER	5429	0,84
II. $\varepsilon \rightarrow \tilde{\omega} / \begin{bmatrix} +cons \\ +haut \\ -cor \end{bmatrix} _ \#$	$\langle kad\tilde{v}ije, kad\tilde{v}ij\tilde{\omega} \rangle$	QUADRILLER	722	0,11
III. $\varepsilon \rightarrow ij\tilde{\omega} / \begin{bmatrix} +cons \\ -voc \\ -son \end{bmatrix} \begin{bmatrix} +cons \\ -voc \\ -nasal \end{bmatrix} _ \#$	$\langle kad\tilde{v}\varepsilon, kad\tilde{v}ij\tilde{\omega} \rangle$	CADRER	289	0,05

Tableau 5 — Classification des paires $\langle \text{IPF.1SG}, \text{IPF.1PL} \rangle$

REGLE	EX. DE PAIRE	EX. DE LEXEME	EFFECTIF	PROBABILITE
I. $j\tilde{\omega} \rightarrow \varepsilon / _ \#$	$\langle kazj\tilde{\omega}, kaze \rangle$	CASER	5429	0,84
II. $\tilde{\omega} \rightarrow \varepsilon / \begin{bmatrix} +cons \\ +haut \\ -cor \end{bmatrix} _ \#$	$\langle kad\tilde{v}ij\tilde{\omega}, kad\tilde{v}ije \rangle$	QUADRILLER	722	0,11
III. $ij\tilde{\omega} \rightarrow \varepsilon / \begin{bmatrix} +cons \\ -voc \\ -son \end{bmatrix} \begin{bmatrix} +cons \\ -voc \\ -nasal \end{bmatrix} _ \#$	$\langle kad\tilde{v}ij\tilde{\omega}, kad\tilde{v}\varepsilon \rangle$	CADRER	289	0,05

Tableau 6 — Classification des paires $\langle \text{IPF.1PL}, \text{IPF.1SG} \rangle$

CLASSE	PATRONS	EX. DE FORME	EX. DE LEXEME	EFFECTIF	PROBABILITE
A	{I}	kaze	CASER	5429	0,84
B	{I,II}	kad \tilde{v} ije	QUADRILLER	722	0,11
C	{I,III}	kad \tilde{v} ε	CADRER	289	0,05

Tableau 7 — Classification de l'IPF.1SG en fonction de sa relation à l'IPF.1PL

CLASSE	PATRONS	EX. DE FORME	EX. DE LEXEME	EFFECTIF	PROBABILITE
A	{I,II}	kazj $\tilde{\omega}$, \tilde{v} asazj $\tilde{\omega}$	CASER, RASSASIER	5893	0,91
B	{I,II,III}	kad \tilde{v} ij $\tilde{\omega}$	QUADRILLER, CADRER	319	0,05
C	{II}	elwaj $\tilde{\omega}$	ELOIGNER	228	0,04

Tableau 8 — Classification de l'IPF.1PL en fonction de sa relation à l'IPF.1SG

La probabilité conditionnelle d'une classe de paires alternantes connaissant une classe de formes d'entrée indique, pour une classe donnée de formes d'entrée, quelle chance on a qu'une forme de cette classe instancie une règle particulière. Dans nos exemples, celle-ci est facilement évaluée à partir d'un examen croisé des deux classifications. Par exemple, parmi les 319 verbes dont l'imparfait 1.PL appartient à la classe B, 30 instancient la règle I. La probabilité conditionnelle de la règle I étant donné la classe B est donc de $30/319 \approx 0,094$. Les tableaux 9 et 10 donnent les détails dans les deux cas qui nous intéressent.

$P(\text{colonne} \text{ligne})$	I	II	III
A	1	0	0
B	0	1	0
C	0	0	1

Tableau 9 — Probabilité conditionnelle de la classe de $\langle \text{IPF.1SG}, \text{IPF.1PL} \rangle$ connaissant la classe de IPF.1SG

$P(\text{colonne} \text{ligne})$	I	II	III
A	0,92	0,08	0
B	0	0,09	0,91
C	0	1	0

Tableau 10 — Probabilité conditionnelle de la classe de $\langle \text{IPF.1PL}, \text{IPF.1SG} \rangle$ connaissant la classe de IPF.1PL

À partir de ces données, il est facile de calculer l'ENTROPIE CONDITIONNELLE D'UNE CASE B DU PARADIGME CONNAISSANT LA CASE A notée $H(B|A)$, et qui nous servira de mesure quantitative de la difficulté à prédire B à partir de A ⁹. L'entropie est une grandeur toujours positive qui évalue quantitativement l'incertitude associée à un certain espace de possibilités. Elle est habituellement calculée en utilisant la formule suivante, où A est une variable aléatoire (un ensemble de possibilités), et $P(a)$ est la probabilité d'une possibilité particulière a .

$$(8) \quad H(A) = - \sum_{a \in A} P(a) \log_2 P(a)$$

La formule en (8) est conçue de manière à ce que

- dans une situation où il n'y a aucune incertitude (l'une des possibilités est certaine, les autres sont impossibles), l'entropie est nulle ;
- dans une situation où il y a deux possibilités également probable, l'entropie est de 1 ; s'il y a 4 situations également probables, elle est de 2 ; etc.¹⁰ Toute augmentation du nombre de possibilités fait augmenter l'entropie.
- tout déséquilibre entre les probabilités des différentes possibilités fait diminuer l'entropie ; par exemple, s'il y a deux possibilités mais que l'une des deux est trois fois plus probable que l'autre, l'entropie est de $-(\frac{1}{4} \log_2 \frac{1}{4} + \frac{3}{4} \log_2 \frac{3}{4}) \approx 0,81$.

L'entropie conditionnelle permet d'évaluer dans quelle mesure la connaissance du choix d'une possibilité dans un premier espace fait diminuer l'incertitude quant aux possibilités dans un deuxième espace. Par exemple, supposons qu'on essaye de deviner la catégorie du prochain mot que va prononcer Paul. L'ensemble des possibilités est $A = \{\text{Nom}, \text{Verbe}, \text{Adjectif}, \text{Adverbe}, \dots\}$. Dans l'absolu, l'entropie de A dépend seulement de la fréquence de chaque catégorie dans la parole de Paul : elle est assez élevée. En revanche, si on connaît la catégorie du mot précédent que

⁹ Les notions discutées ici sont des concepts de base de théorie de l'information, qui sont expliquées en détail dans tous les manuels pertinents—voir par exemple Pierce (1980).

¹⁰ Cette calibration sur les choix binaires est une conséquence de l'utilisation du logarithme binaire \log_2 . L'utilisation d'une autre base logarithmique ne change pas les propriétés mathématiques de l'entropie, mais change la calibration.

Paul a prononcé, l'incertitude diminue : par exemple s'il vient de prononcer un déterminant, il y a peu de chances que le mot suivant soit un verbe.

L'entropie conditionnelle se calcule à partir d'une connaissance des probabilités conditionnelles liant deux variables aléatoires, selon la formule suivante :

$$(9) \quad H(B|A) = - \sum_{a \in A} P(a) \sum_{b \in B} P(b|a) \log_2 P(b|a)$$

(9) consiste à calculer pour chaque valeur de A l'entropie de B connaissant cette valeur particulière de A , puis à faire une moyenne de ces entropies « locales » pondérée par la probabilité de la valeur de A considérée.

Revenons au problème de l'évaluation du niveau de certitude avec lequel on peut prédire la forme occupant une case B du paradigme à partir de la forme occupant une autre case A . Pour ce faire, nous utilisons l'entropie conditionnelle du choix d'une classe de paires $\langle A, B \rangle$ connaissant la classe de A ¹¹. (10) est l'application à notre cas particulier de la définition de l'entropie conditionnelle. En (11) et (12) on applique ces définitions aux deux exemples ci-dessus, en nous appuyant sur les distributions de probabilités qui viennent d'être établies.

$$(10) \quad H(A, B|A) = - \sum_{a \in A} P(A) \sum_{c \in A \times B} P(c|a) \log_2 P(c|a)$$

$$(11) \quad H(\text{IPF.1PL}|\text{IPF.1SG}) = - (0,84 \times (1 \times \log_2 1 + 0 \times \log_2 0 + 0 \times \log_2 0) + \\ 0,11 \times (0 \times \log_2 0 + 1 \times \log_2 1 + 0 \times \log_2 0) + \\ 0,05 \times (0 \times \log_2 0 + 0 \times \log_2 0 + 1 \times \log_2 1)) \\ = -(0,84 \times 0 + 0,11 \times 0 + 0,05 \times 0) \\ = 0.$$

$$(12) \quad H(\text{IPF.1SG}|\text{IPF.1PL}) = - (0,84 \times (0,92 \times \log_2 0,92 + 0,08 \times \log_2 0,08 + 0 \times \log_2 0) + \\ 0,11 \times (0 \times \log_2 0 + 0,09 \times \log_2 0,09 + 0,91 \times \log_2 0,91) + \\ 0,05 \times (0 \times \log_2 0 + 1 \times \log_2 1) + 0 \times \log_2 0) \\ \approx - (0,84 \times 0,338 + 0,11 \times 0,048 + 0,05 \times 0) \\ \approx 0,39.$$

La conclusion est que l'entropie conditionnelle de l'IPF.1PL connaissant l'IPF.1SG est nulle : connaissant l'IPF.1SG d'un verbe, on peut prédire à coup sûr l'IPF.1PL, sur la seule base de la connaissance de la fréquence des différentes classes de paires alternantes. Par contre, l'entropie conditionnelle de l'IPF.1SG connaissant l'IPF.1PL est loin d'être nulle. On notera que l'incertitude est

¹¹ On se sert ici crucialement du fait que pour toutes variables aléatoires A et B , $H(A, B|A) = H(B|A)$. Il revient donc au même de calculer l'entropie conditionnelle de B sachant A et de calculer l'entropie conditionnelle de $A \times B$ sachant A .

ici entièrement due aux effets de deux règles phonologiques générales du français, rappelées en (13), et qui ont pour effet d'opacifier à l'IPF.1PL et l'IPF.2PL la frontière entre le radical et la marque de flexion : à cause de (13a), l'IPF.1PL d'ALLER est indistinguable de celui d'ALLIER, ce qui introduit de l'incertitude quant à la forme de son IPF.1SG. De même, à cause de (13b), l'IPF.1PL de CADRER est indistinguable de celui de QUADRILLER.

(13) a. $j \rightarrow \emptyset / j _$

b. $\emptyset \rightarrow i / C \left[\begin{array}{l} +\text{cons} \\ -\text{voc} \\ +\text{son} \\ -\text{nas} \end{array} \right] _ j$

2.2 Le tableau complet

La méthode de calcul décrite dans le paragraphe 1.4 a été implémentée sous la forme d'un script Python et appliquée systématiquement aux 6 440 verbes non défectifs de la base de données BDLEX. Les transcriptions de BDLEX ont été corrigées à la main, d'une part pour éliminer un certain nombre d'erreurs ponctuelles, d'autre part pour rendre la transcription plus proche de la phonologie de surface¹². Les résultats sont synthétisés dans la figure 1, et le tableau 11 donne les chiffres plus précis pour un sous-ensemble du paradigme.

¹² Dans un certain nombre de cas, BDLEX code des représentations phonologiques qui ne sont jamais prononcées comme telles. Par exemple *sabllez* est codé /sablje/, alors qu'il est systématiquement réalisé /sablje/. Si ce codage est approprié pour certaines applications, il a pour conséquence dans notre cas de faire baisser l'entropie entre certaines paires de cases sur la base d'informations que les locuteurs ne peuvent pas connaître.

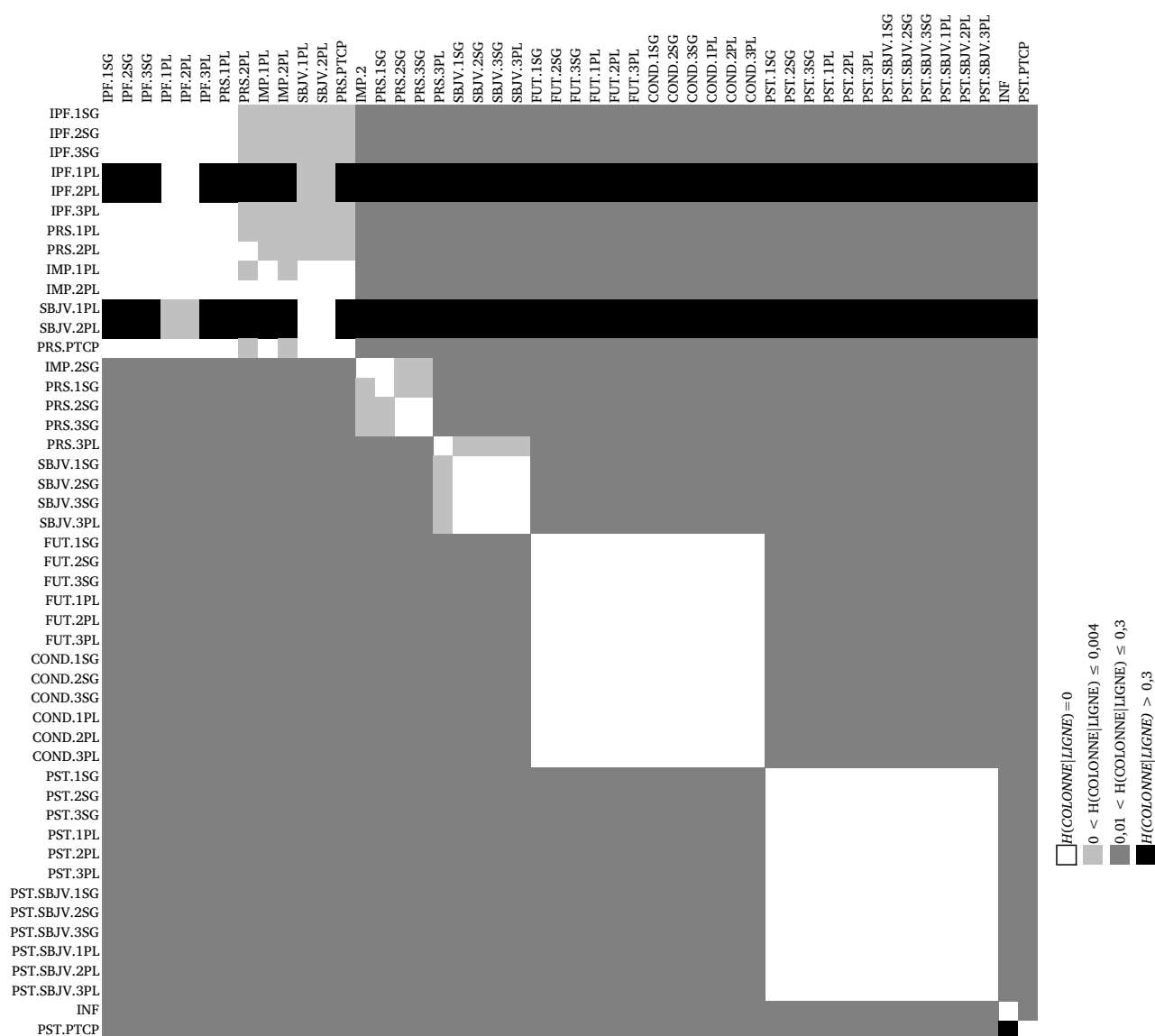


Figure 1 — Entropie conditionnelle entre les 48 formes des 6440 verbes non-défectifs de BDLEX¹³

¹³ Pour les besoins de ce travail, on a ignoré la variation en genre et en nombre des participes passés, pour des raisons techniques: beaucoup de verbes non transitifs directs sont défectifs pour les M.PL, F.SG et F.PL, faute de situation d'accord. La prise en compte des verbes défectifs aurait notablement compliqué les calculs, mais l'élimination des verbes sans accord du participe passé aurait éliminé certaines classes de verbes importantes.

	IPF.1SG	IPF.2SG	IPF.3SG	IPF.1PL	IPF.2PL	IPF.3PL	PRS.1PL	PRS.2PL	IMP.1PL	IMP.2PL	SBVJ.1PL	SBVJ.2PL	PRS.PTCP
IPF.1SG	—	0	0	0	0	0	0	0,0012	0,0003	0,0015	0,0003	0,0003	0,0003
IPF.2SG	0	—	0	0	0	0	0	0,0012	0,0003	0,0015	0,0003	0,0003	0,0003
IPF.3SG	0	0	—	0	0	0	0	0,0012	0,0003	0,0015	0,0003	0,0003	0,0003
IPF.1PL	0,3862	0,3862	0,3862	—	0	0,3862	0,3862	0,3870	0,3865	0,3873	0,0003	0,0003	0,3859
IPF.2PL	0,3862	0,3862	0,3862	0	—	0,3862	0,3862	0,3870	0,3865	0,3873	0,0003	0,0003	0,3859
IPF.3PL	0	0	0	0	0	—	0	0,0012	0,0003	0,0015	0,0003	0,0003	0,0003
PRS.1PL	0	0	0	0	0	0	—	0,0012	0,0003	0,0015	0,0003	0,0003	0,0003
PRS.2PL	0	0	0	0	0	0	0	—	0,0003	0,0003	0,0003	0,0003	0,0003
IMP.1PL	0	0	0	0	0	0	0	0,0012	—	0,0012	0	0	0
IMP.2PL	0	0	0	0	0	0	0	0	0	—	0	0	0
SBJF.1PL	0,3871	0,3871	0,3871	0,0011	0,0011	0,3871	0,3871	0,3881	0,3877	0,3887	—	0	0,3871
SBJV.2PL	0,3871	0,3871	0,3871	0,0011	0,0011	0,3871	0,3871	0,3881	0,3877	0,3887	0	—	0,3871
PRS.PTCP	0	0	0	0	0	0	0	0,0012	0	0,0012	0	0	—

Tableau 11 — Entropie conditionnelle pour quelques cases du paradigme des verbes du français

L'observation des données amène à distinguer trois classes de cas :

- A. L'entropie conditionnelle est nulle : la case en colonne est parfaitement prédictible à partir de la case en ligne
- B. L'entropie conditionnelle est quasi-nulle (inférieure à 0,004) : dans tous ces cas, il existe 1 à 5 lexèmes isolés qui introduisent une très légère incertitude. Le plus souvent l'incertitude est due à une paire de lexèmes qui ont des formes homophones pour la case en colonne mais dont l'un au moins est hautement irrégulier. Ainsi l'entropie conditionnelle de 0,000932 du PRS.1SG au PRS.2SG est-elle due aux homophonies au PRS.1SG, mais pas au PRS.2SG, des paires de lexèmes {être, suivre} (/sɥi/), {avoir, haïr} (/ɛ/), et {aller, vêtir} (/vɛ/)¹⁴.
- C. L'entropie conditionnelle est faible (entre 0,01 et 0,58) : grâce à l'information statistique sur la prévalence des différentes classes de paires alternantes, l'entropie n'est jamais très élevée, mais l'incertitude est significative dans ces cas.

La figure 1 fait en outre ressortir le fait que les cases du paradigme à partir desquelles l'incertitude est la plus élevée sont les IPF.1PL, IPF.2PL, SBVJ.1PL et SBVJ.2PL : à partir de ces cases, l'entropie conditionnelle est toujours supérieure à 0,38, sauf quand il s'agit de prédire une autre case de cette série. La seule autre paire de case qui donne lieu à une entropie supérieure à 0,3 est (PST.PTCP, INF), et encore l'entropie est-elle inférieure à 0,35. À l'examen, on se rend compte que cette situation est entièrement due aux neutralisations résultant des règles phonologiques en (13), qui ont donc un

¹⁴ Les données utilisées ici ne codent pas les « h aspirés ». Leur prise en compte réduirait marginalement l'entropie, puisque *hais* et *ai* ne seraient plus homophones.

effet dévastateur sur les résultats : si on recode les données de manière à faire abstraction des effets de ces règles, l'entropie est partout inférieure à 0,35.

2.3 La partition induite par l'entropie conditionnelle

Les résultats synthétisés dans la figure 1 nous donnent une image des relations d'interprédictivité entre case du paradigme que les espaces thématiques visent à capter. Si la partition induite par les espaces thématique de Bonami & Boyé (2003) est satisfaisante, on s'attendrait donc à ce qu'elle coïncide avec la partition induite par (14) : l'interprédictibilité coïncide avec l'entropie conditionnelle nulle.

(14) Deux cases du paradigme *A* et *B* appartiennent à la même cellule de la partition si et seulement si $H(A|B) = H(B|A) = 0$.

De fait on ne constate pas une telle coïncidence : la partition induite par les paires de cases à entropie nulle, donnée dans le tableau 12, possède 15 cases au lieu de 12.

	1SG	2SG	3SG	1PL	2PL	3PL
PRESENT	3a	3b		1a	1b	2
IMPARFAIT	1a			1c		1a
IMPERATIF		5		6		
SUBJONCTIF	7			8		
FUTUR CONDITIONNEL	10					
PASSE SIMPLE SUBJ. PASSE	11					

INFINITIF	PART. PRESENT	PART. PASSE	M.SG	M.PL	F.SG	F.PL
9	4	12				

TABLEAU 12 — Partition du paradigme des verbes du français induite par (11)

A première vue, cette prolifération du nombre de cases est entièrement due à deux phénomènes bien identifiés. D'une part, les neutralisations phonologiques dues aux règles (13) ont pour effet de séparer l'IPF.1SG et l'IPF.1PL du reste des formes basées chez Bonami & Boyé (2003) sur le radical 1. D'autre part, l'existence de formes fléchies supplétives au PRS.1SG et au PRS.2PL induit une faible entropie qui conduit à isoler ces cases dans une cellule séparée de la partition¹⁵. On est donc fondé à modifier la manière de construire la partition pour prendre en compte ces effets, qui sont

¹⁵ La forme fléchie supplétive *somme* ne conduit pas à isoler le PRS.1PL parce que l'algorithme de généralisation minimale décrit dans le paragraphe 2.1 identifie de manière non ambiguë que le seul radical d'imparfait compatible avec cette forme est *ét-*.

supposés ne pas relever de l'espace thématique par Bonami & Boyé (2003). On peut proposer de substituer à (14) une condition plaçant dans la même cellule deux cases reliées par une entropie minime, inférieure à un seuil conventionnel ε . L'idée est ici que pour la constitution de l'espace thématique, d'une part on fait abstraction des entropies quasi-nulles qui correspondent à ce que modélisent les formes fléchies supplétives, et d'autre part on se contente d'une entropie quasi-nulle dans une des deux directions, de manière à effacer les effets des neutralisations affectant l'une des deux cases et pas l'autre.

(15) Deux cases du paradigme *A* et *B* appartiennent à la même cellule de la partition si et seulement si soit $H(A|B) < \varepsilon$, soit $H(B|A) < \varepsilon$.

Dans le cas présent, l'observation des données de la figure 1 montre qu'il est raisonnable de placer le seuil à $\varepsilon = 0,01$ étant donné qu'il n'y a aucune paire de cases entre lesquelles l'entropie se situe dans l'intervalle $[0,004 ; 0,12]$. Le tableau 13 indique la partition résultant de ce choix. De manière cruciale, cette partition a 7 cases là où l'espace thématique de Bonami & Boyé (2003) en a 12 : le présent singulier tombe dans la même cellule que l'impératif singulier ; le présent 3PL forme une cellule avec le subjonctif singulier et 3PL, alors que le reste du subjonctif rejoint, avec le participe présent, la cellule de l'imparfait.

	1SG	2SG	3SG	1PL	2PL	3PL	
PRESENT				A		B	
IMPERATIF	C						
IMPARFAIT	B			A			
SUBJONCTIF							
FUTUR	D						
CONDITIONNEL							
PASSE SIMPLE							
SUBJ. PASSE	E						
INFINITIF	PART. PRESENT		M.SG		M.PL	F.SG	F.PL
F	A		G				
			PART. PASSE				

TABLEAU 13 — Partition du paradigme des verbes du français induite par (12)

L'explication de cette différence est claire : comme on peut le constater en examinant le tableau 11, les cellules de la partition de Bonami & Boyé (2003) qui sont fusionnées dans le tableau 13 contiennent des cases entre lesquelles l'entropie est très basse, d'un niveau analogue à l'entropie qui est induite par les formes fléchies supplétives. Les relations d'interprédictibilité sont donc excellentes, même si elles ne sont pas parfaites—le peu d'entropie que l'on constate est due au comportement d'une poignée de lexèmes très irréguliers et très fréquents.

Ce constat étant fait, il est utile de revenir sur la motivation initiale, chez Bonami & Boyé (2003), des distinctions qui sont absentes du tableau 13. Celles-ci n'étaient pas motivées par l'interprédictibilité entre les cases, mais par la tension entre deux objectifs : (i) réduire le nombre de cases de l'espace thématique, et (ii) autoriser une analyse constructive en radical et marque de flexion à chaque fois qu'une marque de flexion est identifiable. La tension entre (i) et (ii) est particulièrement claire si on examine par exemple le rapport entre participe présent et IPF.3SG. Les seuls lexèmes à posséder deux radicaux distincts dans les cases concernées sont SAVOIR (*sach-ant* vs. *sav-ait*), AVOIR (*ay-ant* vs. *av-ait*) et leurs dérivés. Dans tous les cas de ce type, Bonami & Boyé (2003) arbitrent en faveur de la segmentation, et décident donc de postuler deux cases séparées de l'espace thématique, malgré l'excellente interprédictibilité. Ce biais est clairement induit par l'esprit constructiviste de Bonami & Boyé (2003). Dans le cadre qu'ils définissent, la seule alternative serait d'abandonner (ii) et de générer *sachant* comme une forme fléchie supplétive inanalysable—une hypothèse jugée contre-intuitive par la plupart des morphologues, les auteurs de cet article compris.

3 Efforts en termes

La section précédente a confronté les hypothèses de Bonami & Boyé (2003) sur l'espace thématique des verbes du français avec les résultats d'une enquête empirique quantitative sur les relations d'interprédictibilité entre cases du paradigme. On a vu que les discordances entre les partitions du paradigme des verbes induites par l'analyse de Bonami & Boyé (2003) et par l'enquête empirique s'expliquaient par une tension entre la conception constructiviste de la morphologie implicitement adoptée par Bonami & Boyé (2003) et la prise en compte de l'ensemble des relations d'interprédictibilité. La question se pose alors de savoir si une analyse abstraite au sens de Blevins (2006) de la conjugaison du français souffrirait du même problème. Dans cette section nous montrons par l'exemple que ce n'est pas le cas.

Nous explicitons une méthode d'analyse abstraite conduisant d'une observation de surface des formes fléchies à une caractérisation de la partition des paradigmes verbaux du français. Cette méthode dérive de celle proposée par Boyé (2000) pour déterminer la « grille thématique » des verbes, et systématise les heuristiques implicitement utilisées par Bonami & Boyé (2003) pour réduire l'influence des irrégularités isolées. Le résultat est que l'on aboutit, sur une base conceptuellement bien fondée, à une partition du paradigme des verbes du français qui coïncide avec celle que permet d'observer l'approche purement observationnelle exposée dans la section 2.

3.1 Paradigmes et classes flexionnelles

Les espaces thématiques étant principalement une manière de décrire la variabilité des paradigmes de flexion, la base de la procédure tient dans l'identification des classes flexionnelles, qui repose elle-même sur une définition précise des paradigmes de flexion.

(16) La définition d'un paradigme met en jeu trois notions¹⁶ :

- a. Le paradigme abstrait d'un lexème est l'ensemble des combinaisons de propriétés morphosyntaxiques que ce lexème est susceptible de distinguer en vertu de son appartenance à une catégorie lexicale donnée.
- b. Le paradigme concret d'un lexème est l'ensemble des formes phonologiques qu'il peut manifester.
- c. Une fonction paradigmatique lie le paradigme abstrait d'un lexème à son paradigme concret. On note $FP(L, \sigma) = \rho$ le fait que la forme ρ est la réalisation de la combinaison de propriétés morphosyntaxiques σ pour le lexème L .

Pour définir de manière générale les systèmes de classes flexionnelles, il est important d'éviter tout biais concaténatif dans les formulations. On le fait ici en partant des fonctions morphologiques reliant les cases du paradigme d'un lexème, sans supposer que ces fonctions doivent prendre la forme de substitution d'un affixe à un autre.

- (17) a. Soit S un ensemble des segments phonologiques. On appelle FONCTION MORPHONOLOGIQUE toute fonction partielle de l'ensemble S^* des séquences de segments dans S^* .
- b. Un PATRON D'ALTERNANCE pour le paradigme abstrait Σ est une fonction partielle π de $\Sigma \times \Sigma$ dans l'ensemble des fonctions morphologiques tel que pour tous σ et σ' , $\pi(\sigma, \sigma')$ est définie si et seulement si $\sigma \neq \sigma'$.
- c. Un patron d'alternance π s'applique au lexème L si pour toutes paires de cases σ et σ' du paradigme abstrait de L , $\pi(\sigma, \sigma')(FP(L, \sigma)) = FP(L, \sigma')$.
- d. Pour tout ensemble Λ de lexèmes partageant le même paradigme abstrait Σ , un ensemble Π de patrons d'alternance est CLASSIFICATOIRE pour Λ si un et un seul patron de Π s'applique à chaque lexème de Λ .

Tout ensemble classificatoire de patrons d'alternance induit une partition de l'ensemble des lexèmes qu'il classe ; ce sont les cellules de cette partition qui servent de base à la classification en classes flexionnelles :

- (18) On appelle SYSTEME DE CLASSES FLEXIONNELLES ELEMENTAIRES pour un ensemble de lexèmes Λ , la partition de Λ induite par un ensemble de patrons d'alternance classificatoires pour Λ .

¹⁶ Notre paradigme abstrait est le « paradigme 1 » de Carstairs (1987) et Fradin (2003), notre paradigme concret est leur « paradigme 2 ». La notion de fonction paradigmatique est empruntée à Stump (2001), mais elle a ici une valeur descriptive plus que théorique : Stump (2001) constitue un ensemble de propositions spécifiques sur la manière dont les fonctions paradigmatiques doivent être modélisées. Nous nous contentons ici de supposer qu'une telle fonction existe sans nous commettre à une modélisation particulière.

Tout ensemble de lexèmes peut être caractérisé par des ensembles de patrons d'alternance divers, et donc donner lieu à des systèmes de classes élémentaires distincts—le choix d'un système particulier met en jeu des arbitrages entre la taille des classes et la naturalité des fonctions morphologiques postulées. De plus, les classes flexionnelles élémentaires peuvent partager des propriétés qui légitiment leur regroupement en macro-classes (Dressler & Thornton, 1996), dont l'inventaire est encore plus ouvert. Pour choisir dans la forêt de classifications possibles, il est utile de disposer de la notion de CLASSE FLEXIONNELLE IRREDUCTIBLE :

- (19) a. Deux classes flexionnelles c_1 et c_2 sont UNIFIABLES si il existe un critère phonologique ou morphosyntaxique qui permet de prédire de manière catégorique l'appartenance des lexèmes de $c_1 \cup c_2$ à l'une ou l'autre classe¹⁷.
- b. Une classe c_1 est IRREDUCTIBLE s'il n'existe pas de classe c_2 telles que c_1 et c_2 sont unifiables.

Les classes flexionnelles irréductibles sont celles dont aucun système de classe ne peut se passer ; elles constituent la base empiriquement stable de tout système de classification. D'après les définitions données ici, sauf à postuler (comme par exemple Boyé 2000) des distinctions phonologiques abstraites non observables indépendamment, les verbes TAPISSER et TAPIR ne peuvent pas appartenir à la même classe élémentaire, parce qu'il ne peut pas exister une unique fonction morphologique passant de *nous tapissons* à *je tapis* d'une part et à *je tapisse* d'autre part, et donc il ne peut exister un patron d'alternance unique s'appliquant aux deux lexèmes. En revanche les deux classes concernées sont unifiables sur la base de la différence systématique entre la forme de leurs infinitifs.

Pour les besoins des paragraphes suivants, il est utile de pouvoir caractériser les classes flexionnelles en fonction de leur niveau de peuplement :

- (20) a. Une classe flexionnelle est TRIVIALE si elle comporte un seul membre plus ses éventuels dérivés (par exemple DIRE et REDIRE).
- b. Une classe flexionnelle est SUBSTANTIELLE si elle n'est pas triviale et peut voir son effectif augmenter.

Les classes ouvertes sont substantielles ; pour les classes fermées mais bien peuplées, comme la classe des verbes français du deuxième groupe, on peut débattre de la question de savoir si elles doivent être considérées comme substantielles ou non (Kilani-Schoch & Dressler, 2005 ; Bonami *et al.* 2008).

L'hypothèse centrale de Bonami & Boyé (2003) peut être formulée comme l'idée qu'il existe en français une unique classe substantielle irréductible.

¹⁷ Notre notion de classes irréductibles est le complémentaire de la notion de classes en compétition chez Carstairs-McCarthy (1994 : 744-745).

3.2 La segmentation

La méthode d'analyse repose sur une méthode de segmentation des formes héritée de Boyé (2000) et qui peut être systématisée comme en (21) :

- (21) a. Pour toute case du paradigme abstrait σ , la fonction de flexion caractéristique de cette case est une fonction morphologique f telle que :
- i. pour tout lexème L appartenant à une classe flexionnelle non-triviale, il existe une séquence phonologique non-nulle φ telle que $f(\varphi)$ est la forme de L pour la case σ ;
 - ii. l'allongement produit par f est maximal : il n'existe pas de fonction f' qui vérifie la condition (i) et qui produise un allongement plus grand de son input que f ¹⁸.
- b. Pour tout lexème L (y compris les lexèmes appartenant à une classe flexionnelle triviale) et toute case du paradigme abstrait σ , le radical de L pour la case σ est la séquence φ telle que $f(\varphi)$ est la forme de L pour σ .

Dans le cas concaténatif, (21a) revient à dire que l'on maximise la longueur des marques de flexion. La procédure suggérée ici contraste donc fortement avec celle de Spencer (2012) : chez Spencer, on maximise la longueur du radical tant que la marque de flexion induite exprime un ensemble cohérent de propriétés morphosyntaxiques. Ici, on maximise la longueur de la marque de flexion tant que celle-ci est uniforme pour les classes flexionnelles non-triviales. Les deux procédures produisent des résultats différents, par exemple, dans le cas du r du futur et du conditionnel, qui est une marque de flexion selon (21a) et une portion de radical selon Spencer¹⁹.

L'exclusion des classes flexionnelles triviales en (21a) capturent une partie des exemples que Bonami & Boyé (2003) traitent comme des formes fléchies supplétives : les formes *sommes*, *êtes*, *faites* et *dites* ne sont pas prises en compte pour la détermination des marques de flexion parce qu'elles appartiennent à des lexèmes qui sont dans des classes flexionnelles triviales. (21) appliqué aux verbes du français aboutit donc bien à l'inventaire de marques de flexion détaillé dans le tableau 3. Cependant, bien que les classes triviales ne soient pas prises en compte pour établir les

¹⁸ Techniquement, les fonctions candidates à constituer une fonction de flexion sont toutes les fonctions morphologiques telles que pour tout L appartenant à une classe flexionnelle non triviale, $f^1(\text{FP}(L, \sigma))$ est définie. Une telle fonction est maximale s'il n'existe pas d'autre fonction candidate f' telle que

$$\mu_{L \in \Lambda} \left(\left| f'^{-1}(\text{PF}(L, \sigma)) \right| \right) < \mu_{L \in \Lambda} \left(\left| f^{-1}(\text{PF}(L, \sigma)) \right| \right)$$

Autrement dit, la fonction est maximale si elle revient en moyenne (notée μ) à supposer de plus grands allongements quand on passe du radical à la forme fléchie. Il est à noter que cette définition ne garantit pas l'unicité de la fonction maximale—à vrai dire sans contrainte supplémentaire il existe une infinité de fonctions maximales pour toute case du paradigme. En pratique cependant, nous n'avons pas rencontré de cas où le choix parmi les fonctions maximales avait des conséquences empiriques notables.

¹⁹ En supposant que le conditionnel et le futur ne forment pas une classe morphosyntaxique naturelle, ce qui ne va pas de soi — voir par exemple Verkuyl *et al.* (2004).

marques de flexion, rien n'empêche de supposer que leurs formes sont segmentables quand cela est compatible avec leur aspect phonologique ; ainsi AVOIR et ALLER, qui appartiennent à des classes triviales, sont segmentables et segmentés par (21b).

3.3 Abstraire la partition

On peut maintenant identifier une partition du paradigme abstrait partagé par un ensemble de lexèmes partitionné en classes flexionnelles en étendant la stratégie qui a été utilisée pour la segmentation.

(22) Soient σ et σ' deux cases du paradigme abstrait d'un ensemble de lexèmes de même paradigme abstrait, f et f' les fonctions de flexion associées. σ et σ' appartiennent à la même cellule de la partition si et seulement si pour tous les lexèmes appartenant à une classe flexionnelle non triviale, $f^1(\text{PF}(L,\sigma)) = f'^1(\text{PF}(L,\sigma'))$

(22) aboutit précisément, dans le cas du français, et en supposant les marques de flexion du tableau 3, à la partition du tableau 13. On a donc reconstruit, sur une base conceptuellement fondée, une partition du paradigme des verbes du français qui coïncide avec celle qu'a conduit à observer l'enquête quantitative de la section 2.

Il est important de noter que la procédure définie dans cette section, étant une procédure abstractive, conduit à des formulations qui peuvent sembler contre-intuitives mais qui ne sont pas contradictoires. Reprenons l'exemple du participe présent de SAVOIR *sachant*. (21a) conduit à conclure que la marque de flexion du participe présent est la suffixation de *-ant*, (21b) à conclure que le radical de *sachant* est *sach-*. La procédure aboutit donc à une segmentation complète de *sachant* en *sach+ant*, tout comme elle aboutit à une segmentation complète de *savait* en *sav+ait*. Pour autant, (22) conduit à conclure que le participe présent et l'imparfait partagent une même cellule de la partition. Nommer cette partition un espace thématique serait donc trompeur, dans la mesure où les cellules de la partition ne coïncident pas avec l'utilisation d'un radical unique. Nous l'appellerons donc la PARTITION THEMATIQUE. En outre la partition thématique, contrairement à l'espace thématique, ne fait sens que pour un ensemble de lexèmes, et pas pour un lexème individuel : il n'y a pas de sens à dire que le paradigme de SAVOIR en tant que tel place son participe présent et son imparfait dans la même cellule de la partition. En revanche on peut caractériser avec précision ce qui fait de SAVOIR un lexème remarquable :

(23) Un lexème L est INCONGRU si il existe deux cases de son paradigme abstrait σ et σ' appartenant à la même cellule de la partition thématique mais telles que $\text{FP}(L,\sigma)$ et $\text{FP}(L,\sigma')$ n'ont pas le même radical.

4 À terme, des formes?

La section précédente a montré qu'il est possible de construire une analyse abstractive de la conjugaison du français qui aboutit précisément à la postulation d'une partition du paradigme des verbes qui coïncide avec celle qui a été observée dans l'enquête empirique de la section 2. À ce point du travail, il est tentant de conclure que l'approche abstractive a fait la preuve de sa supériorité : non seulement elle capte le même effet de partition du paradigme que l'hypothèse des espaces thématiques, mais d'une part elle propose une partition qui correspond plus étroitement aux relations d'interprédictibilité observées, et d'autre part elle caractérise correctement la classe des lexèmes qui interdisent à l'approche constructive un tel succès : ce sont les lexèmes incongrus, qui connaissent un patron d'alternance radicale trop rare pour avoir une incidence sur la partition. La conclusion s'impose que le caractère constructif du cadre des espaces thématique leur interdit de capter l'ensemble des relations d'interprédictibilité. On est donc amené à conclure, comme Blevins (2006), que l'approche constructive n'apporte rien qui ne se déduise de la seule modélisation des relations de forme à forme.

Bibliographie

- Ackerman, Farrell, James Blevins & Robert Malouf. 2009. Parts and wholes: Implicative patterns in inflectional paradigms. In James Blevins & Juliette Blevins (éds), *Analogy in Grammar: Form and Acquisition*, pp. 54–82. Oxford : Oxford University Press.
- Albright, Adam. 2002. The identification of bases in morphological paradigms. PhD thesis, UCLA.
- Anderson, Stephen R. 1992. *A-morphous morphology*. Cambridge : Cambridge University Press.
- Aronoff, Mark. 1994. *Morphology by Itself*. Cambridge : The MIT Press.
- Arrivé, Michel. 1997. *La conjugaison pour tous*. Hatier : Bescherelle.
- Baayen, Harald, Ton Dijkstra & Robert Schreuder. 1997. Singulars and plurals in Dutch : Evidence for a parallel dual route model. *Journal of Memory and Language*, 37:94–117
- Baerman, Matthew, Dunstan Brown & Greville Corbett. 2005. *The Syntax-Morphology Interface: A study of syncretism*. Cambridge : Cambridge University Press.
- Blevins, James. 2003. Stems and paradigms. *Language* 79(4): 737–767.
- Blevins, James. 2006. Word-based morphology. *Journal of Linguistics* 42:531–573.
- Bonami, Olivier & Gilles Boyé. 2002. Suppletion and dependency in inflectional morphology. In Franck Van Eynde, Lars Hellan & Dorothee Beerman (éds.), *Proceedings of the HPSG'01 Conference*, pp. 51–70. Stanford : CSLI publications.
- Bonami, Olivier & Gilles Boyé. 2003. Supplétion et classes flexionnelles dans la conjugaison du français. *Langages* 152:102–126.
- Bonami, Olivier & Gilles Boyé. 2005. Construire le paradigme d'un adjectif. *Recherches linguistiques de Vincennes* 34:77–98.
- Bonami, Olivier and Gilles Boyé. 2006. Deriving inflectional irregularity. *Proceedings of the 13th International Conference on HPSG*, pp 39–59. Stanford : CSLI Publications.

- Bonami, Olivier & Gilles Boyé. 2007. Remarques sur les bases de la conjugaison. In Delais-Roussarie, Elisabeth & Laurence Labrune (éds), *Des sons et des sens : données et modèles en phonologie et en morphologie*, pp. 77–90. Hermès.
- Bonami, Olivier, Gilles Boyé, Hélène Giraudo & Madeleine Voga. 2008. Quels verbes sont réguliers en français ? *Actes du premier Congrès Mondial de Linguistique Française*, pp. 1511–1523. Paris : Institut de Linguistique Française.
- Bonami, Olivier, Gilles Boyé & Fabiola Henri. 2011. Measuring inflectional complexity: French and Mauritian. Présentation au colloque Quantitative Measures in Morphology and Morphological Development, San Diego, 2011.
- Bonami, Olivier, Gilles Boyé & Françoise Kerleroux. 2009. L'allomorphie radicale et la relation flexion-construction. In Bernard Fradin, Françoise Kerleroux & Marc Plénat (éds.), *Aperçus de morphologie du français*, pp. 103–125. Saint-Denis : Presses Universitaires de Vincennes.
- Boyé, Gilles 2000. *Problèmes de morpho-phonologie verbale en français, en espagnol et en italien*. Thèse de doctorat, Université Paris 7.
- Boyé, Gilles. 2011. Régularités et classes flexionnelles dans la conjugaison du français. In Michel Roché, Gilles Boyé, Nabil Hathout, Stéphanie Lignon & Marc Plénat (éds.), *Des unités morphologiques au lexique*, pp. 41–68. Paris : Hermès.
- Boyé, Gilles & Patricia Cabredo Hofherr. 2005. Régularité, irrégularité et exception dans la flexion verbale. *Faits de Langue* 25:161–164.
- Boyé, Gilles & Patricia Cabredo Hofherr. 2006. The structure of allomorphy in spanish verbal inflection. *Cuadernos de Lingüística* 13:9–24
- Brown, Dunstan 1998. Stem Indexing and Morphological Selection in the Russian Verb. In Ray Fabri, Albert Ortmann & Teresa Parodi (éds.), *Models of Inflection*, pp. 161–221. Tübingen : Niemeyer.
- Brown, Dunstan & Andrew Hippisley. 2012. *Network Morphology*. Cambridge : Cambridge University Press.
- de Calmès, Martine & Guy Pérennou. 1998. BDLEX: a Lexicon for Spoken and Written French. 1st *International Conference on Language Resources and Evaluation*, pp 1129–1136. ELRA : Grenade.
- Cameron-Faulkner, Thea & Andrew Carstairs-McCarthy. 2000. Stem alternants as morphological signata: evidence from blur avoidance in Polish nouns. *Natural Language and Linguistic Theory*, 18:813–835.
- Carstairs, Andrew. 1987. *Allomorphy in Inflexion*. Londres : Croom Helm.
- Carstairs-McCarthy, Andrew. 1994. Inflection Classes, Gender, and the Principle of Contrast. *Language*, 70(4):737–788.
- Chomsky, Noam & Morris Halle. 1968. *The Sound Patterns of English*. Cambridge : The MIT Press
- Dressler, Wolfgang & Anna Thornton. 1996. Italian nominal inflection. *Wiener Linguistische Gazette*,

55–57:1–26.

- Fradin, Bernard. 2003. *Nouvelles approches en morphologie*. Paris : Presses Universitaires de France.
- Hippisley, Andrew. 1998. Indexed stems and Russian word formation: a Network Morphology account of Russian personal nouns. *Linguistics* 36:1039–1124.
- Kerleroux, Françoise. 2007. On a subclass of non-affixed deverbal nouns in French. In Geert Booij, Bernard Fradin, Angela Ralli & Sergio Scalise (éds.), *On-line Proceedings of the 5th MMM*.
- Kilani-Schoch, Marianne & Wolfgang Dressler. 2005. *Morphologie naturelle et flexion du verbe français*. Tübingen : Gunter Narr Verlag.
- Maiden, Martin 1992. Irregularity as a determinant of morphological change. *Journal of Linguistics* 28:285–312.
- Matthews, Peter. 1972. *Inflectional morphology: A theoretical study based on aspects of Latin verb conjugation*. Cambridge : Cambridge University Press.
- Montermini, Fabio & Gilles Boyé. 2012. Stem relations and inflection class assignment in Italian. *Word Structure* 5.1 : 69-87.
- Pierce (1980). *An introduction to Information Theory*. New York : Dover Publications.
- Pirrelli, Vito & Battista, Marco 2000. The Paradigmatic Dimension of Stem Allomorphy in Italian Verb Inflection. *Rivista di Linguistica* 12(2). pp. 307–380.
- Plénat, Marc. 2008. Le thème L de l'adjectif et du nom. *Actes du premier Congrès Mondial de Linguistique Française*, pp. 1613–1626. Paris : Institut de Linguistique Française.
- Plénat, Marc & Gilles Boyé. Sous presse. Le choix des thèmes dans les dérivés désadjectivaux en français. In Bernard Tranel (éd.), *Understanding Allomorphy: Perspectives from Optimality Theory*. Sheffield : Equinox Publishing.
- Roché, Michel. 2010. Base, thème, radical. *Recherches Linguistiques de Vincennes* 39:95–133.
- Spencer, Andrew. 2012. Identifying Stems. *Word Structure* 5.1 : 88-108.
- Stump, Gregory. 2001. *Inflectional morphology*. Cambridge : Cambridge University Press.
- Tribout, Delphine 2010. *Les conversions de nom à verbe et de verbe à nom en français*. Thèse de doctorat, Université Paris 7.
- Tribout, Delphine. 2012. Verbal stem space and verb to noun conversion in French. *Word Structure* 5.1 : 109-128
- Verkuyl, Henk, Co Vet, Andrée Borillo, Myriam Bras, Anne Le Draoulec, Arie Molendijk, Henriëtte de Swart, Carl Vetter & Laure Vieu. 2004. Tense and aspect in sentences. In Francis Corblin & Henriëtte de Swart (éds.), *Handbook of French Semantics*, pp. 233–270. Stanford : CSLI Publications.