Un algorithme universel pour l'abstraction automatique d'alternances morphophonologiques

Sacha Beniamine

Un algorithme universel pour l'abstraction automatique d'alternances morphophonologiques

Sacha Beniamine
(1) Laboratoire de Linguistique Formelle, Labex EFL, Université Paris Diderot 7, France
sbeniamine@linguist.univ-paris-diderot.fr

RÉSUMÉ

Cet article présente un algorithme implémenté pour l'inférence de patrons d'alternances morphophonologiques entre mots-formes. Il est *universel* au sens où il permet d'obtenir des classifications comparables d'une langue à l'autre sans préjuger des types d'alternances. Les patrons constituent une première étape pour les travaux quantitatifs dans l'approche Mot et Paradigme de la morphologie.

ABSTRACT

A universal algorithm for the automatical abstraction of morphophonological alternations.

We present an implemented algorithm for the inference of morphophonological alternation patterns between word-forms. It is *universal* in that it leads to comparable classifications across languages without expectations on the shape of the alternations it searches for. Alternation patterns are a necessary first step for the quantitative study of morphology in the Word and Paradigm framework.

MOTS-CLÉS : Flexion, morphophonologie, linguistique quantitative, typologie.

KEYWORDS: Inflection, morphophonology, quantitative linguistics, typology.

1 Introduction

L'étude quantitative et typologique de la morphologie présuppose l'extraction des données morphophonologiques à partir de lexiques. Les travaux existant en morphologie quantitative reposent sur des grammaires élaborées manuellement (Ackerman & Malouf, 2013; Stump & Finkel, 2013) difficilement comparables d'une langue à l'autre, ou sur des heuristiques fixant des biais linguistiquement motivés qui ne se prêtent pas aux comparaisons typologiques (Albright & Hayes, 2002).

Cet article présente un algorithme pour l'abstraction de patrons morphophonologiques permettant de caractériser des alternances entre mots-formes. Il est implémenté et distribué sous licence libre [1]. Il est *universel* au sens où il permet d'obtenir des classifications comparables d'une langue à l'autre sans préjuger des types d'alternances. Il a été développé dans le cadre de l'approche « mot et paradigme » pour l'analyse des paradigmes de flexion et peut être utilisé pour décrire d'autres alternances morphophonologique entre mots-formes.

Nous utilisons le vocabulaire standard en morphologie flexionnelle (Matthews, 1991). Un paradigme est l'ensemble des formes d'un lexème, structuré en fonction des propriétés morphosyntaxiques exprimées par ces formes (par exemple « présent première personne » ou PRS.1SG). On appelle *case*

1. Documentation et téléchargement : http://drehu.linguist.univ-paris-diderot.fr/qumin/. Le programme est écrit en Python3.

du paradigme un ensemble de propriétés à exprimer. Dans ce contexte, un système flexionnel peut être représenté sous la forme d'une table à double entrée, où chaque ligne représente le paradigme d'un lexème et chaque colonne une case de paradigme.

Les modèles « Mot et Paradigme » (WP) suivent la tradition grammaticale européenne et considèrent le mot et non le morphème comme l'unité fondamentale de la morphologie (Hockett, 1954; Robins, 1959). Ces modèles décrivent traditionnellement les systèmes de flexion en se fondant sur l'analogie à un petit nombre de paradigmes exemplaires. Les développements récents mettent au centre de l'attention les relations implicatives entre cases du paradigme (Wurzel, 1989) et la mesure quantitative de la fiabilité de ces implications (Ackerman *et al.*, 2009; Sims, 2010; Ackerman & Malouf, 2013; Blevins, 2016; Ackerman & Malouf, 2016). La formulation précise de ces implications nécessite d'identifier des patrons d'alternance entre formes de surface (Bonami, 2014).

/amɛn/ PRS.1SG 'amène' ⟷ _ɛ_ ⇌ _ø_e / am_n_ ⟷ /amøne/ PRS.2PL 'amenez'

FIGURE 1 – Un patron d'alternance

Les patrons d'alternances tels que nous les formulons ici prennent la forme d'alternances bidirectionnelles du type « *X alterne avec Y dans le contexte Z* », noté $X \rightleftharpoons Y/Z$, comme illustré en figure 1. X et Y constituent l'alternance et Z constitue le contexte d'application du patron.

L'une des questions centrales étudiées dans ce cadre est celle du problème de remplissage des cases de paradigme (« Paradigm Cell Filling Problem » ou PCFP). Étant donnée la distribution zipfienne des mots, comment les locuteurs qui n'ont rencontré que certaines formes d'un lexème peuvent-ils produire les formes inconnues ? Ackerman *et al.* (2009) évaluent la difficulté de ce problème à travers l'entropie des distributions de patrons d'alternance dans les systèmes flexionnels. Les résultats de ce type de calcul dépendent crucialement de la façon dont les patrons sont inférés.

Nous présentons d'abord (§2) les stratégies existantes pour extraire des alternances morphophonologiques. Puis (§3) nous présentons le problème comme une recherche d'alignements entre formes qui soient optimaux à la fois localement à ces formes et globalement à travers le système. L'algorithme (§4) procède en trois étapes : aligner localement les formes, généraliser les patrons, sélectionner les patrons globalement. Enfin nous évaluons la qualité des généralisations capturées par les patrons (§5) dans le cadre d'un problème du type PCFP idéalisé.

2 État de l'art

Deux traditions ont contribué à l'inférence automatique à grande échelle de règles morphologiques à même de capter les implications paradigmatiques. En linguistique, le *Minimal Generalization Learner* (MGL, Albright & Hayes 2002, 2003, 2006) constitue le premier modèle quantitatif des tests psycholinguistiques d'élicitations morphologiques dits *wug tests*, et a inspiré des travaux qui s'appuient sur des biais linguistiquement motivés pour inférer efficacement des patrons d'alternance dans un ensemble de données spécifique. En TAL, le PCFP se pose dans le contexte de la complétion automatique de lexiques et s'est traduit par la formulation d'une tâche de réinflection.

Le MGL est un programme qui infère incrémentalement, à partir de couples de formes, un ensemble

de patrons d'alternance unidirectionnels. Il infère d'abord des patrons spécifiques à une seule paire de formes, puis crée des patrons de généralité croissante au fil de la découverte de nouvelle formes. Pour chaque paire de formes dans les données d'apprentissage, on obtient un ensemble de patrons de généralité variable à même de les dériver. L'ensemble de ces règles sont scorées pour l'application aux formes inconnues. Les patrons générés par le MGL sont capables de découvrir un changement unique dans les formes d'entrées, cherchant dans l'ordre un changement suffixal, préfixal puis interne.

Bonami & Boyé (2014); Bonami & Luís (2014) combinent l'idée directrice du MGL et celle d'Ackerman *et al.* (2009), en évaluant la prédictibilité au sein des paradigmes au moyen de l'entropie conditionnelle sur la base de patrons d'alternance bidirectionnels. Un seul patron est alors associé à toute paire de cases afin de fournir une classification des lexèmes. Ce choix permet d'appliquer ces programmes à des larges ensembles de données et nombre d'alternances, ce qui n'était pas envisageable pour le MGL. Pour identifier les alternances, ces modèles s'appuient sur des biais adaptés à un ou plusieurs ensembles de données. Ces stratégies nécessitent la conception d'un nouveau programme *ad hoc* pour chaque nouveau système. Elles ne se prêtent donc pas à la typologie quantitative.

L'extraction de lexiques à partir de corpus produit des paradigmes incomplets en raison de la distribution zipfienne des formes. Des efforts se sont donc tournés vers l'inférence des formes manquantes en s'appuyant sur la structure implicative des paradigmes (Durrett & DeNero, 2013; Ahlberg *et al.*, 2014; Nicolai *et al.*, 2015). La campagne d'évaluation SIGMORPHON (Cotterell *et al.*, 2016) a formalisé une famille de tâches de TAL dites de « réinflection » répondant à cette manifestation concrète du PCFP. Le problème qui nous intéresse le plus consiste à prédire une forme cible à partir d'une forme déjà fléchie. Les approches par réseaux neuronaux sont les plus performantes (Kann & Schütze, 2016). D'autres systèmes, suivant Durrett & DeNero (2013), extraient des opérations d'édition de chaînes puis entraînent des transducteurs à les appliquer. Enfin certains reposent sur des heuristiques linguistiques et extraient soit des segmentations affixales (Taji *et al.*, 2016) soit des « paradigmes abstraits », règles morphologiques formulées pour l'ensemble du paradigme (Sorokin, 2016; Ahlberg *et al.*, 2014). Le meilleur modèle parvient à 95% d'exactitude en s'entraînant uniquement sur des paires de formes, montrant que l'information nécessaire à la résolution du PCFP peut être extraite de données éparses.

Cependant, aucun de ces systèmes ne permet d'extraire le type de représentations dont nous avons besoin, c'est-à-dire des patrons d'alternance entre deux cases, interprétables de façon transparente comme des règles linguistiques, calculables pour toutes les paires de cases sur de larges lexiques, ne s'appuyant pas sur des données externes à la paire de case concernée, et fournissant une classification catégorique des lexèmes selon leur comportement morphophonologique. Nous proposons donc une méthode à même d'inférer automatiquement de telles représentations.

3 Les patrons dépendent d'alignements entre formes

Déterminer le patron qui relie une paire de formes revient à trouver un alignement optimal entre ces formes : si deux segments alignés ensemble sont identiques, ils font partie du contexte, sinon ils font partie de l'alternance. Cet alignement ne peut pas, dans une perspective multi-lingue, être fixé à l'avance. Différents types d'alternances requièrent des stratégies d'alignements distinctes. Les alternances suffixales, par exemple dans la flexion verbale du français (figure 2), sont captées par un alignement à gauche. Une alternance préfixale nécessiterait un alignement à droite. Certaines alternances discontinues nécessitent un alignement qui ne peut être pré-défini (flexion verbale de

l'arabe, figure 2).

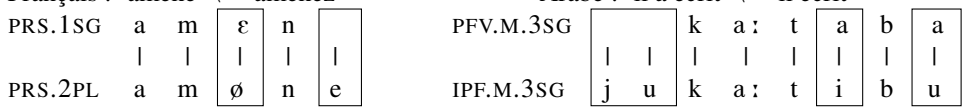

FIGURE 2 – Les alignements optimaux dépendent du type d'alternance.

Il existe parfois plusieurs alignements plausibles (Bonami, 2014, pp.104-106). La figure 3 présente trois alignements des formes imaginaires /baba/ et /ba/ et trois patrons qui offrent respectivement une analyse préfixale, suffixale ou infixale. Seul l'examen d'autres formes permet de décider de l'analyse préférable. Pour chacun des systèmes A, B et C de la figure 3, une seule des trois analyses, respectivement infixe, préfixe et suffixe, permet de rendre compte de l'ensemble des lexèmes.

Alignement					Patron	A.		B.		C.	
	b	a	b	a		SG	PL	SG	PL	SG	PL
	b	a	b	a		ba	baba	ba	baba	ba	baba
A.Infixe	b	_	_	a	$\epsilon \rightleftharpoons$ ab / b_a	ri	rabi	ri	bari	ri	riba
B.Préfixe	_	_	b	a	$\epsilon \rightleftharpoons$ ba / _ba	su	sabu	su	basu	su	suba
C.Suffixe	b	a	_	_	$\epsilon \rightleftharpoons$ ba / ba_	ne	nabe	ne	bane	ne	neba

FIGURE 3 – Trois alignements et systèmes imaginaires.

4 Méthode

Notre système prend en entrée des paradigmes dont les formes sont transcrites en notation phonétique ainsi qu'une spécification de la valeur des phonèmes utilisés en traits. Il calcule, pour chaque paire de cases possible parmi l'ensemble des cases du paradigme, un ensemble de patrons d'alternance qui relient les paires de formes. Malgré la complexité de cette combinatoire, les besoins de l'analyse linguistique de type Mot et Paradigme imposent de calculer les patrons pour toutes les paires de cases. Cependant chaque patron se fonde strictement sur la connaissance des formes des deux cases qu'il relie.

L'algorithme procède en trois étapes : produire pour chaque paire de formes l'ensemble des alignements localement optimaux et en déduire des patrons élémentaires ; généraliser les patrons élémentaires en fusionnant les alternances structurellement identiques ; choisir les patrons selon leur pouvoir descriptif global pour l'ensemble des lexèmes.

Nous commençons par trouver pour chaque paire de formes un ensemble d'alignements qui minimisent une distance d'édition. Nous considérons deux mesures de distance. La plus simple est la distance de Levenshtein, qui maximise les segments identiques alignés ensemble. Les trois alignements linguistiquement motivés entre /ba/ et /baba/ sont aussi ceux qui présentent la plus petite distance de Levenshtein. Considérons cependant l'alternance en Tchèque entre le nominatif singulier « čivava » prononcé /tʃɪvava/ et le génitif pluriel du nom « čivav » (chihuahua), prononcé /tʃɪvaf/. L'alignement intuitif pour un linguiste, en raison du dévoisement final, aligne [f] avec le deuxième

[v]. Cependant, il existe un deuxième alignement, non linguistiquement fondé, qui suppose un infixe /-av-/ et présente la même distance de Levenshtein.

Afin d'éviter ce type d'alignement fallacieux, et d'évaluer leur impact, nous implémentons également une seconde distance d'édition pondérée par la similarité phonologique (Albright & Hayes, 2006). (Frisch *et al.*, 2004) proposent une similarité phonologique fondée sur la proportion de classes naturelles partagées entre deux segments. En phonologie, une classe naturelle est un ensemble de segments phonologiques partageant un ou plusieurs traits (Chomsky & Halle, 1968). Les segments d'une même classe se comportent identiquement relativement à certains processsus phonologiques. Soit $C(a)$ l'ensemble des classes naturelles d'un segment a, la similarité phonologique peut se définir $\text{sim}(a,b) = \frac{|C(a) \cap C(b)|}{|C(a) \cup C(b)|}$. Nous fixons le coût de la substitution à $1 - \text{sim}(a,b)$. Le coût de l'insertion est un paramètre que nous fixons à $\frac{1}{3}$ du coût moyen de la substitution dans le système phonologique. Soit I l'ensemble des phonèmes connus et $I \times I$ le produit cartésien sur cet ensemble, le coût moyen d'une substitution dans ce système est $\frac{1}{I \times I} \cdot \sum_{(a,b) \in I \times I} 1 - \text{sim}(a,b)$.

Nous obtenons un ensemble de patrons concurrents pour chaque paire de formes. L'alignement de /ba/ et /baba/ produit l'ensemble des trois patrons : {$\epsilon \rightleftharpoons$ ba / _ba, $\epsilon \rightleftharpoons$ ba / ba_, $\epsilon \rightleftharpoons$ ab / b_a}. Ces patrons sont spécifiques aux formes dont ils proviennent, car leurs contextes contiennent des segments phonologiques spécifiques.

La seconde étape consiste à fusionner les patrons présentant la même alternance structurelle de façon à capter des généralisations sur les contextes d'application et sur les alternances. Les contextes sont alors exprimés sous la forme d'expressions régulières où les classes de caractères, entre crochets, correspondent à des classes naturelles de segments phonologiques et les quantificateurs « ?, +, * » rendent compte de la longueur des séquences concernées. Nous fusionnons enfin les patrons dont les alternances constituent différentes instanciations d'une même opération phonologique régulière.

Alternances	Contextes							Lexèmes
ɛ _ ⇌ ə_õ /			a	m	_	n	_	AMENER
ɛ _ ⇌ ə_õ /	p	ʁ	o	m	_	n	_	PROMENER
ɛ _ ⇌ ə_õ /				s	_	v ʁ	_	SEVRER
	{p} ?	{ʁ} ?	{o,a} ?	{m,s}	_	{n,v,ʁ}+	_	(i)
	{p,ʁ,o,a}*			{m,s}	_	{n,v,ʁ}+	_	(ii)
ɛ _ ⇌ ə_õ /	[eɛoɔøœabdflpstvzəʁ]*			[flmnsvz]	_	[lmnvzʁ] +	_	(iii)

TABLE 1 – Généralisation du contexte de trois patrons

La généralisation des contextes suit la procédure décrite par Albright & Hayes (2002), adaptée pour une généralisation n-par-n plutôt que 2 par 2. Étant donné un ensemble de patrons partageant une même alternance structurelle, on souhaite déterminer ce que ces contextes ont en commun. Pour cela, on aligne les segments à travers les contextes, puis les segments alignés sont traduits sous la forme d'une expression régulière.

La table 1 présente trois patrons d'alternance dont les contextes doivent être fusionnés. (i) Les séquences sont d'abord alignées du côté des blancs : à gauche du premier blanc, les séquences « am », "pʁom et « s » sont alignées à gauche. Entre les deux blancs, les séquences « n »,« n » et « vʁ » sont alignées au centre. (ii) Les ensembles de segments alignés facultatifs consécutifs sont ensuite fusionnés. (iii) Pour chaque ensemble de segments alignés, on détermine l'ensemble des segments qui partagent les mêmes traits phonologiques. On en déduit une expression régulière qui caractérise

Langue	Lexique	Traits phonologiques
Français	Vlexique (Bonami *et al.*, 2014)	Dell 1973
Anglais	CELEX2 (Baayen *et al.*, 1995)	Halle & Clements 1983
Chatino	Oto-Manguean Inflectional Class Database (Feist & Palancar, 2015)	Campbell 2014
portugais	Dictionnaire de prononciation (Veiga *et al.*, 2013)	Bonami & Luís 2014
Arabe	Unimorph (Kirov *et al.*, 2016)	Hayes 2012

TABLE 2 – Sources pour les ensembles de données utilisés dans l'évaluation.

les trois contextes.

Nous obtenons plusieurs patrons concurrents pour une même paire de formes, parmi lesquels il nous faut choisir celui qui est le plus approprié au vu du reste du système. Pour cela, nous attribuons des scores aux patrons. Intuitivement, un patron est meilleur s'il est souvent correct et un patron général est préférable à un patron spécifique. La précision d'un patron pour son application dans une direction $C_1 \rightarrow C_2$ est le rapport entre le nombre de formes appartenant à la case C_1 dont il dérive correctement la forme pour la case C_2 sur le nombre de formes appartenant à la case C_1 auxquelles ce patron est applicable. Sa couverture est le rapport entre le nombre de formes de la case C_1 auxquelles il est applicable et le nombre de formes total. Les patrons, qui sont bidirectionnels, sont scorés par la moyenne harmonique des couvertures et précisions dans les deux sens d'application. La généralisation du patron infixe de la figure 3 produira un patron du type $\epsilon \rightleftharpoons \text{ab}/C\_V$. Il aura un score de 1, car il dérive toutes les formes correctement. Les deux autres patrons ne sont pas généralisables et obtiendront un score inférieur.

Pour chaque paire de formes, nous choisissons parmi l'ensemble des patrons qui les dérivent correctement celui dont le score est le meilleur. Ce patron peut avoir été inféré à partir d'autres paires de formes. Ceci permet d'obtenir des patrons robustes à bas coût computationnel.

5 Évaluation

Notre algorithme vise à la comparaison quantitative de systèmes flexionnels. La pertinence des patrons inférés se jugera donc *in fine* aux généralisations trans-linguistiques qu'ils permettront de dégager. Cependant, leur capacité à capter des généralisations internes à un système peut d'ores et déjà être évaluée au travers d'une tâche de prédiction.

Nous avons mené une évaluation pour les systèmes verbaux du français, de l'anglais, du portugais européen, du zenzontepec chatino et de l'arabe classique. Les tables de paradigmes sont issues de lexiques flexionnels. Au besoin, nous avons transcrit automatiquement les formes en alphabet phonétique international depuis d'autres notations (zenzontepec chatino, anglais, arabe). La table 2 présente leur sources ainsi que celles des traits distinctifs. La première partie de la table 3 rapporte la taille des paradigmes.

Nous avons mené une validation croisée en 10 plis. L'ensemble des patrons est chaque fois appris sur les formes d'entraînement. Puis, sur ces mêmes données, nous déterminons les classes de patrons applicables (Bonami, 2014), c'est-à-dire les ensembles de patrons susceptibles d'être appliqués à chaque forme. Nous calculons la probabilité conditionnelle $P(\text{patron}|\text{classe})$ d'après les fréquences relatives des patrons et des classes ainsi que la probabilité simple $P(\text{patron})$ pour servir de repli. Lors de l'évaluation, nous considérons des formes uniques, dont nous connaissons la case de paradigme, et

	Chatino	Français	portugais	Arabe	Anglais
Lexèmes d'entraînement	352	4688	1797	576	5450
Lexèmes d'évaluation	39	521	199	64	605
Cases de paradigme	4	51	69	109	8
Alignement à gauche (suffixe)	23.39%	94.35%	93.92%	40.53%	94.03%
Alignement à droite (préfixe)	54.12%	23%	18.04%	24.27%	31.09%
Albright & Hayes 2002	53.67%	94.35%	93.67%	42%	94.01%
Distances d'édition simples	**57.37%**	**94.55%**	**92.77%**	**82.42%**	**94.18%**
Distances d'édition pondérées	**57.45%**	**94.53%**	**92.86%**	**82.58%**	**94.18%**

TABLE 3 – Résultats de l'évaluation : pourcentage d'exactitude moyenne.

tentons de prédire la forme cible dans la case cible. À partir de l'ensemble des patrons applicables à cette forme, nous choisissons le patron qui maximise $P(\text{patron}|\text{classe})$ si la classe est connue, sinon celui qui maximise $P(\text{patron})$.

Nous comparons les résultats du même algorithme en faisant varier la stratégie d'alignement (table 3). Afin de comparer notre proposition à des alternatives plus simples, nous évaluons les alignements fixes à gauche ou à droite ainsi que l'alignement décrit dans Albright & Hayes (2002). Nous évaluons également les deux alignements fondés sur des distances d'édition décrits dans cet article. De façon générale, les résultats sont nettement meilleurs sur les lexiques de taille importante (français, portugais, anglais) que sur les plus petits échantillons (chatino, arabe). Comme attendu, les stratégies d'alignement fixe sont efficaces exclusivement pour certains ensembles de données. Seul l'alignement par distances d'éditions produit de bons résultats sur l'ensemble des langues. Nos résultats ne montrent pas de différence importante entre les deux distances d'édition. La distance de Levenshtein a l'avantage de la simplicité. Elle produit initialement plus de patrons concurrents, ce qui conduit à un temps d'exécution plus long. Cependant l'algorithme est robuste à ce bruit, la sélection des patrons se faisant *in fine* selon leur adaptation à l'ensemble du paradigme.

6 Conclusion

Nous avons présenté un algorithme permettant d'inférer des patrons d'alternances à partir de paires de formes. Cet algorithme est robuste et applicable à travers les langues sans présumer du type d'alternances. Il peut constituer le fondement d'une analyse quantitative des alternances morphophonologiques, aussi bien dans le domaine de la flexion, pour lequel il a été développé, que pour l'étude des changements diachroniques, dialectaux, et des alternances dérivationnelles.

Remerciements

Nous remercions les trois relecteurs anonymes pour leur nombreuses remarques, Olivier Bonami pour ses conseils et suggestions, Enrique Palancar pour la préparation des données du chatino et Kenza Ouldhamouda pour la validation manuelle de la transcription des données de l'arabe.

Références

ACKERMAN F., BLEVINS J. P. & MALOUF R. (2009). *Parts and wholes : Patterns of relatedness in complex morphological systems and why they matter*, In J. P. BLEVINS & J. BLEVINS, Eds., *Analogy in Grammar : Form and Acquisition*, p. 54–82. Oxford University Press.

ACKERMAN F. & MALOUF R. (2013). Morphological organization : The low conditional entropy conjecture. *Language*, **89**(3), 429–464.

ACKERMAN F. & MALOUF R. (2016). Word and pattern morphology : An information-theoretic approach. *Word Structure*, **9**(2), 125–131.

AHLBERG M., FORSBERG M. & HULDEN M. (2014). Semi-supervised learning of morphological paradigms and lexicons. In *Proceedings of the 14th Conference of the European Chapter of the Association for Computational Linguistics, Gothenburg, Sweden 26–30 April 2014*, p. 569–578.

ALBRIGHT A. & HAYES B. (2002). Modeling english past tense intuitions with minimal generalization. In *Proceedings of the ACL-02 Workshop on Morphological and Phonological Learning - Volume 6*, MPL '02, p. 58–69, Stroudsburg, PA, USA : Association for Computational Linguistics.

ALBRIGHT A. & HAYES B. (2003). Rules vs. analogy in english past tenses : A computational/experimental study. *Cognition*, **90**, 119–161.

ALBRIGHT A. & HAYES B. (2006). Modeling productivity with the gradual learning algorithm : The problem of accidentally exceptionless generalizations. *Gradience in grammar : Generative perspectives*, p. 185–204.

BAAYEN R., PIEPENBROCK R. & GULIKERS L. (1995). Celex2.

BLEVINS J. P. (2016). *Word and Paradigm Morphology*. Oxford University Press (OUP).

BONAMI O. (2014). La structure fine des paradigmes de flexion. Mémoire d'habilitation U. Paris Diderot.

BONAMI O. & BOYÉ G. (2014). De formes en thèmes. In F. VILLOING, S. LEROY & S. DAVID, Eds., *Foisonnements morphologiques. Etudes en hommage à Françoise Kerleroux*, p. 17–45. Presses Universitaires de Paris Ouest.

BONAMI O., CARON G. & PLANCQ C. (2014). Construction d'un lexique flexionnel phonétisé libre du français. In F. NEVEU, P. BLUMENTHAL, L. HRIBA, A. GERSTENBERG, J. MEINSCHAEFER & S. PRÉVOST, Eds., *Actes du quatrième Congrès Mondial de Linguistique Française*, p. 2583–2596.

BONAMI O. & LUÍS A. R. (2014). Sur la morphologie implicative dans la conjugaison du portugais : une étude quantitative. In J.-L. LÉONARD, Ed., *Morphologie flexionnelle et dialectologie romane. Typologie(s) et modélisation(s).*, number 22 in Mémoires de la Société de Linguistique de Paris, p. 111–151. Leuven : Peeters.

CAMPBELL E. (2014). *Aspects of the phonology and morphology of Zenzontepec Chatino, a Zapotecan language of Oaxaca, Mexico*. PhD thesis, University of Texas at Austin.

CHOMSKY N. & HALLE M. (1968). *The Sound Pattern of English*. Studies in language. New York, NY : Harper & Row.

COTTERELL R., KIROV C., SYLAK-GLASSMAN J., YAROWSKY D., EISNER J. & HULDEN M. (2016). The sigmorphon 2016 shared task—morphological reinflection. In *Proceedings of the 14th Annual SIGMORPHON Workshop on Computational Research in Phonetics, Phonology, and Morphology*, p. 10–22, Berlin, Germany : Association for Computational Linguistics.

DELL F. (1973). *Les règles et les sons*. Collection Savoir. Hermann.

DURRETT G. & DENERO J. (2013). Supervised learning of complete morphological paradigms. In *Proceedings of the 2013 Conference of the North American Chapter of the Association for Computational Linguistics : Human Language Technologies*, p. 1185–1195, Atlanta, Georgia : Association for Computational Linguistics.

FEIST T. & PALANCAR E. L. (2015). Oto-manguean inflectional class database. University of Surrey.

FRISCH S. A., PIERREHUMBERT J. B. & BROE M. B. (2004). Similarity avoidance and the ocp. *Natural Language & Linguistic Theory*, **22**(1), 179–228.

HALLE M. & CLEMENTS G. N. (1983). *Problem Book in Phonology : A Workbook for Introductory Courses in Linguistics and in Modern Phonology*. A Bradford book. A Bradford.

HAYES B. (2012). Spreadsheet with segments and their feature values. Distributed as part of course material for Linguistics 120A : Phonology I at UCLA. These are the features used in FeaturePad.

HOCKETT (1954). two models of grammatical description.

KANN K. & SCHÜTZE H. (2016). Med : The lmu system for the sigmorphon 2016 shared task on morphological reinflection. *ACL 2016*, p.6̃2.

KIROV C., SYLAK-GLASSMAN J., QUE R. & YAROWSKY D. (2016). Very-large scale parsing and normalization of wiktionary morphological paradigms. In N. C. C. CHAIR), K. CHOUKRI, T. DECLERCK, S. GOGGI, M. GROBELNIK, B. MAEGAARD, J. MARIANI, H. MAZO, A. MORENO, J. ODIJK & S. PIPERIDIS, Eds., *Proceedings of the Tenth International Conference on Language Resources and Evaluation (LREC 2016)*, Paris, France : European Language Resources Association (ELRA).

MATTHEWS P. H. (1991). *Morphology*. Cambridge University Press, second edition. Cambridge Books Online.

NICOLAI G., CHERRY C. & KONDRAK G. (2015). Inflection generation as discriminative string transduction. In *Proceedings of the 2015 Conference of the North American Chapter of the Association for Computational Linguistics : Human Language Technologies*, p. 922–931, Denver, Colorado : Association for Computational Linguistics.

ROBINS R. H. (1959). In defence of WP. *Transactions of the Philological Society*, **58**(1), 116–144.

SIMS A. (2010). Probabilistic paradigmatics : Principal parts, predictability and (other) possible pieces of the puzzle. In *14th International Morphology Meeting, Budapest*, p. 13–16.

SOROKIN A. (2016). Using longest common subsequence and character models to predict word forms. *ACL 2016*, p.5̃4.

STUMP G. & FINKEL R. (2013). *Morphological Typology : From Word to Paradigm*. Cambridge Studies in Linguistics. Cambridge University Press.

TAJI D., ESKANDER R., HABASH N. & RAMBOW O. (2016). The columbia university-new york university abu dhabi sigmorphon 2016 morphological reinflection shared task submission. *ACL 2016*, p.7̃1.

VEIGA A., CANDEIAS S. & PERDIGÃO F. (2013). Generating a pronunciation dictionary for european portuguese using a joint-sequence model with embedded stress assignment. *Journal of the Brazilian Computer Society*, **19**(2), 127–134.

WURZEL W. (1989). *Inflectional Morphology and Naturalness*. Studies in Natural Language and Linguistic Theory. Springer.