

Aux confins de la morphologie du français : description et analyse des formations du type *rififi*, *roudoudou* et *raplapla*

Laurence Labrune¹

Ecole des Hautes Etudes en Sciences Sociales & CRLAO (UMR 8563), Paris, France

Abstract. Ce travail vise à dégager les principales caractéristiques des mots du type *rififi* et *raplapla* et à proposer une analyse formelle de leurs mécanismes de formation. Nous identifions quatre modes de formation des *rVCVCV* : conventionnel, onomatopéique, apophonique et acronymique. Nous montrons qu'un mot *rVCVCV* est le résultat de l'association à un gabarit d'une racine issue d'un radical lui-même extrait d'un lemme-base, selon un processus relevant de la morphologie gabaritique non-concaténative. L'étude met également au jour des tendances nettes de préférence ou d'évitement des consonnes et des voyelles qui apparaissent dans ces mots et révèle la présence des phonesthèmes /p/ et /u/, très fréquents dans ces formations, de même que /p/, /k/, /b/, /g/, /f/, /s/, /ð/ et /ẽ/ tandis que d'autres segments, notamment /r/, /v/, /n/, /s/, /y/ et /e-ɛ/ sont nettement évités. La formation des *rVCVCV* peut donc s'analyser, pour ce qui est des aspects segmentaux, comme le résultat d'une morphologisation de paramètres phonologiques.

1 Introduction²

1.1 Objet et cadre général de la recherche

Il existe en français non conventionnel des formations expressives familières ou ludiques telles que *rififi*, *roudoudou*, *ribouiboui* ou *raplapla*, peu recensées par les dictionnaires, et qui n'ont fait l'objet d'aucune étude à notre connaissance malgré leur grande productivité. Cet article vise à combler cette lacune en documentant ces formations pour en dégager les principaux traits caractéristiques et proposer une analyse formelle de leurs mécanismes de formation. Dans ce travail, nous utilisons la formule *rVCVCV* pour désigner ces formes de manière générique. Elles présentent en effet un patron remarquablement uniforme et

¹ Adresse de l'auteur : laurence.labrune@ehess.fr

² Je remercie Fabio Montermini et Hélène Giraudo qui m'ont suggéré des pistes explicatives pour l'absence de la voyelle /y/ dans les *rVCVCV* lors du colloque Ismo tenu à Lille en octobre 2025, ainsi que Camille Monnot pour son calcul des statistiques de fréquence lemmatique des phonèmes du français général à partir des données de la base Lexique 4.00 (<http://www.lexique.org/>), projet sous licence CC-BY-NC (auteurs : Boris New & Christophe Pallier). Je suis également reconnaissante à deux collègues anonymes pour leurs remarques sur une version préliminaire de cet article.

contraint : trisyllabiques, elles sont constituées d'une reduplication d'une syllabe de structure C_1V ou C_1C_2V (C = Consonne ou Glide), elle-même précédée d'une syllabe dont l'attaque est /r/. La voyelle est identique dans les trois syllabes. Quelques exemples sont fournis en (1). En (1)a, la syllabe redupliquée est de forme CV, en (1)b, elle contient une attaque branchante avec une seconde consonne ou un glide.

(1)

a. $\sigma_1rV1. \sigma_2C1V1. \sigma_3C1V1$: *rififi, roudoudou, radada, rototo, rinquinquin, ragnagna*, etc.

b. $\sigma_1rV1. \sigma_2C1C2V1. \sigma_3C1C2V1$: *raplapla, ribouiboui, routioutiou, ronflonflon*, etc.

En (2)a-i, nous présentons quelques exemples en contexte relevés sur la Toile.

(2)

- a. « On a droit au festival des **reuneuneu** dans les commentaires » <https://www.ladepeche.fr/2021/02/23/essonne-ce-que-lon-sait-des-deux-adolescents-morts-apres-des-rixes-entre-bandes-rivales-9390254.php>
- b. « De premier abord, les Verron composent une famille **ringlinglin**, qui n'inspire guère la sympathie » <https://www.ouest-france.fr/pays-de-la-loire/quand-une-bd-lave-son-linge-sale-en-anjou-2979963>
- c. « **Ranlanlan** comment elle t'irait trop Kellian la GMT en or blanc que t'auras pas, le pepsi, le poids, la légère chaleur de l'or blanc » https://forum.hardware.fr/hfr/Discussions/Loisirs/horlogerie-montres-cb-sujet_1750_15265.htm
- d. « Pou était un pou qui s'appelait Pou. Pas Poupou, pas **Roupoupou**, juste Pou. » <https://www.decitre.fr/livres/guerre-et-poux-9782211308359.html>
- e. « Oui, c'est un moment '**rouniouniou**' » https://www.lemonde.fr/m-perso/article/2019/12/22/baby-yoda-une-mascotte-interplanetaire_6023757_4497916.html
- f. « **Rapiapia** du porte monnaie finalement je me passerai du nutz et coca fais le colis en recommandé pour le reste » <https://cattlaelia.forumactif.org/t30634-neofinetialand>
- g. « Hier j'ai fait volé le **ribouiboui** électrique de mon Père est pour une première fois ça donne vraiment envie de passer sur du plus gros » <http://www.pilote-virtuel.com/viewtopic.php?id=39736>
- h. « Si ça vous rend **rongnongnon** de voir Louis La Brocante incarner votre docteur favori, vous faites comme les détracteurs de Ben Affleck concernant Batman, vous passez votre chemin » <https://www.lecinemaestpolitique.fr/quoi-de-neuf-docteur-docteur-who-et-m-moffat/>
- i. « Fan de rugby, le **reukeukeu** n'est pas mort » <https://www.facebook.com/watch/?v=1375051002569380>

Nous considérons que les $rVCVCV$ relèvent d'un mode de formation gabaritique de type *root and pattern*, cas rare en français, mais qu'on retrouve dans les formations hypocoristiques comme *Totor* (< *Victor*) ou *Guiguite* (< *Marguerite*) (Plénat 1984, 1999). Au-delà des aspects proprement gabaritiques et *output oriented* (cf. section 2 ci-dessous), nous montrerons également que ces formations sont soumises à des contraintes segmentales fortes affectant la voyelle et l'attaque des syllabes 2 et 3 (section 3).

1.2 Le corpus

Pour mener cette étude, nous avons constitué un corpus de formations $rVCVCV$ rassemblées à partir d'une collecte manuelle complétée par des recherches systématiques dans les dictionnaires et sur la Toile de tous les patrons possibles (à savoir, toutes les voyelles et

consonnes possibles du français en V, C1 et le cas échéant C2). Il est à noter que ce que nous appelons dans la suite de ce travail « formation » ou « forme » *rVCVCV* peut correspondre à plus d'un lemme-base et/ou porter plus d'une signification. Lorsque nous aurons besoin de désigner strictement un mot *rVCVCV* en tant que signifiant pour un signifié particulier, nous parlons de mot *rVCVCV*. Ainsi, la forme ou formation *ronflonflon* a cinq significations différentes (correspond à cinq mots) : 1) « musique de fanfare » ; 2) « ronflement » ; 3) « ornements exhubérants » ; 4) « discours grandiloquent et convenu » ; 5) « personne ennuyeuse », pas toutes issues de la même base lexicale. Une autre précision utile est que nous avons ignoré les différences graphiques dans les rares cas où il y en a : *riquiqui* et *rikiki*, *rouniouniou* et *rounougrou*, etc. sont considéré comme représentant la même formation.

Nous souhaitons souligner que les données ne sont sans doute pas exhaustives. Outre les lacunes inhérentes à la constitution de ce type de corpus, nous supposons que certaines formes *rVCVCV* tout à fait possibles à l'oral n'ont pas été relevées à cause d'incongruités graphiques qui les pénalisent dans les supports écrits. Ainsi, la forme /rɛwɛwɛ/ nous semble parfaitement acceptable, mais sa graphie qui pourrait être *rinouinouin* fait qu'elle a peu de chances d'être employée à l'écrit, et encore moins choisie comme pseudonyme sur les réseaux sociaux par exemple. C'est aussi peut-être cette raison graphique qui explique l'absence totale de la semi-consonne /ɥ/ dans les formes relevées, même si cette absence peut aussi résulter de la très faible fréquence générale de ce phonème en français. Comment transcrire /riɥiɥi/ par exemple ? *Riuiui* a peu de chance de passer le test graphique.

Quoi qu'il en soit, on peut estimer que le nombre de formes *rVCVCV* relevées, qui est de 103, pour 159 signifiés différents, constitue un échantillon suffisant et représentatif pour une étude générale des principes régissant ce mode de formation.

1.3 Remarques grammaticales, lexicales, sémantiques et pragmatiques générales

Bien que le focus de cet article soit la morpho-phonologie des mots *rVCVCV*, quelques remarques générales sur les aspects autres que morphologiques et phonologiques seront utiles. Les *rVCVCV* sont majoritairement des noms communs (*ragnagna*, *rototo*) ou des noms propres. Les noms propres, une sous-catégorie très productive, désignent des noms de marques (*rinzinzin*, entreprise de développement d'applications téléphoniques) ou de gammes de produits (*ronbonbon*, un modèle de bonbon, *rancancan*, un modèle de robe), des pseudonymes sur les réseaux sociaux (*Raclacla*, pseudo de Clara, *Ricricri*, de Christelle), des noms de personnages de livres pour enfants (*Ranpanpan*, *Roupoupou*) ou d'animaux domestiques ou d'élevage (*Rilili des Aigrettes*, nom d'un cheval de course, *Rouzouzou*, nom d'un chat). On trouve aussi quelques adverbes (par exemple *voler en* (ou *au*) *radada*) et beaucoup d'adjectifs (*riquiqui*, *raplapla*, *rondondon*), en principe invariables.

Il arrive que certains *rVCVCV* soient morphologiquement bien intégrés dans le lexique, à l'instar de *rococo*, qui donne *rococoterie*, *rococotier*, *rococoter* et qui a aussi une forme féminine *rococote* (TLFi). Ces cas sont cependant très rares. À l'exception d'un <s> ou <x> graphique marquant le pluriel dans certains exemples écrits, les variations morphologiques de genre et les formes dérivées sont peu fréquentes (précisons cependant que nous n'avons pas mené de recherche systématique sur les dérivés possibles). Concernant les aspects sémantiques et pragmatiques, il faut souligner la dimension expressive, évaluative et ludique très marquée de ces formations qui relèvent le plus souvent (mais pas toujours) d'un registre familier. Leur productivité, entendue ici comme la disponibilité du patron dans la langue et le fait que les locuteurs sont capables de le manipuler pour créer facilement de nouveaux mots *rVCVCV*, est très élevée. Ceci explique que l'on trouve énormément d'occasionalismes (dans le sens de Dal 2003 et Dal et Namer 2018) comme l'attestent la plupart des exemples en (2), choisis à dessein. Enfin, sur le plan lexico-sémantique, nous observons que les

polysémies sont fréquentes. Par exemple, la forme *ronflonflon*, déjà citée, n'a pas moins de cinq sens distincts, de même que *riquiri*. Nous reviendrons en 4.4 sur la raison de ce phénomène. La polyonymie est aussi répandue. Ainsi le référent 'seins' compte cinq exponents différents dans notre corpus (*rinplinplin*, *rondondon*, *rototo*, *roploplo*, *ronplonplon*) et le référent 'menstruations' en a quatre (*ragnagna*, *reugneugneu*, *rangnangnan*, *ringningnin*).

Dans la suite de cet article, nous nous intéresserons d'abord aux principes gabaritiques et structurels qui régissent la formation des *rVCVCV* (section 2). Nous traiterons ensuite les aspects mélodiques, à savoir les préférences ou évitement de certains phonèmes observables dans les dérivés (section 3). L'on verra que des résultats inattendus se dégagent. Enfin, nous aborderons quelques questions connexes (section 4) avant de conclure (section 5).

2 Aspects gabaritiques: une morpho-phonologie *output oriented*

Examinons pour commencer les aspects prosodiques et gabaritiques des formations *rVCVCV*, qui sont les plus caractéristiques de ces formes, et leur « marque de fabrique » la plus saillante. Auparavant, il est nécessaire de présenter les principaux modes de formation des *rVCVCV* du point de vue morpho-lexical.

2.1 Principaux modes de formation des *rVCVCV* : conventionnel, onomatopéique, apophonique, acronymique

Nous avons identifié quatre grands modes de formation des mots *rVCVCV*. Le mode 1, le plus fourni, concerne les *rVCVCV* dérivés à partir d'une base du lexique conventionnel, comme par exemple *rototo* < *rot*, *raplapla* < *plat*, *ronchonchon* < *ronchon*, *roudioudiou* < *boudiou* (« bon dieu »), *ronflonflon* (dans le sens « sommeil ») à partir de la base *ronfler*, etc. Le mode 2 comprend les *rVCVCV* formés depuis une base onomatopéique, incluant des refrains de chansons et d'autres formations similaires, par exemple *reupeupeu* (bruit d'un moteur qui démarre avec peine), à partir de *reup reup*, *rinchinchin* (« instrument de musique à manivelle »), *rintrintrin* (refrain de chanson), *rouclouclou* (roucoulement), *routoutou* (canard nomonyx dominicus, à partir de son cri ?), *raflafla* (roulement de caquarinettes, un genre de castagnettes), à partir de l'onomatopée *fla fla*, *ronflonflon* (dans le sens « musique de fanfare ») à partir de la base *flonflon*, *rohoho* < *ho ho* (interjection)... Le mode 3 est constitué de mots dérivés secondairement par apophonie à partir d'un *rVCVCV* existant. A partir de *ragnagna*, on aura ainsi *reugneugneu*, *rangnangnan*, *ringningnin* ; à partir de *ranplanplan* (« ennuyeux »), qui relève du type 2, on a *ronplonplon*, même sens ; à partir de *roploplo* (« seins »), on a formé *ronplonplon* et *rinplinplin*, etc. On peut inclure dans cette catégorie la série *rohoho*, *rihihi*, *rahaha* (noms de personnages fictifs), sans qu'il soit possible de déterminer lequel est à l'origine des autres, mais c'est finalement sans grande importance. Ce qui compte, c'est le fait qu'il existe des micro-séries basées sur une alternance vocalique avec préservation du sens global du *rVCVCV* malgré le changement vocalique. Plus marginalement, nous avons identifié également un mode 4, constitué de *rVCVCV* issus de l'oralisation plus ou moins approximative d'un acronyme commençant par la lettre <r> ou par Voyelle + <r> : *reuneuneu*, pour RN (Rassemblement National ; il y a aussi dans l'exemple une interférence probable avec le mot *neuneu*), *reuneuneu* à partir de ARN messenger, *reuteuteu*, pour RTT, et *reuneumeu* pour RMM (« rémunération mensuelle minimale »). Nous estimons que l'oralisation de ces sigles a été facilitée par l'attraction du patron *rVCVCV* et les connotations qui y sont associées. Dans les quatre cas, une connotation péjorative est nettement perceptible dans le ton et le contexte des exemples.

Nous avons enfin dans le corpus un grand nombre de *rVCVCV* dont nous n'avons pu établir avec certitude le mode de formation, faute d'avoir réussi à en identifier la base.

2.2 Caractère gabaritique des formations *rVCVCV*

Les formations *rVCVCV* relèvent d'une morpho-phonologie orientée vers l'output, c'est-à-dire que la bonne formation d'un mot *rVCVCV* dépend de sa forme de surface, qui doit se conformer au patron *rVCVCV*, mais aussi du rapport que celle-ci entretient avec sa base. Ce rapport, cependant, n'est pas toujours capturable d'une manière identique dans tous les *rVCVCV*, et c'est ce qui définit le caractère gabaritique et non-concaténatif de la formation. Autrement dit, il n'est pas possible de formuler un scénario dérivationnel linéaire des formes *rVCVCV*. Comme le montrent les exemples en (3), la partie du *rVCVCV* qui correspond au mot de départ n'est pas toujours la même. (Nous indiquons ici en caractères gras la portion du *rVCVCV* en correspondance avec la base, donnée après le symbole <.)

(3)

- a) **Rififi** < rif
- b) **Ronchonchon** < ronchon
- c) **Rototo** < rot
- d) **Raplapla** < plat
- e) **Reuneuneu** < neuneu
- f) **Reudeudeu** (« personnes de la noblesse ») < de
- g) **Rococo** ; **rococo** > rocaille, roc ; barocco ?
- h) **Rongnongnon** < grognon
- i) **Ronflonflon** (sens 1) < ronfler
- j) **Ronflonflon** (sens 2) < flonflon
- k) **Radada** (*voler en radada*, terme d'aéronautique) < ras du sol, ras des pâquerettes

Le matériel commun entre le lemme de départ et le dérivé n'est pas toujours ni de la même taille, ni de même statut ou position selon les *rVCVCV*. Dans certains cas, le lemme-base est entièrement repris dans le dérivé (ex. a, b, c, d, e, f, j), dans d'autres cas, il est tronqué (ex. g, h, i, k). Le *r* initial du dérivé est parfois déjà présent dans le lemme-base (ex. a, b, c, g, i, k), parfois non (ex. d, e, f, h, j). Le matériel repris dans le dérivé correspond parfois à une seule syllabe, ouverte ou fermée (ex. a, c, d, f, h, i), ou bien à deux syllabes ouvertes (b, e, j), et cette / ces syllabe(s) ne sont pas forcément les syllabes initiales du lemme-base (ex. h). Enfin, l'exemple (3)c nous montre qu'une consonne latente peut être récupérée pour servir de consonne fixe réalisée dans le dérivé quand la seule consonne disponible dans le radical serait /r/. *Rototo* est le seul exemple de ce type, mais on doit mentionner le cas assez proche de *radada*, à partir de « ras du sol » dans l'un de ses trois sens, dont le radical commence aussi par /r/, ce qui empêche que cette consonne soit récupérée dans le dérivé. La stratégie consiste alors à utiliser la consonne (non-latente) du mot qui suit au début de la préposition *de*.

2.3 Catégories opératoires pour l'analyse

Nous inspirant des travaux de Roché (2010), nous adoptons la terminologie en (4) pour décrire les objets morphologiques identifiés dans le processus de formation des *rVCVCV*, à savoir le lemme-base, le dérivé et le radical. Les termes de racine et de gabarit sont tirés de la terminologie classique employée pour l'analyse des langues gabaritiques.

(4)

- **Lemme, base, ou lemme-base** : le lemme de départ existant dans le lexique au sens large (incluant des noms propres) à partir duquel est dérivée la formation *rVCVCV*.
- **Dérivé** : la forme *rVCVCV* de surface (que nous appellerons aussi parfois **output**).
- **Radical** : « la chaîne segmentale correspondant à la base dans le mot construit » (Roché 2010). Il s'agit du matériel mélodique, constitué de deux ou trois phonèmes,

qui se trouve en correspondance avec le lemme-base de manière linéaire. Le radical est noté en petite capitales.

- **Racine** : mélodie extraite du radical qui est associée au gabarit des formations *rVCVCV* ; la mélodie est constituée de deux éléments phonologiques, l'un consonantique qui a vocation à être associé à une attaque de syllabe, l'autre vocalique, associé à un noyau de syllabe. Il faut noter que l'élément consonantique peut être complexe car constitué de deux consonnes mais qu'il fonctionne comme un constituant phonologique unique. La racine sera donnée entre accolades.
- **Gabarit, schème, patron** : nous le représentons à l'aide de la formule *rNANAN* (cf. Plénat 1984)), où *r* = /r/ (c'est une mélodie fixe), N = noyau syllabique, A = attaque syllabique.

Comme nous le verrons, il est indispensable de distinguer le radical de la racine, car dans le dérivé, le radical ne se retrouve pas toujours dans son entièreté, ni avec le même ordre linéaire. Dans notre analyse, la racine est extraite du radical, et non directement du lemme-base. La racine n'est en correspondance avec le lemme-base que par le truchement du radical.

2.4 Analyse formelle

Le lemme-base a pu être identifié dans bon nombre de dérivés *rVCVCV* mais pas dans tous. Les 103 formes *rVCVCV* que compte notre corpus correspondent à 159 signifiés distincts (comme nous l'avons dit ci-dessus, les homonymies sont nombreuses, d'où ce ratio élevé). Nous avons identifié avec certitude les lemmes-bases de 59 *rVCVCV*. Pour une poignée de *rVCVCV*, nous avons une hypothèse plus ou moins solide. Assez souvent, malheureusement, le lemme-base n'est pas identifiable. C'est le cas pour *riquiriqui* ou *rikiki* (qui a cinq sens différents), de *ringlinglin* « vin de pays aigrelet », de *rouniouniou* « mignon », de nombre de pseudos internet, et de bien d'autres mots. Dans d'autres cas, plusieurs lemmes-bases sont envisageables, voire, existent en parallèle en double motivation (ainsi, *roc*, *rocaille* et *barocco* sont tous trois signalés comme possibles bases de *rococo* par le TLFi).

Tableau 1. Quelques exemples de lemme-base, dérivé, radical et racine du patron *rVCVCV*

Lemme-base	Dérivé	Radical	Racine
<i>rif</i>	rififi	RIF	{f, i}
<i>rot</i>	rototo	ROT	{t, o}
<i>ras du sol, ras des pâquerettes</i>	radada	RAD	{d, a}
<i>plat</i>	raplapla	PLA	{pl, a}
<i>flonflon</i>	Ronflonflon (sens 1)	FLONFLON	{fl, õ}
<i>ronfler</i>	Ronflonflon (sens 2)	RONFL	{fl, õ}
<i>ronchon</i>	ronchonchon	RONCHON	{ʃ, õ}
<i>roc, rocaille</i>	rococo	ROC	{k, o}
<i>barocco</i>	rococo	ROCO	{k, o}
<i>chou ou chouchou</i>	rouchouchou	CHOU ou CHOUCOU	{ʃ, u}
<i>oh la la</i>	ralala	LALA	{l, a}
<i>ho ho</i>	rohoho	HOHO	{h, o}
<i>RN</i>	reuneuneu	RØN	{n, ø}

Comme nous l'avons dit en 2.2, le rapport entre un dérivé *rVCVCV* et son lemme-base n'est pas capturable en une suite d'opérations ou de règles séquentielles ordonnées à partir d'un radical. Autrement dit, les radicaux des *rV₁C₁(C₂)V₁C₁(C₂)V₁* sont en correspondance (dans

le sens technique qu'a ce terme en Théorie de l'Optimalité, cf. par exemple Kager 1999) avec différentes portions de leur base.

Le Tableau 1 de la page précédente présente quelques exemples qui illustrent plus clairement la différence entre radical et racine. (NB : Le mode 3 n'est pas directement concerné ici, puisqu'il est dérivé secondairement d'un *rVCVCV* déjà créé.)

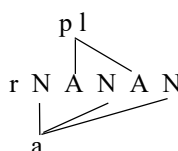
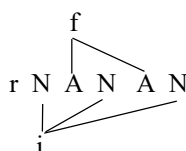
On note qu'un cluster consonantique présent dans le radical est conservé dans le dérivé (*plat* → *raplapla*), ce qui montre la nécessité de recourir à un modèle centré sur la syllabe : ce sont des constituants de syllabes qu'on manipule, et non des suites de phonèmes. Quand le radical commence par /r/, ce /r/ n'est jamais repris dans la racine, pour éviter un dérivé qui répèterait une unique consonne. Dans ce cas, c'est une consonne du radical qui est utilisée.

Voici les grandes lignes de notre analyse du mécanisme de formation des *rVCVCV*, que nous formaliserons dans le cadre de la phonologie auto-segmentale (Goldsmith 1979) : le *r* initial des *rVCVCV* est une mélodie fixe, fournie par le patron. Le matériel mélodique représenté dans les deux éléments de la racine est associé aux positions N et A du gabarit *rNANAN*, la voyelle aux N, l'élément consonantique (qui peut être double) aux A. Ce mécanisme est représenté en (5)a et (5)b à travers deux exemples : *rififi* et *raplapla*.

(5)

a. *rif* → *rififi* : {f, i}

b. *plat* → *raplapla* : {pl, a}



3 Aspects mélodiques : les préférences segmentales

Nous poursuivons l'investigation des principes qui régissent la formation des *rVCVCV* en nous intéressant aux aspects segmentaux, c'est-à-dire les consonnes et les voyelles figurant dans l'une des positions V et C des formes *rVCVCV*. Nous verrons que des principes assez robustes, et parfois inattendus, conditionnent les possibilités d'apparition des segments dans les *rVCVCV*. Le trait le plus remarquable est que certaines consonnes ou voyelles sont nettement favorisées, tandis que d'autres sont évitées.

3.1 Les attaques

Nous nous intéressons d'abord aux consonnes figurant en attaque de la deuxième et troisième syllabe des formes *rVCVCV*³. Nous distinguons dans un premier temps les attaques simples avec le patron *rVC1VC1V* (type *rififi*, en Fig. 1) des attaques complexes avec le patron *rVC1C2VC1C2V* (type *raplapla*, en Fig. 2).

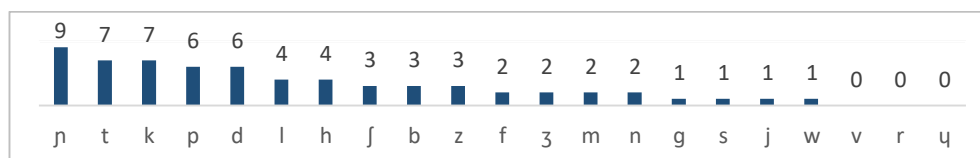


Fig.1 Attaques simples : classement par fréquence⁴ (total = 64)

³ La théorie syllabique adoptée pour nos analyses est celle, tout à fait classique, où Syllabe = Attaque + Rime, avec la Rime elle-même subdivisée en Noyau et Coda (cf. Labrune 2005).

⁴ Le mot orthographié <rouniouniou> a été classé comme contenant la consonne simple /ɲ/.

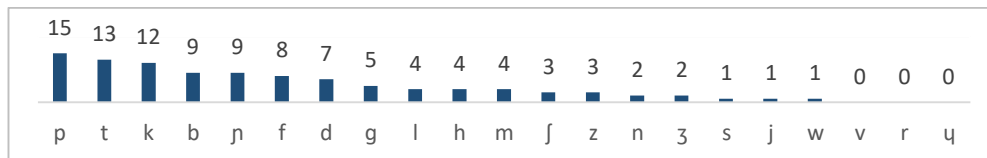


Fig. 2 Attaques simples et complexes : classement par fréquence de C1 (total = 103)

Le point le plus remarquable de la Fig. 1 est la surfréquence de /p/, pourtant la consonne la plus rare du français (avec /ɲ/, ce dernier n'apparaît jamais en attaque de syllabe), et la totale absence de /r/, pourtant la consonne la plus fréquente du français, quel que soit le type de comptage retenu (voir ci-dessous). La fréquence de /p/ résulte selon nous du statut phonostylistique de cette consonne en français. L'absence de /r/ s'explique par la nécessité d'éviter que la consonne initiale du mot ne soit répétée, ce qui créerait des tripliques (*rororo, à partir de *rot*) et brouillerait le schéma.

On notera également la présence d'une attaque que nous notons /h/ (c'est la forme graphique qu'elle reçoit dans nos données, par ex. *rahaha*, *rohoho*), qui correspond phonétiquement à la présence d'un hiatus ou d'un léger coup de glotte. Cette présence est assez inattendue étant donné le faible nombre d'hiatus en position interne de mot en français.

De manière générale, on constate, concernant les attaques simples (Fig. 1) que les préférences suivent à peu près l'échelle de sonorité (du moins au plus sonore), sauf pour /f/, /l/ et /h/, placées un peu plus haut qu'attendu si l'on se réfère à cette échelle, et /g/ qui attire ici l'attention par sa faible présence, comparée aux consonnes qui lui sont phonologiquement proches /k/, /d/ et /b/, de même que /v/ totalement absent. L'hypothèse globale qui permet d'expliquer ces tendances fréquentielles est qu'il s'agit de rechercher un contraste phonologique maximum avec le /r/ initial, d'où la fréquence de /t/, /p/ et /k/, et que la consonne /p/ est favorisée en raison de la valeur impressive et affective des formations *rVCVCV*. Il n'est pas anodin non plus de relever que les trois occlusives sourdes /p, t, k/ en haut de la hiérarchie en Fig. 1 sont, d'une part, les sons les plus bas dans l'échelle de sonorité, et aussi les trois consonnes parmi les plus fondamentales et les plus répandues dans les langues du monde. Nous verrons plus loin qu'une tendance similaire à privilégier des segments universellement basiques se retrouve avec les voyelles. Par contraste, la partie droite de la Fig. 1 ne comporte que des consonnes fricatives et des sonantes, c'est-à-dire des consonnes qui partagent au moins les traits de frication [+continu] ou de sonorance [+sonant] avec /r/, à l'exception de /g/ (nous verrons plus loin la raison expliquant la présence inattendue de /g/ en bas de la hiérarchie, dont la position sera d'ailleurs révisée).

Si l'on observe à présent les résultats de la Fig. 2 qui agrège toutes les consonnes d'attaque, que celle-ci soit simple ou complexe, on constate que les mêmes tendances se retrouvent. Les consonnes aptes à être suivies d'une liquide (les *mutae cum liquida*) ou d'un glide gagnent naturellement quelques points, mais les préférences générales restent globalement les mêmes une fois la pondération des *mutae cum liquida* effectuée.

Nous fournissons également en Fig. 3 et 4 les données concernant les deuxième membres des attaques complexes. On constate que /l/ et /j/ sont les deux éléments préférés dans cette position, au détriment de /r/, /w/ et /ɥ/. Les deux derniers segments sont de toute façon rares dans le corpus, et plus généralement en français général, donc seule la sous-fréquence de /r/ retient l'attention. Cette sous-fréquence s'explique par les mêmes raisons que celle vues ci-dessus pour /r/ en attaque simple, mais ces raisons s'imposent avec une moindre force quand une autre consonne accompagne le /r/.

Si l'on agrège les résultats des attaques palatalisées représentées par *Cj* (12 occurrences) avec ceux de la nasale palatalisée /ɲ/ (9 occurrences), on obtient un nombre de 19 attaques palatalisées, un nombre particulièrement élevé, mais pas surprenant s'agissant de formations

expressives. Car, comme le signale Fónagy (1991), la palatalisation est la marque, dans de nombreuses langues, d'une connotation enfantine et affective.

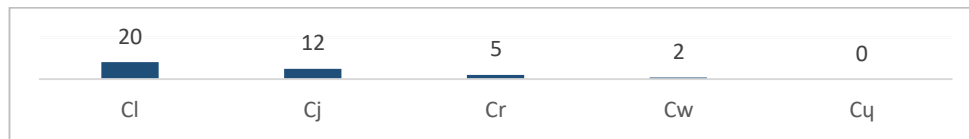


Fig. 3 Attaques complexes: classement par fréquence de C2 (total = 39)

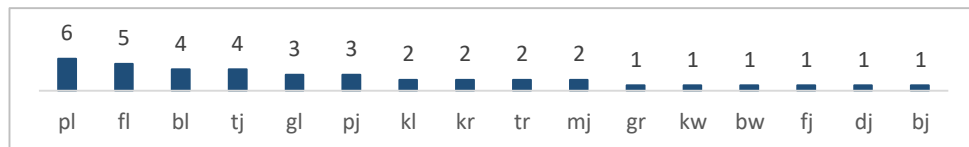


Fig. 4 Attaques complexes: classement général par fréquence (total = 39)

Afin de permettre une appréciation plus nuancée des résultats segmentaux présentés dans les Fig. 1 à 4, il est nécessaire de les comparer aux données générales des consonnes en français. Comme tous les phonèmes n'ont pas la même fréquence, il se pourrait que les préférences mises au jour ne soient que le reflet de préférences générales au niveau de la langue et ne soient pas propres aux formations *rVCVCV*. Pour ce que nous appelons ici « français général », nous proposons la comparaison avec deux types de données : d'une part, des statistiques par occurrence fournies par Wioland (1972), qui compile plusieurs études concernant les fréquences textuelles ou en discours, d'autre part, des statistiques lemmatiques, c'est-à-dire dans les lemmes (sans doublons) extraits de la base de données lexicale Lexique3, qui nous ont été aimablement fournies par Camille Monnot. Il faut noter que ces données du français général que nous utilisons ici concernent les fréquences des consonnes en toutes positions au sein des syllabes (attaque et coda), tandis que les fréquences dans le corpus des *rVCVCV* ne concernent que la position d'attaque. Il est probable que certaines consonnes présentent des fréquences différentes en fonction de leur position syllabique mais nous n'avons pas les moyens d'opérer cette distinction à ce stade de l'étude pour le français général.

La Fig. 5 présente les résultats généraux ramenés en pourcentages pour les trois populations, celle des formes *rVCVCV* (types *rififi* et *raplapla*) et celles du français général (FrG), d'une part en fréquence dans le discours, d'autre part en fréquence lemmatique.⁵

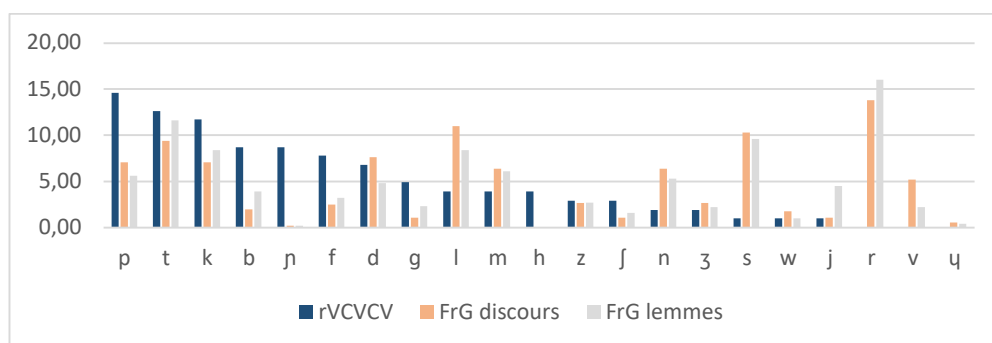


Fig. 5 Comparaison des fréquences de consonnes des syllabes 2 et 3 dans le corpus *rVCVCV* et des consonnes dans le français général en discours et dans les lemmes, exprimées en pourcentage. Pour le corpus *rVCVCV*, les deuxièmes segments des attaques branchantes ne sont pas pris en compte.

⁵ /h/ ne figure pas dans les données du français général dont nous disposons.

Des analyses statistiques plus poussées seraient évidemment nécessaires pour une appréciation plus sûre et plus fine de ces résultats, mais les différences importantes dans les fréquences de certaines des consonnes permettent des commentaires sans grand risque d'erreur. Nous constatons que les données de la Fig. 5 confirment nos observations précédentes tout en permettant de les affiner : les consonnes qui sur-performent nettement sont effectivement /p/, les occlusives /p/, /k/, /b/, /g/, ainsi que /f/ et /ʃ/. Les résultats ne sont pas nets pour /t/ et /d/, une analyse statistique plus fine serait nécessaire. L'occlusive vélaire sonore /g/ qui apparaissait comme faiblement recherchée dans les *rVCVCV* se révèle en réalité une consonne plutôt favorisée si on tient compte de sa fréquence relativement basse en français général. Les coronales, sauf /ʃ/ (mais /ʃ/ est-il vraiment coronal ?) et peut-être /t/ ne semblent finalement pas aussi appréciées que les consonnes d'autres points d'articulation. Ceci est peut-être l'indice de ce que le patron *rVCVCV* remonte à une époque où le /r/ du français était encore majoritairement apical et que la recherche d'un contraste maximum avec le /r/ initial concernait aussi le point d'articulation. Ce pourrait être aussi une manifestation du statut phonologique spécial des coronales comme consonnes non-marquées ou sous-spécifiées (Paradis & Prunet 1991).

Les consonnes qui sous-performent nettement sont /r/, /v/, /s/, /n/, /m/ et /l/. Les consonnes que l'on peut qualifier de neutres, c'est-à-dire qui ne présentent apparemment pas de biais fréquentiel par rapport à leur fréquence dans le français général sont /j/, /w/, /z/, /d/, peut-être /t/, et sans doute /q/ dont la fréquence dans le lexique général est cependant trop faible pour que sa relative rareté dans les *rVCVCV* soit interprétable.

Pour résumer, la nasale /ɲ/ et la rhotique /r/ confirment leur statut particulier dans les *rVCVCV*, la première comme consonne nettement recherchée, la seconde, comme consonne nettement évitée. La tendance globale qui est de privilégier les consonnes qui contrastent le mieux sur le plan phonétique et phonologique avec le /r/ initial parce qu'elles sont positionnées en bas de la hiérarchie de sonorité et/ou ne partageant pas les traits phonologiques de voisement, de sonorance et de frication, est confirmée sauf pour /t/ et /d/, pour lesquels une analyse statistique plus fine serait nécessaire. La très faible fréquence de /s/, comparée à celle de /f/ et /ʃ/ qui présentent pourtant les mêmes traits fricatif et sourd qu'elle, interroge. Elle est sans doute à mettre au compte de la coronalité de /s/.

3.2 Les noyaux

Nous examinons ici les voyelles présentes dans les formations *rVCVCV*. Comme nous l'avons vu, les *rVCVCV* ne comportent qu'une seule voyelle, identique dans les trois syllabes. Quelques précisions méthodologiques : pour l'analyse des voyelles, nous avons neutralisé les oppositions entre /a/ et /ɑ/ et entre les voyelles mi-hautes et mi-basses. Le /ə/ est agrégé à /æ/ et /ø/. Nous n'avons pas tenu compte, non plus, de la voyelle /œ/, d'ailleurs absente dans nos données. Ceci nous donne un système de 10 valeurs vocaliques en noyau. Les fréquences brutes dans les *rVCVCV* sont données dans la Fig. 6.

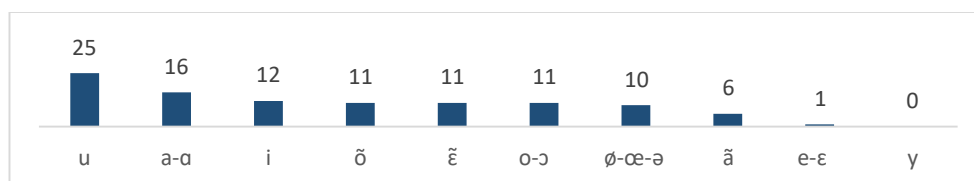


Fig. 6 Classement par fréquence des noyaux dans le corpus *rVCVCV* (total 103)

Le trio de tête, /u/, /a/ et /i/, correspond au triangle vocalique supposé quasi-universel, et fait en quelque sorte pendant au trio consonantique /p, t, k/ discuté plus haut, mais nous n'en

tirerons pas de conclusion hâtive, car nous verrons que ces résultats seront nuancés par la comparaison avec le français général. On note surtout, dans la Fig. 6, la surfréquence de /u/ – dans quasiment un *rVCVCV* sur quatre – ainsi que l’absence totale de /y/, inattendue. La faiblesse de /e-ɛ/ est également surprenante. Il convient, ici encore, de comparer ces fréquences brutes avec celles du français général, en discours et en lemmes (Fig. 7).

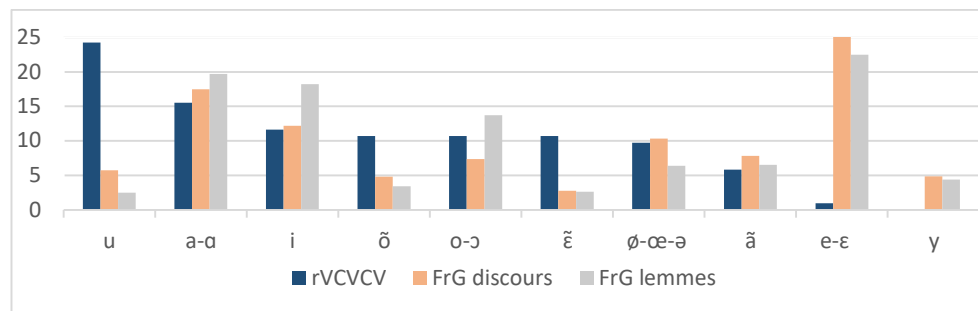


Fig. 7 Comparaison des fréquences de voyelles dans le corpus *rVCVCV* et dans le français général, en discours et dans les lemmes, exprimées en pourcentage.

On constate à l’examen de la Fig. 7 que les résultats pour /u/, /e-ɛ/ et /y/ sont effectivement nettement discordants par rapport aux fréquences de ces voyelles en français général. Nous pensons que la préférence marquée pour /u/ dans les formations *rVCVCV* provient du statut phonostylistique particulier de cette voyelle en français. Elle possède en effet une fonction impressive (pour reprendre le terme proposé par Léon 1969) qui fait d’elle un phonsthème avec une charge affective marquée associée à une connotation d’affection, de tendresse, de mignonnerie, particulièrement adaptée à certains des emplois observés des *rVCVCV* notamment les pseudonymes, les noms d’animaux, de personnages de livres pour enfants, et certains adjectifs à valeur évaluative. On peut en quelque sorte considérer que /u/ est aux voyelles ce que /p/ est aux consonnes en termes de charge impressive. L’autre point d’intérêt dans la Fig. 7 est la sur-représentation des nasales /ɔ̃/ et /ẽ/ (mais pas de /ă/...). À l’exception de /u/, nous ne voyons pas de raison phonologique à ces préférences ou évitements vocaliques. Notamment, l’absence totale de /y/ et quasi-totale de /e-ɛ/ restent une énigme. Nous esquisserons l’hypothèse, qui nous a été suggérée par Fabio Montermini, d’une difficulté articulatoire accrue de la combinaison /ry/, les sons [ʁ], [ʀ] et [y] étant notoirement parmi les plus difficiles du français au niveau acquisitionnel, qu’il s’agisse d’acquisition chez les natifs ou les non natifs. En outre, suivant une autre suggestion d’Hélène Giraudo, la valeur expressive et familière des *rVCVCV* fait peut-être d’elles des formations perçues comme enfantines, et pour cette raison, on y éviterait les combinaisons phoniques qui posent des difficultés aux enfants⁶. Ces hypothèses n’expliquent cependant pas pourquoi c’est la consonne la plus ardue du français, /r/, qui figure en initiale de la formation *rVCVCV*, ni pourquoi /e-ɛ/, qui n’est pas en soi une voyelle difficile, est évitée. Nous reviendrons sur le statut de la consonne /r/ dans la section 4.2.

La formation des *rVCVCV* peut donc s’analyser, pour ce qui est des aspects segmentaux, comme le résultat d’une morphologisation de paramètres phonologiques. Les *rVCVCV*instancient une morphologie fortement conditionnée par la phonologie.

⁶ Une évaluatrice anonyme observe également que les mots contenant deux /y/ pourraient être rares en français. Nous ne disposons pas de statistiques précises sur cette question, mais on peut noter que dans les mots à valeur évaluative, l’occurrence de plusieurs /y/ ne pose pas problème : *turlututu*, *nunuche*, *tutu*, *cucul*... Dans le lexique général, on trouve *culture*, *tumulte*, *futur*, *multitude*, *truculent*, etc.

4 Remarques complémentaires

4.1 Le statut du *r* initial

Il est sans doute intéressant, dans une perspective plus large et pour une comparaison avec d'autres procédés de formation des mots en morphologie et morpho-phonologie du français, de considérer que les *rVCVCV* sont décomposables en une syllabe de structure AN redoublée précédée d'un affixe constitué de la consonne fixe /r/ suivie d'une copie de la voyelle en N: donc *r-* + *CVx2*.

Pour ce qui est de l'initiale /r/, seul élément fixe du patron et qu'on peut considérer comme un préfixe, nous pensons qu'il a peut-être à voir avec le préfixe dit « de renforcement » *ra-* ou *re-* du français discuté dans Roché (2008), qui cite aussi Nyrop (1936), à ceci près que dans les *rVCVCV* la voyelle n'est pas fixe mais recopiée depuis la racine. Un paramètre ayant pu favoriser le choix de /r/ pourrait aussi venir de son statut hors système dans le système phonologique du français. Nous pourrions ajouter également à la liste des particularités concernant ce phonème qu'il présente une tendance certaine à être évité à l'initiale des hypocoristiques selon Plénat (1999). Il faudrait voir plus en détail comment cet ensemble de propriétés fait sens par rapport aux *rVCVCV*.

4.2 Y-a-t-il un déterminisme au niveau des lemmes-bases ?

Il est intéressant d'observer que les lemmes-bases qui commencent par /r/ semblent privilégiés, de même que ceux comportant une reduplication. Même si les données concernant les lemmes-bases ne sont certaines que pour 59 mots de notre corpus, nous relevons que 21 d'entre eux commencent par *r* ou par Voyelle + *r* (ex. *rif*, *ronchon*, *arrache*, *RN...*). On peut supposer que les bases offrant potentiellement une bonne correspondance entre input et output (dans un sens OT) parce que la base commence déjà par *r* semblent plus enclines à subir la dérivation. On aurait donc, en quelque sorte, un déterminisme segmental au niveau de la forme phonologique de la base. Ce type de déterminisme segmental a été identifié pour d'autres langues, notamment le japonais (Labrune 2006).

4.3 Remarques phonologiques

Que nous apprennent les *rVCVCV* sur le plan strictement phonologique ? Tout d'abord, que dans les formations comme *raplapla* ou *Ricricri* (pseudo de Christelle), les attaques dites branchantes se comportent comme des attaques simples, ce qui va dans le sens de l'analyse proposée par Lowenstamm (2003). Au contraire, les suites *sC*, quant à elles, ne fonctionnent pas comme un constituant unique puisqu'on n'a aucun *rVCVCV* de la forme *rVsCVsCV* (non seulement **riskiski* n'existe pas, mais il le pourrait difficilement car il est très mal formé et presque imprononçable pour un natif). Ceci confirme l'analyse selon laquelle les séquences *sC* sont hétérosyllabiques en français (cf. Labrune 2024). Les glides (G) /w/, /j/ et /ɥ/, quant à eux, sont à considérer comme faisant partie de l'attaque et non du noyau, puisqu'on n'a aucun *rVCVCV* de forme *rGVCGVCGV* (**rouibouiboui*) alors que des formes telles que *ribouiboui* ou *rapiapia* sont courantes. Enfin, il est intéressant de noter que les formations *rVCVCV* permettent les hiatus internes au mot. Nous avons trouvé quatre exemples de ce type : *rohoho*, *rahaha*, *rihihi* et *rouhouhou*. C'est un cas rarissime en français conventionnel quand aucune frontière morphologique n'est présente.

4.4 L'homonymie

L'homonymie fréquente au sein du corpus *rVCVCV*, évoquée plus haut, s'explique par le fait que la racine (dans le sens technique du terme adopté ici) ne comporte que deux éléments. Les possibilités sont donc limitées, car non seulement des bases différentes peuvent fréquemment produire une racine identique, mais le gabarit étant fortement contraint, le nombre de formes existantes issues du patron *rVCVCV* est en nombre fini. Les évitements de certains segments font que les possibilités sont encore moindres que si l'on avait une répartition par le hasard des phonèmes. C'est cet ensemble de contraintes qui explique les nombreuses homophonies. Ceci étant, lors de notre recherche systématique, nous avons constaté des lacunes : nous n'avons pas trouvé de *rVCVCV* tels que **richichi* ou **roujoujou*, mais nous estimons que ces formes peuvent exister car elles remplissent les conditions et pourraient venir facilement occuper des cases encore vides de la matrice.

4.5 Scrogneugneu, zigouigoui, ratatouille et les autres

Enfin, il faut mentionner l'existence de mots tels que *scrogneugneu*, *zigouigoui*, *tralala*, *zigligli*, *aglagla*, *taratata*, *ratatouille*, *ratatiner*, *perlimpimpin*, *patatras*, *radadame*, *rododome* (onomatopées), *rititine* (petite mésange des bois), *rondoudou* (nom d'un Pokémon), *Saint-Glinglin*, *Radadou* (nom d'une poupée en vogue au début du XXe siècle, fils de *Rintintin*) et sans doute plusieurs autres encore. Nous pensons que ces mots doivent être analysés à l'aune de la formation *rVCVCV* car ils partagent plusieurs propriétés avec celle-ci. En effet, ils ont en commun de posséder une reduplication interne, de ne pas contenir de *r* interne en tant qu'attaque simple de syllabe, et de ne présenter aucun des phonèmes évités dans les *rVCVCV* à savoir /v/, /y/, /e-ε/. La différence avec le patron *rVCVCV* est que la forme ne commence pas par /r/ (*zigouigoui*, *Saint-Glinglin*, *aglagla*, *tralala*, *zigligli*), que la voyelle la plus à gauche du mot n'est pas forcément une copie de celle de la partie redupliquée (*scrogneugneu*, *rogneugneu*, *rondoudou*), qu'il y a un autre préfixe avant la partie *rVCVCV* (*taratata*), que la reduplication interne est imparfaite (*patatras*), qu'il y a un suffixe (*ratatouille*, *ratatiner*, *radadame*, *rododome*, *rititine*, *radadou*) ou bien une combinaison de deux de ces propriétés (*perlimpimpin*, et certains des exemples déjà cités). Pour un petit nombre de ces mots, les dictionnaires (TLFi ou FEW), les travaux de linguistes (Roché 2008 sur *ratatouille*), ou d'autres sources telles que Wikipedia (*Saint-Glinglin*) recensent des étymologies plus ou moins convaincantes, mais il semble indéniable que l'existence du patron *rVCVCV* a contribué à la création et à la bonne implantation linguistique de ces formes.

5 Pour conclure : bilan, limites et ouvertures

Ce travail nous a permis d'explorer une dimension peu (re)connue de la morphologie du français qui, à la différence d'autres phénomènes de la morphologie dite non-conventionnelle comme les troncations, les hypocoristiques, les langages secrets, les onomatopées ou les mots valises, n'avait jamais fait l'objet d'une étude détaillée. On retrouve assez naturellement, dans les *rVCVCV*, des propriétés fréquemment observées dans ces autres types de formations : instanciation de gabarits, copie et propagation de matériel segmental, relation particulière de correspondance entre la base et le dérivé, reduplication, préférence pour certains segments, usage de phonesthèmes, valeur affective, expressive et évaluative, etc.

Le principal résultat de ce travail, basé sur un corpus de formes comprenant 103 *rVCVCV*, est qu'il existe en français contemporain un patron *rVCVCV* productif (dans le sens que nous donnons à ce terme, voir la section 1.3) et générateur de nombreux occasionalismes. Suivant les termes de Guiraud (1986), on peut parler d'une matrice

dynamique et générative, par l'existence de laquelle de nouveaux mots sont précipités dans les cases vides. Il y a un invariant commun, le patron *rVCVCV* (ou macro-signe), et une variable, la racine {A, N}, autrement dit un dispositif de type *root and pattern*. Pour rendre compte du mécanisme de formation des *rVCVCV*, nous avons proposé une analyse dans le cadre auto-segmental et qui recourt aux concepts opératoires de lemme-base, radical, racine, et gabarit *rNANAN*. Pour synthétiser, nous dirons qu'un mot *rVCVCV* est le résultat de l'association à un gabarit *rNANAN* d'une racine issue d'un radical lui-même extrait d'un lemme-base, selon un processus non-concaténatif rappelant ceux des langues sémitiques. Ce qui importe, c'est le résultat en surface, d'où la qualification du phénomène comme *output-oriented*. C'est en effet la bonne formation morpho-phonologique de l'output et sa conformité au patron qui conditionnent la possibilité de la dérivation à partir d'un lemme-base.

Nous avons mis au jour également des biais assez robustes au niveau des phonèmes présents en surface dans les *rVCVCV*. Ces biais s'expliquent d'une part, pour les consonnes, par la recherche d'un contraste maximal avec le *r* initial, contraste basé sur l'échelle de sonorité et l'évitement des traits phonologiques de sonorance, de continuance et de voisement, d'autre part par le statut phonostylistique marqué en français des phonèmes /p/ et /u/, très fortement présents dans les *rVCVCV* et auxquels on peut attribuer le statut de phonesthèmes. Autrement dit, les consonnes « très consonantiques » sont recherchées, les consonnes « moins consonantiques » évitées (sauf /p/) car elles permettent un contraste optimal avec la rhotique. Les formations *rVCVCV* se présentent donc comme un lieu de morphologisation de paramètres phonologiques et phonostylistiques.

Pour ce qui est des limites de la présente étude, il faut souligner qu'il aurait fallu ne pas se cantonner à des statistiques descriptives mais utiliser des modèles de statistique analytique plus fins pour l'analyse des préférences segmentales. C'est un volet du travail que nous souhaitons développer dans un avenir proche. Comme suggéré par un ou une évaluatrice, il serait nécessaire également de faire les calculs de fréquences phonémiques en sous-catégorisant les *rVCVCV* selon que leur base est identifiée ou non, et selon que leur mode de formation est conventionnel, onomatopéique, apophonique ou acronymique. Faute de temps et d'espace, nous remettons cette entreprise à un travail futur.

Enfin, parmi les observations tirées de l'étude du corpus, il en est pour lesquelles nous n'avons aucune explication à proposer. La principale concerne l'évitement des voyelles /y/ et /e-e/ dans les *rVCVCV*. Nous ne voyons pas de raison qui explique que la voyelle antérieure écartée /i/ soit acceptable mais pas l'autre voyelle antérieure écartée /e-e/ pourtant phonétiquement et phonologiquement proche.

Pour la suite, nous pensons qu'il serait intéressant de mener une analyse formelle dans un cadre basé sur des contraintes, de type Théorie de l'Optimalité. Le phénomène semble s'y prêter parfaitement. Il faudra également prolonger la réflexion en se demandant ce que les *rVCVCV*, et leur morphologie si particulière, peuvent nous apprendre sur les interactions entre phonologie et morphologie, étant donnée l'imbrication forte des deux dans la formation des *rVCVCV*. Phonologie et morphologie sont-elles à voir comme deux modules disjoints et autonomes, ou bien comme deux sous-systèmes en interaction, où l'une « regarde » ce que fait – ou ce que va faire – l'autre avant d'opérer ? Une autre piste prometteuse intéressant la phonologie théorique serait de creuser la question du poids des traits distinctifs et la manière dont ils se combinent ou se cumulent pour expliquer l'évitement de certaines consonnes. Nous avons vu que trois traits, communs à /r/ et aux consonnes peu représentées dans les *rVCVCV* étaient actifs : [continu], [voisé] et [sonant]. Il conviendrait aussi analyser plus finement le statut du trait [coronal]. Le cadre que nous avons adopté est un cadre assez classique issu des approches phonologiques en vogue dans la seconde moitié du XXe siècle, mais il faudrait examiner plus en détail ce que les différentes théories des traits distinctifs ont à dire sur ce sujet et comparer leurs prédictions.

Enfin, nous n'avons pu qu'effleurer les dimensions sémantiques, socio-linguistiques et connotatives. Si, suivant Guiraud (1986), les régularités formelles peuvent être considérées comme des régularités sémantiques, alors il conviendrait d'étudier plus finement ces aspects. C'est un sujet que nous laisserons aux spécialistes de ces domaines. La question des phonesthèmes, notamment, mériterait une analyse poussée en français.

References

1. Dal, G. 2003. Productivité morphologique : définitions et notions connexes. *Langue française* 140. pp. 3-23.
2. Dal, G. et F. Namer. 2018. Les occasionnalismes et la question de la productivité : le locuteur à l'œuvre. Pourquoi ? Comment ? *Neologica* 12. pp. 71-89.
3. Fónagy, I. 1991. *La vive voix. Essais de psycho-phonétique*. 2e éd. Paris : Payot.
4. Goldsmith, J. 1979. *Autosegmental Phonology*. New York: Garland.
5. Guiraud, P. 1986. *Structures étymologiques du lexique français*. Paris : Payot.
6. Kager, R. 1999. *Optimality Theory*. Cambridge: Cambridge University Press.
7. Labrune, L. 2005. Autour de la syllabe ; les constituants prosodiques mineurs en phonologie. *Phonétique et phonologie, approches contemporaines*, N. Nguyen, S. Wauquiers, J. Durand (éds), Paris : Hermès. pp. 95-116.
8. Labrune, L. 2006. Phonemic preferences in Japanese non-headed binary compounds : what *waa-puro*, *mecha-kucha* and *are-kore* have in common. *Gengo Kenkyū, Journal of the Linguistic Society of Japan* 129. pp. 3-41.
9. Labrune, L. 2024. Causes and effects of misreported syllable structures. *Theoretical Linguistics*, vol. 50, n° 3-4. pp. 233-253.
10. Léon, P. 1969. Principes et méthodes en phonostylistique. In *Langue française* n°3. *La stylistique*, sous la direction de M. Arrivé et J-Cl. Chevalier. pp. 73-84.
11. Lowenstamm, J. 2003. Remarks on Mutae cum Liquida and Branching Onsets. In *Living on the Edge, 28 Papers in Honor of Jonathan Kaye*, Stefan Ploch (éds.), *Studies in Generative Grammar* 62, Mouton de Gruyter, Berlin New York. pp. 339-363.
12. Nyrop, K. 1936. *Grammaire historique de la langue française, t. 3 : Formation des mots*, 2e éd. revue, Copenhagen-Paris : Picard
13. Paradis, C. et J.F. Prunet (éds). 1991. *The special status of coronals, internal and external evidence*. San Diego, CA: Academic Press.
14. Plénat, M. 1984. Toto, Fanfa, Totor et même Guiguitte sont des ANARs. *Forme sonore du langage*. Paris: Hermann. pp. 161-181.
15. Plénat, M. 1999. Prolégomènes à une étude variationniste des hypocoristiques à redoublement en français. *Cahiers de grammaire* 24. pp. 183-219.
16. Roché, M. 2008. Quelques exemples de morphologie non conventionnelle dans les formations construites ; à partir d'un mot en ouille(r). In *La raison morphologique, Hommage à la mémoire de Danielle Corbin*, B. Fradin (éd). Amsterdam : Benjamins.
17. Roché, M. 2010. Base, thème, radical. *Recherches linguistiques de Vincennes* 39. pp. 95-134.
18. von Wartburg, W. et O. Bloch (dir.). 2004. *Dictionnaire étymologique de la langue française* (2^e édition). coll. « Quadriges. Paris : Presses Universitaires de France.
19. Wioland, F. 1972. Estimation de la « fréquence » des phonèmes en français parlé. *Travaux de l'Institut de Phonétique de Strasbourg* (4). pp. 177-204.