

De l'incertitude en linguistique : le cas des variantes

Michel Mathieu-Colas

▶ **To cite this version:**

Michel Mathieu-Colas. De l'incertitude en linguistique : le cas des variantes. Gaston Gross ; Klaus U. Schulz. Linguistics, Computer Science and Language Processing. Festschrift for Franz Guenthner on the Occasion of his 60th Birthday, College Publications (London), pp.203-217, 2008. halshs-00410941

HAL Id: halshs-00410941
https://shs.hal.science/halshs-00410941

Submitted on 30 Nov 2009

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

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

DE L'INCERTITUDE EN LINGUISTIQUE : LE CAS DES VARIANTES

Michel Mathieu-Colas

LDI (UMR 7187), CNRS et Université Paris 13

La présence de variantes à tous les niveaux du fonctionnement linguistique, loin d'être un épiphénomène, représente un paramètre essentiel des langues naturelles, au même titre que la polysémie ou la synonymie. Il est vrai que les dictionnaires et les grammaires tendent souvent à sous-estimer cet aspect, appauvrissant ainsi, par excès de normativité, la représentation de la langue. Mais le phénomène se manifeste subrepticement sous une autre forme, dès lors que l'on compare les ouvrages de référence : à côté des variantes déclarées, explicitement décrites, il existe de nombreuses variantes « cachées », non reconnues comme telles, et détectables seulement par superposition des sources. Une méthodologie fondée sur la comparaison permet de prendre conscience de cet espace de variation inscrit au cœur de la langue.

On pourrait croire que cette dimension représente un handicap du point de vue de l'informatisation de la langue, comme si l'informatique exigeait une « réduction » de la richesse linguistique, pour rejoindre la simplicité formelle des langages artificiels. Or le travail sur les variantes prouve le contraire. L'effet de « flou » qui leur est propre n'est pas incompatible avec le traitement automatique : bien au contraire l'informatique, par son exigence de précision, incite à décrire avec rigueur le jeu apparemment irrationnel des variations et permet par là même de le domestiquer.

1. Variations formelles

Du point de vue formel, les variations opèrent à tous les niveaux de la langue.

Les variantes phonétiques

Rappelons des faits connus. Les variantes vocaliques affectent notamment le *e* « muet » ([fənɛtʀ] ou [fnɛtʀ]), les deux timbres du *a* [1], la nasalisation ([mɔ̃nami] ou [mɔnami]) et le degré d'aperture : *maison* peut se prononcer [mezɔ̃] (*Lexis*) ou [mɛzɔ̃] (*Robert & Collins*, comme dans le film *E.T.*). Pour ce qui est des consonnes, l'hésitation concerne en premier lieu le redoublement ([ilymine] ou [illymine]) et la prononciation des finales ([anana] ou [ananas]), sans même parler des variantes contextuelles (« ils étaient *six* : *six* enfants, *six* garçons »).

On observe également des flottements plus originaux, comme les suivants :

[1] Qu'on me permette d'évoquer un souvenir personnel. Un de mes enfants, alors en cours élémentaire, devait trier une série de mots selon la prononciation de la lettre *a* [a / ɑ]. Je me revois hésitant sur la prononciation d'un nom aussi simple que *bras* et consultant immédiatement, par réflexe professionnel, deux dictionnaires... pour découvrir, sans véritable surprise, qu'ils étaient en désaccord : le *Lexis* prononce [bra] ; le *Petit Robert* [brɑ]. J'ai dû me rendre à l'évidence : la question était mal posée et portait en fait sur la présence ou non d'un accent circonflexe !

- lecture de lettres initiales : le *Petit Larousse* de 1992 indique bien, pour *hiéroglyphe*, la présence d'un h muet, mais cette information est simultanément démentie par la légende de l'illustration (« exemples *de* hiéroglyphes »), rejoignant ainsi l'interprétation du *Petit Robert* (h aspiré) . De même opposera-t-on la lecture des mots en *x-* dans le *Lexis* ([ksenofɔb]) et dans le *Petit Robert* ([gzenɔfɔb]) ;
- interprétation d'un digramme, par ex. *gn* dans *prégnant* : le *Petit Robert* opte pour [pregnã], là où le *Lexis* préfère [preɲã] ;
- découpage syllabique : la prononciation [jɛʀ], choisie par le *Petit Robert*, est considérée comme « rare » par le *Lexis*, qui privilégie [ijɛʀ]. On comparera de même, pour *fiancé*, les transcriptions [fjãse] (*Petit Robert*) et [fijãse] (*Lexis*) ;
- variations propres aux emprunts, comme dans le cas de *dominion* : [dɔminjɔ̃] (*TLF*) ou [dɔminjɔn] (*Petit Robert*).

Il ne s'agit pas ici de faire un recensement systématique de tous les types de variantes phonétiques, mais simplement d'insister sur l'existence du phénomène et sur son importance pour le traitement automatique : la reconnaissance vocale implique que toutes les formes attestées soient connues et reconnues, ce qui est loin d'être toujours le cas dans les dictionnaires traditionnels [2].

Les variantes graphiques

Ayant longuement travaillé sur les variantes graphiques, nous résumerons ici l'essentiel de nos observations. Le phénomène affecte aussi bien la langue générale (*clé* ou *clef*) que l'argot (voir plus loin l'exemple de *gnôle*) ou les langues spécialisées. Citons, pour ces dernières, l'exemple du terme *intentionnalité*, écrit avec *nn* dans les œuvres de Sartre ou de Merleau-Ponty, mais un seul – *intentionalité* – chez Husserl (*Méditations cartésiennes*), Cuvillier (*Vocabulaire philosophique*) ou J.-F. Lyotard (*La Phénoménologie*).

On attend donc des dictionnaires qu'ils prennent en compte ces variations. Le *TLF*, de ce point de vue, occupe une place privilégiée : les variantes y sont très nombreuses et décrites avec soin. On regrette seulement que leur présentation s'avère par trop hétérogène. Elles peuvent, selon les cas, figurer en entrée – AÏOLI, AILLOLI ; APPOG(G)IATURE – ou être renvoyées en commentaire dans la rubrique « Prononc. et Orth. », à moins que les deux systèmes ne se combinent, comme il arrive pour *gnôle*, dont l'orthographe autorise presque toutes les combinaisons théoriquement possibles (*gn/gni/ni+o/ô/au*) :

> GNOLE, GNÔLE, subst. fém.
> **Prononc. et Orth.** Comme d'autres mot pop. il connaît de nombreuses var. graph. : *gnole (Lar. 20e, Lar. encyclop., Lar. Lang. fr.), gnôle (Lar. 20 e, Lar. encyclop., Lar. Lang. fr.,* ROB.), *gniole (Lar. Lang. fr.), gniôle (Lar. 20 e, Lar. encyclop.,* ROB.), *gniaule (Lar. 20 e, Lar. encyclop.), gnaule* (ROB., *Lar. Lang. fr.), niaule (Lar. 20 e, Lar. encyclop., Lar. Lang. fr.), niôle (Lar. encyclop., Lar. Lang. fr.,* ROB.), *gniolle* (LE BRETON).

On notera toutefois que des graphies usuelles se trouvent reléguées en commentaire (par ex. *allégro* ou *acuponcteur*), tandis que des variantes rares sont mises

[2] La meilleure référence, de ce point de vue, reste le *Dictionnaire de la prononciation française dans son usage réel* d'André Martinet et Henriette Walter (1973).

en vedette (BÉRET, *BERRET)*. Madame Catach, qui connaissait le sujet mieux que quiconque (voir entre autres son étude de 1971), avait attiré notre attention sur le cas étonnant de l'entrée DAME, *DAM*...

Les autres dictionnaires pèchent plutôt par défaut. Si certaines variantes sont bien représentées (on trouve partout *clé* ou *clef*, *saoul* ou *soûl*, *trimbaler* ou *trimballer*, *téléphérique* ou *téléférique*), d'autres sont masquées, perdues dans un exemple ou le corps d'un article. A moins qu'elles n'apparaissent seulement par superposition. Tel ouvrage opte pour une graphie, sans se préoccuper des discordances qui l'opposent à d'autres sources. Le *Petit Larousse* écrit *kasher, casher* ou *cachère,* le *Petit Robert casher* ou *kascher* (une seule graphie commune). La confusion est encore plus grande pour les mots composés, qu'il s'agisse de la présence du trait d'union (*fille-mère* ou *fille mère*) ou des marques de nombre (*marine à voile* ou *marine à voiles, boutons de manchette* ou *boutons de manchettes*). Et que dire des majuscules ? Elles flottent au gré des références : *jeux olympiques, jeux Olympiques, Jeux olympiques, Jeux Olympiques* [3]...

Les dictionnaires électroniques conçus dans le cadre du TAL ne peuvent se permettre de telles incertitudes. Il importe que toutes les formes soient recensées et décrites avec rigueur. Rien n'empêche si on le souhaite, en prévision de certaines applications, de marquer comme telles les graphies rares ou archaïques (ce qui permettrait de distinguer *béret* de *berret,* ou de privilégier *belote* par rapport à *belotte*). L'essentiel est d'être le plus précis possible afin de limiter, autant que faire se peut, le nombre de formes non reconnues.

Les variantes morphologiques

Deux composantes sont en jeu : le genre des substantifs et la flexion des mots variables. S'agissant des hésitations relatives au genre – dans le cadre d'un même emploi –, on connaît les cas les plus fréquents (*un* ou *une après-midi*) ou tout au moins les mieux décrits dans les dictionnaires [4]. Mais la confrontation de sources différentes ferait apparaître, ici aussi, un certain nombre de désaccords : *tennis* (au sens de « chaussure ») est masculin pour le *Lexis*, le *TLF*, le *Robert & Collins*, mais féminin pour le *Petit Larousse* et le *Robert Brio*. Il faut compter aussi avec les incertitudes liées à la féminisation : *Madame le député* ou *Madame la députée* ? Les deux formes, naturellement, doivent être enregistrées.

Pour ce qui est de la flexion, l'hésitation ne se réduit pas à un problème orthographique (des *presse-citron* ou *des presse-citrons*). Il arrive qu'elle se fasse entendre : le *Grand Robert* écrit *nasaux*, mais admet, en composition, *semi-nasals* ou *semi-nasaux,* quand le GDEL, à l'inverse, accepte *nasals* ou *nasaux,* mais ne reconnaît que le composé *semi-nasaux* ! On marche sur la tête...

[3] Pour les sources de ces graphies et, plus généralement, sur l'idée de « majuscule flottante », v. Mathieu-Colas 1998.

[4] Voir par exemple les mots suivants, dans le *Petit Robert* : *acmé, aglycone, antienzyme, apoenzyme, après-midi, arolle (arole), avant-guerre, avant-midi, barbouze, baston, carpocapse, chaintre, chapska (schapska), chibouque, chistera, coenzyme, cotyle, country, fétuque, goulache, happy end, harissa, holding, jaque, jet-set, limule, livedo (livédo), manse, métaldéhyde, météorite, mirepoix, nasique, notonecte, oolithe (oolite), ope, ophrys, parka, pentatome, perce-neige, phalène, phonolithe (phonolite), poiscaille, poliste, protogine, protophyte, psylle, quatre-épices, schappe, sex-shop, squatine, stout, tachine, tanagra, thermos, township, vérétille, wateringue.*

Il en va de même, toutes proportions gardées, pour la conjugaison. Témoin le verbe *bruire* : « Après s'être conjugué comme *fuir* (cf. bruyant), explique le *Petit Robert*, ce verbe se conjugue comme *finir* (p. prés. *bruissant*). » Pourtant le *Petit Larousse*, dans ses tableaux de conjugaison, mentionne encore l'imparfait *bruyait*. Plus généralement, les variantes verbales peuvent relever de plusieurs types :

- formes isolées : l'infinitif (*fiche* ou *ficher*), le présent de l'indicatif (*il éclôt* ou *il éclot*), le participe (*rassi, ie* ou *rassis, ise*), etc. ;
- alternances propres à un verbe, comme il arrive pour *asseoir* (*assois* ou *assieds*, *assoyais* ou *asseyais*, *assoirai* ou *assiérai*..., avec mélanges possibles pour un même locuteur : *je m'assois, nous nous asseyons, ils s'assoient, assieds-toi*) ;
- hésitation entre deux modèles : *harceler*, selon les sources, se conjugue comme *geler* ou comme *appeler* (*il harcèle* ou *il harcelle*) ; quelques verbes, dont *breveter*, peuvent se fléchir comme *jeter* ou *acheter* (*je brevète* ou *je brevette*) ; *départir* se conjugue comme *partir* « selon l'Académie » et comme *finir* « selon l'usage » (GDEL) ;
- alternances plus régulières affectant un ensemble de verbes. C'est le cas, par exemple, pour les verbes en *–ayer* (*je paie, je paierai* ou *je paye, je payerai*). On peut mentionner aussi le futur et le conditionnel des verbes en *–é..er* (*céder, célébrer, régner, révéler, rapiécer, assiéger*, etc.) ; si la tradition et les guides de bon usage imposent ici l'accent aigu (*je céderai, il régnera*), l'accent grave, plus conforme à la prononciation, tend à se généraliser (*je cèderai, il règnera*), au point d'être retenu comme seule graphie par la dernière édition du Dictionnaire de l'Académie. On admirera au passage la prudence du *Petit Robert* qui, dans le tableau de conjugaison correspondant, privilégie l'accent aigu (malgré la variation phonétique : [sedʀɛ ; sɛdʀɛ]), en se bornant à signaler en bas de page l'autre possibilité (« la prononciation actuelle appellerait plutôt l'accent grave »). Le *Petit Larousse*, encore moins engagé, évoque dans une simple note le choix de l'Académie, sans le prendre à son compte.

Aucun ouvrage n'énumérant toutes les variantes, seule une collation des sources disponibles permet de faire émerger l'ensemble des possibilités. Il est souhaitable d'y ajouter les marques d'usage, afin de différencier les formes « usuelles » des formes rares ou vieillies.

Les variantes syntaxiques

La syntaxe offre également son jeu d'options. Rappelons-en quelques exemples :

- alternances prépositionnelles : *d'ici demain* ou *à demain, habiter Paris* ou *à Paris, continuer à* ou *de* [5]... ;
- système interrogatif : *vous êtes satisfait ? êtes-vous satisfait ? est-ce que vous êtes satisfait ?*;
- négations : *je ne sais (pas) ce qui m'a pris*. Il faut compter aussi avec les oscillations du *ne* « explétif » : *je crains qu'il (ne) vienne* ;

[5] Il est difficile de trouver entre ces deux constructions une différence sémantique : « On a le choix ; c'est parfois l'oreille qui choisit » (Hanse : 268).

- inversions : *ainsi, il pourra* ou (sans virgule) *ainsi pourra-t-il...* ;
- accords : *c'est eux* ou *ce sont eux qui me l'ont dit ; la majorité des blessés survit* ou *survivent ; le peu d'observations que j'ai fait* (ou *faites*) *à cette occasion s'est effacé* (ou *se sont effacées*) *de ma mémoire* (Hanse 1983 : 680). On notera à ce propos que les problèmes d'accord ne concernent pas seulement l'orthographe, même s'il est vrai que la variation est souvent inaudible : *les épreuves que j'ai eu* (ou *eues*) *à surmonter ; je suis tout* (ou *toute*) *à vous* [6]*...*

Il existe bien d'autres variantes, que nous ne pouvons développer ici. Elles sont plutôt mieux décrites que les autres types de fluctuation, ne serait-ce qu'en raison de l'intégration plus ou moins systématique des variantes dans certaines grammaires. On songe en particulier au *Bon Usage* de Grevisse qui multiplie, chaque fois qu'il est possible, les exemples contradictoires :

> « *Vous trouverez* CI-INCLUSE *la copie que vous m'avez demandée* (Académie, s.v. *ci*). [...] *Vous trouverez* CI-INCLUS *une lettre de votre père* (Académie, s.v. *inclure*).
>
> L'Acad. (aux mots *inclure* et *joindre*) pose en règle que *ci-inclus* et *ci-joint* restent invariables quand ils précèdent le nom auquel ils se rapportent. Cette règle n'est pas sûre, et l'Acad. elle-même ne la respecte pas, comme le montre un des exemples que nous avons reproduits. [7] »

Mais tous les ouvrages ne procèdent pas ainsi et beaucoup (à l'instar de l'Académie !) préfèrent édicter des règles univoques – ce qui ne les empêche pas, à l'occasion, de se contredire mutuellement. Il faut écrire, selon Thomas (1956 : 236), *La première et la deuxième déclinaison latine*, alors que, pour Hanse (1983 : 39), « il est logique de mettre au pluriel le troisième adjectif dans une phrase comme celle-ci où le nom reste au singulier : *La première et la deuxième déclinaison latines*. Les deux déclinaisons sont latines. » Les comparaisons, en révélant des désaccords, mettent au jour des variantes qu'on s'évertuait à nous dissimuler.

La langue orale aussi a sa propre syntaxe, plus familière (Leeman 1994). On ne saurait omettre les variantes les plus usuelles : les négations tronquées (*il est pas venu*), les substitutions de préposition (*aller chez le coiffeur > aller au coiffeur*), les changements de mode (*après qu'il a* ou *qu'il eut parlé > après qu'il ait parlé* [8])... Il suffit simplement, pour plus de précision, de les assortir d'une marque, comme on le fait pour les mots (*fam., région., rare, vx,* etc.).

Toutes ces variations doivent être identifiées d'autant plus clairement que l'on souhaite soumettre des textes (ou des discours) à un analyseur syntaxique. L'ordinateur est totalement tributaire des descriptions qu'on lui fournit, ce qui implique, en amont, un long travail d'analyse linguistique.

---

[6] A moins qu'on ne veuille établir ici, quand le sujet est féminin, une nuance sémantique d'une grande subtilité : *Je suis toute à vous*, formule réservée à la passion, mais *Je suis tout à vous*, simple politesse (Thomas : 410) !

[7] 12e édition, p. 1368.

[8] Pour éviter l'emploi jugé incorrect du subjonctif avec *après que* (*après qu'il ait parlé, tout le monde s'est levé*), tout en évitant le mode indicatif de moins en moins perçu comme naturel (*après qu'il a parlé...*), certains recourent systématiquement au passé antérieur (*après qu'il eut parlé, tout le monde s'est levé*). Ils mélangent ainsi deux systèmes temporels différents, fondés respectivement sur le passé composé (*après qu'il a parlé, tout le monde s'est levé*) et sur le passé simple (*après qu'il eut parlé, tout le monde se leva*). Le remède n'est-il pas pire que le mal ?

2. Interférences sémantiques

Dans les exemples précédents, la variation n'affectait que la forme ou la construction des unités lexicales, le sens sous-jacent restant stable et clairement défini. Or ce n'est pas toujours le cas : la polysémie et l'homonymie inhérentes aux langues naturelles introduisent un nouveau facteur de complexité. Il en résulte de singulières interférences, notamment sur les plans graphique et morphologique.

Tout est simple, naturellement, quand les formes concurrentes renvoient à des emplois différents, qu'il s'agisse d'homonymes natifs, issus d'étymons distincts (*mer* / *mère* / *maire*), ou d'homonymes obtenus par différenciation graphique ([kɔ̃te], du lat. *computare*, s'écrit *compter* ou *conter* selon le sens). A chaque emploi sa forme, à chaque forme son emploi.

Mais l'interprétation est souvent plus délicate, ainsi lorsque deux formes fonctionnent en même temps comme homonymes et comme variantes :

VARIANTES				HOMONYMES	
abside	ou	*apside* (vx)	archit.	*apside*	astron.
différentier	ou	*différencier*	math.	*différencier*	langue gén.
disharmonie	ou	*dysharmonie*	langue gén. ou géol.	*dysharmonie*	méd., etc.

On notera, pour le dernier cas, les contradictions entre dictionnaires. Le *Petit Larousse* admet les deux graphies pour le sens général (« absence d'harmonie entre des choses, des personnes »), mais impose *disharmonie* dans le domaine géologique. Inversement, le GDEL n'accepte les deux variantes qu'au sens géologique et choisit *dysharmonie* dans tous les autres cas (langue gén., anthropol. soc., méd., psychiatr.). Le *Petit Robert*, quant à lui, exige *dysharmonie* en médecine, mais laisse le choix pour le sens général !

Si le nombre d'emplois et de graphies augmente, les désaccords peuvent aboutir à une véritable cacophonie :

- le *Lexis* oppose *arcure* (arboriculture) et *arqûre* (médecine vétérinaire) ; mais d'autres dictionnaires proposent des graphies différentes (*arcûre* et *arqure*) et ajoutent trois autres domaines (armement, zootechnie, marine), sans qu'il y ait unanimité sur la répartition sémantique des formes : il en résulte autant de variantes potentielles ;
- le cas de *tête de maure* (*more* ou *mort*) n'est pas moins éloquent. En lisant de près les dictionnaires et en les comparant, on ne relève pas moins de 10 graphies différentes, avec présence ou non de traits d'union et/ou de majuscules – *tête-de-maure, tête de maure, tête-de-Maure, tête de Maure, tête-de-more, tête de more, tête-de-More, tête de More, tête-de-mort, tête de mort* –, lesquelles graphies recouvrent, selon des distributions variables d'un ouvrage à l'autre, 6 significations distinctes (sans compter le sens usuel de *tête de mort*), du fromage de Hollande au muflier, en passant par un couvercle, un type de nœud, un papillon, une couleur [9]...

[9] Pour une description plus détaillée, voir Mathieu-Colas 1995. Nous présentons ici même, de façon plus développée, un autre exemple de « cacophonie » (*infra*, Annexe §1).

Ce type de « mise à plat » est nécessaire, ne serait-ce que pour prendre conscience de la part d'arbitraire de certains choix dictionnairiques et leur substituer une représentation plus cohérente.

Les verbes offrent aussi quelques interférences, au croisement de la sémantique et de la morphologie. Si les deux verbes *ressortir* – « sortir de nouveau » ou « être du ressort de » – ont un comportement clair, chaque sens développant ses propres formes (*il ressort de chez lui* vs *cette affaire ressortit à*), d'autres verbes sont problématiques, comme le montrent les formes *repartir* et *répartir* (voir aussi, en annexe §2, les verbes *fleurir* et *saillir*).

Repartir, dans l'usage courant (« partir de nouveau »), se conjugue comme *partir* (*repartons, repartant, repartait*), cependant que *répartir*, au sens de « partager », suit le modèle de *finir* (*répartissons, répartissant, répartissait*). Or voici que tout se complique avec la prise en compte du couple de variantes *repartir*/*répartir* synonyme de « répondre », « répliquer » :

> « Il ne tiendra qu'à vous, beau Sire,
> D'être aussi gras que moi, lui repartit le chien[10]. »

Il y a hésitation, pour la conjugaison, entre les modèles *partir* et *finir* [11], même si la difficulté est masquée pour les formes les plus fréquentes (le passé simple et le participe passé étant identiques dans les deux paradigmes).

Il faut donc se résoudre à prendre à bras-le-corps ce type d'imbroglios, si embarrassants qu'ils puissent paraître. Si l'on veut y voir clair dans la langue – mais comment, sans cela, prétendre y appliquer un traitement informatique ? –, tout doit être identifié, analysé, interprété.

Ajoutons que le même principe pourrait s'appliquer, *mutatis mutandis*, aux interférences entre synonymie et polysémie : tout comme pour les variantes, il s'agit de mettre en correspondance des ensembles de formes et des séries d'emplois, par-delà les divergences des dictionnaires. Nous en donnerons un exemple en annexe (§3).

Conclusion

Les analyses précédentes ont été développées dans le cadre de travaux sur l'informatisation de la langue, en particulier l'élaboration de dictionnaires électroniques conçus en vue du traitement automatique. Mais elles dépassent le simple champ des applications et prennent, à notre sens, une valeur méthodologique et épistémologique. Sur un plan fondamental, l'informatique est pour le linguiste une école de rigueur : elle ne saurait se satisfaire de l'implicite ou de l'approximation. On peut songer à

[10] Les éditions imprimées des *Fables* de La Fontaine, pour « le Loup et le Chien », privilégient généralement la graphie sans accent. S'agissant du Web, on notera que, sur 240 occurrences de ce vers (selon Google, à l'heure où nous écrivons), *repartir* apparaît 172 fois et *répartir* 68 fois. La prononciation ne correspond pas exactement à l'orthographe, à en croire le *Petit Robert* : *repartir* peut se prononcer à volonté [ʀ(ə)-] ou [ʀe-], alors que le choix n'est pas possible pour *répartir* [ʀe-] ; quant au substantif dérivé, il se prononce toujours [ʀe-] quelle que soit la graphie (*repartie* ou *répartie*).

[11] Selon Grevisse 1986, *repartir*, même au sens de « répondre », se conjugue comme *partir*. Il constate toutefois « une tendance à conjuguer *repartir* 'répondre' (qu'on écrit parfois °*répartir*)... comme *finir* » (p. 1253).

Heisenberg et au « principe d'incertitude » qui régit la microphysique : à une échelle microscopique, tout semble se troubler ; mais le « flou » est lui-même susceptible d'une représentation fine. S'il est vrai que la langue n'a rien à envier à la complexité de l'univers physique [12], elle exige, tout comme lui, une description cohérente et minutieuse.

RÉFÉRENCES

Dictionnaires

Grand Dictionnaire encyclopédique Larousse [**GDEL**], Larousse, 1982-1985.
Grand Larousse encyclopédique, Larousse, 1960-1964
Grand Robert de la langue française, Le Robert, 1985
Larousse du XX^e siècle, Larousse, 1927-1933
Le Nouveau Petit Robert, Le Robert, 2002
Le Petit Larousse illustré 1992, Larousse, 1991
Le Petit Larousse 2003, Larousse, 2002 (sauf mention spéciale, les références au *Petit Larousse* se rapportent à cette édition).
Le Robert Brio, Le Robert, 2004.
Lexis, Larousse, 1979.
Nouveau Larousse illustré, Larousse, 1897-1904
Le Robert & Collins Senior, 3e édition, Le Robert, Collins Publishers, 1993
Trésor de la langue française [**TLF**], Ed. du C.N.R.S., 1971-1983, Gallimard, 1985-1994.
HANSE, J. *Nouveau Dictionnaire des difficultés du français moderne*, Duculot, 1983.
LITTRÉ. *Dictionnaire de la langue française*, Paris, Hachette, 1863-1872.
THOMAS, A. V. *Dictionnaire des difficultés de la langue française*, Larousse, 1956.

Autres références

CATACH, N. *Orthographe et lexicographie*, tome 1, Didier, 1971.
GREVISSE, M. *Le Bon Usage*, 12e édition refondue par A. GOOSE, Duculot, 1986.
LEEMAN, D. *Les fautes de français existent-elles* ?, Seuil, 1994.
MARTINET, A. et H. WALTER. *Dictionnaire de la prononciation française dans son usage réel*, France-Expansion, 1973 (CILF).
MATHIEU-COLAS, M. *Les mots à trait d'union. Problèmes de lexicographie informatique*, CNRS-INaLF, Didier érudition, coll. « Études de sémantique lexicale », 1994.
— « Orthographe et informatique : Établissement d'un dictionnaire électronique des variantes graphiques », *Langue française*, n° 87, 1990, pp. 104-111.
— « Un dictionnaire électronique des mots à trait d'union », *Langue française*, n° 108, 1995, pp. 76-85.
— « La majuscule flottante. Remarques sur l'orthographe des noms propres composés (type *NAdj*) », *BULAG*, Université de Franche-Comté, Besançon, 1998.

(Voir l'annexe page suivante)

[12] Le regretté Alain Guillet me disait un jour, alors que je lui faisais part de mon émerveillement devant les richesses de l'univers révélées par l'astrophysique : « La langue est tout aussi complexe ! »

ANNEXE : ÉTUDES DE CAS

Les analyses suivantes illustrent, chacune à sa manière, le problème soulevé par l'imprécision des sources traditionnelles, dès lors qu'on les soumet à l'épreuve de la comparaison. Les données lexicographiques ne peuvent être utilisées pour l'élaboration d'un dictionnaire électronique qu'après un travail minutieux de « mise au point » (au sens photographique du terme).

1. Graphie et sémantique : *espar, épar, épart, espart, espars*

Si l'adjectif *épars* reçoit un traitement univoque, il n'en va pas de même, malheureusement, pour les substantifs homophones. Initialement, tout semble simple : plusieurs dictionnaires distinguent clairement les mots *espar* (terme de marine désignant une pièce du gréement d'un bateau) et *épar* (techn. : traverse ou barre servant à fermer une porte). Toutefois, le *Petit Robert* et le *Petit Larousse* ajoutent, pour le deuxième sens, une variante *épart*, cependant que le *Grand Larousse encyclopédique* et le *Lexis* mentionnent un troisième terme : *espart* ou *espars* (« *Text*. Cheville de bois servant à tordre les écheveaux et les tissus au sortir des bains de teinture. »). Ce qui donne, provisoirement, la répartition suivante :

1. *espar* (mar.)
2. *épar* ou *épart* (techn. : barre)
3. *espart* ou *espars* (text.)

Mais la consultation du *Grand Robert* (1985) ébranle cet édifice : *espar* y est analysé selon plusieurs emplois, dont l'un, nous apprend-on, équivaut à *épar* (techn.), tandis que l'autre (terme de marine) « s'écrit aussi *espars* » ; à quoi s'ajoute un énième terme, écrit *épart* ou *épars* (« Vx ou régional. Éclair de chaleur »). En résumé, pour le *Grand Robert* :

1. *espar* ou *espars* (mar.)
2. *épar*, *épart* ou *espar* (techn. : barre)
3. *épart* ou *épars* (vx ou région. : éclair)

En prenant au pied de la lettre les indications du *Grand Robert* (GR) et en les fusionnant avec les données précédentes, en obtiendrait la combinatoire suivante :

	mar.	text.	techn. (barre)	vx ou rég. (éclair)
épar			+	
épars				(GR)
épart			+	(GR)
espar	+		(GR)	
espars	(GR)	+		
espart		+		

Soit 6 graphies pour 4 emplois, avec de nombreux recouvrements (sans compter l'adjectif *épars*). Quoi qu'il en soit de l'usage réel, ce tableau formalise les données lexicographiques et structure le flou : condition indispensable pour un traitement informatisé. Il rend aussi possible, s'il en était besoin, une enquête plus approfondie.

2. Morphologie et sémantique : *fleurir* et *saillir*

Le problème n'est pas ici graphique, mais concerne les rapports entre la sémantique et la morphologie flexionnelle, en l'occurrence la conjugaison.

Si *fleurir*, au sens propre, se conjugue comme *finir*, l'emploi figuré a une flexion plus hésitante, qui emprunte en partie à l'ancienne variante *florir*. Cela vaut notamment pour l'imparfait : « Le romantisme *fleurissait,* ou *florissait* en France au XIX^{e} siècle » (*Grand Robert*). Au participe présent – et toujours dans le cadre de l'emploi figuré –, *fleurissant* semble plus rare, alors qu'aux autres temps, ce sont les formes en *flor-* qui sont exceptionnelles (*florir, florissent, florirent, flori* ; cf. Grevisse 1986 : 1201). Il est possible d'accueillir toutes les variantes attestées, à condition de les assortir, chaque fois qu'il y a lieu, de marques explicites (*rare*, *littér*., etc.).

Saillir, au sens de « jaillir » ou de « s'accoupler », se conjugue théoriquement comme *finir* (*il saillissait* ; TLF, GDEL, etc.), mais on trouve aussi des formes empruntées à *assaillir* ou à *cueillir*. A l'inverse, quand il signifie « être en saillie », il se modèle en général sur *assaillir* et/ou sur *cueillir* (*il saillait, il saillera* ou *il saillira*), ce qui n'empêche pas certains de le conjuguer comme *finir* (cf . Grevisse 1986 : 1251). La description doit tenir compte de toutes ces possibilités, si l'on veut rendre possible la reconnaissance des formes et leur interprétation.

3. Polysémie et synonymie : *chasse-rond, chasse-rondelle, chasse-roue*

Ce dernier exemple ne concerne pas directement les variantes, mais pose un problème analogue. Il s'agit, ici aussi, de comprendre la relation entre des formes et des emplois, compte tenu des discordances entre les dictionnaires.

Partons du *Larousse du XX^{e} siècle* (LXX) et du *Grand Larousse encyclopédique* (GLE), qui distinguent clairement les deux unités suivantes :

> *chasse-rond(s),* Outil servant à pousser les moulures appelées *congés.*
> *chasse-roue(s),* Borne ou arc métallique destiné à empêcher les roues de détériorer les murs.

Or tout se brouille avec le TLF qui, après avoir défini *chasse-roues*, ajoute cette précision surprenante :

> Rem. Donné par Littré comme synon. de *chasse-rondelle,* « outil de charron qui sert à pousser les moulures appelées 'congés' » (*Lar. 20^{e}* et *Lar. encyclop.*).

En réalité, ni le GLE ni le LXX ne mentionnent *chasse-rondelle*. Et si Littré identifie *chasse-rondelle* et *chasse-roue*, c'est pour les définir comme « outils de charron ». Bilan provisoire :

chasse-rond(s)	outil servant à pousser les moulures (LXX, GLE)
chasse-roue(s)	borne protégeant les murs (LXX, GLE ; cf. GR et TLF)
chasse-roue = chasse-rondelle	outil de charron (Littré)
chasse-rondelle	« outil de charron servant à pousser les moulures » (TLF)

A ce niveau microscopique, les emplois semblent se dérober... Mais la consultation du *Nouveau Larousse illustré,* vieux de cent ans, remet les choses à leur place, en attribuant à chaque mot ses emplois spécifiques :

> CHASSE-ROND, Outil servant à pousser les moulures concaves en quarts de rond, appelées *congés*.
> CHASSE-ROUE ou CHASSE-RONDELLE, Outil de charron. ‖ Borne métallique ou arc destinés à empêcher les roues de détériorer les murs.

Au terme de l'analyse, on peut donc distinguer trois emplois et trois formes (sans compter les variantes graphiques) ainsi répartis :

	outil de menuisier	outil de charron	borne de protection
chasse-rond(s)	NLI, GLE, LXX		
chasse-rondelle		Littré, NLI	NLI
chasse-roue(s)		Littré, NLI	NLI, GLE, LXX , GR, TLF

Quant à l' « outil de charron servant à pousser les moulures », il semble résulter d'une erreur de lecture...

On voit ainsi que les interférences entre synonymie et polysémie imposent le même travail de clarification formelle et sémantique que les variantes. C'est à ce prix seulement qu'on peut envisager, sur des bases précises, l'informatisation de la langue.