



UFR HSS
HUMANITÉ
& SCIENCES SOCIALES

UNIVERSITÉ
CAEN
NORMANDIE

MORPHOLOGIE DES LANGUES

COURS

RICHARD RENAULT

Maître de conférences



TABLE DES MATIÈRES

| | |
|---|----|
| 1. La morphologie : définitions..... | 4 |
| 1.1. Études des mots..... | 4 |
| 1.2. Mot, forme de mot et lemme..... | 4 |
| 1.3. Études des morphèmes..... | 6 |
| 2. Morphologie et lexicologie..... | 8 |
| 2.1. Intuition et étymologie..... | 8 |
| 2.2. Mots possibles et mots attestés..... | 10 |
| 3. Le mot..... | 11 |
| 3.1. Le signe linguistique, sens, signification et valeur grammaticale..... | 11 |
| 3.2. L'autonomie du mot..... | 13 |
| 3.3. Les mots composés..... | 15 |
| 3.4. Les clitiques..... | 16 |
| 3.5. Le mot et les niveaux d'analyse..... | 17 |
| 3.6. La forme écrite des mots..... | 18 |
| 4. Les morphèmes..... | 18 |
| 4.1. Les mots et les morphèmes..... | 19 |
| 4.2. Morphèmes libres et morphèmes liés..... | 20 |
| 4.3. Radical, base et affixes..... | 21 |
| 4.4. Les affixes..... | 23 |
| 4.4.1. Les préfixes..... | 23 |
| 4.4.2. Les suffixes..... | 23 |
| 4.4.3. Les infixes..... | 24 |
| 4.4.4. Les circonfixes..... | 25 |
| 4.4.5. Les suprafixes..... | 26 |
| 4.5. Les radicaux comme morphèmes liés..... | 27 |
| 4.6. Morphèmes et allomorphes..... | 28 |
| 4.6.1. La variation contextuelle en phonologie..... | 28 |
| 4.6.2. La variation contextuelle en morphologie..... | 29 |
| 5. La structure interne des mots complexes..... | 32 |
| 5.1. La commutation en phonologie..... | 33 |
| 5.2. La commutation en morphologie..... | 34 |
| 5.3. Les relations syntagmatiques hiérarchisées..... | 36 |
| 5.4. justification de la segmentation..... | 38 |
| 5.5. Base et radical..... | 43 |
| 6. Concaténation et concomitance..... | 43 |
| 6.1. Paradigme morphologique et paradigme de mots..... | 43 |
| 6.2. Morphèmes et segments..... | 46 |
| 6.3. L'analyse en traits morphologiques..... | 49 |
| 7. Les types de mots..... | 51 |
| 7.1. Les types de mots lexicaux..... | 51 |

| | |
|---|----|
| 7.1.1. Les mots composés..... | 51 |
| 7.1.2. Les mots dérivés..... | 52 |
| 7.1.3. Les mots savants..... | 53 |
| 7.1.4. Les mots-valises..... | 53 |
| 7.1.5. Les mots amalgamés..... | 54 |
| 7.2. Mots construits et mots non construits..... | 54 |
| 7.2.1. Mots dérivés construits et non construits..... | 55 |
| 7.2.2. Mots non dérivés et non construits..... | 56 |
| 7.2.3. Synthèse..... | 57 |
| 8. Flexion et dérivation..... | 58 |

1. LA MORPHOLOGIE : DÉFINITIONS

Pour commencer, deux définitions de la morphologie :

1. Étude des mots et de la forme des mots.
2. Étude des morphèmes, de leurs réalisations, de leurs fonctions et de leurs relations.

La première définition repose sur la notion de mot tandis que la seconde repose sur la notion de morphème.

1.1. ÉTUDES DES MOTS

La première définition a l'inconvénient de reposer sur la notion de mot et le mot est quelque chose qui se laisse difficilement définir d'un point de vue général. En fait, il est pratiquement impossible de proposer une définition simple du mot qui soit valable pour toutes les langues. On peut cependant dire que le mot est un concept commun à toutes les langues, mais chaque langue le définit à sa manière en fonction de ses propriétés phonologiques, morphologiques, syntaxiques et lexicales. Il n'est donc pas surprenant que certains linguistes rejettent le mot comme unité d'analyse au profit du morphème.

Cependant le mot existe, et j'en donnerai pour preuve le fait qu'il est présent dans la conscience des locuteurs même dans les langues qui n'ont pas de tradition écrite, car il est vrai par ailleurs que le mot est souvent présenté comme une notion fortement marquée par la forme écrite de la langue.

1.2. MOT, FORME DE MOT ET LEMME

Dès lors qu'on aborde l'étude des mots, il convient de faire la distinction entre mot, forme de mot et lemme.

Selon la première définition, on est amené à distinguer les mots et les formes de mot. Pour comprendre la différence entre les deux, nous devons partir de la phrase ou du texte. Dans la phrase qui suit :

Je me suis donné le temps de lire la préface.

nous comptons 10 mots. Il s'agit en fait de 10 formes de mot car les constituants *je* et *me*, de même que *le* et *la* peuvent être considérés respectivement comme deux formes du même pronom et comme deux formes du même article.

Si nous considérons maintenant la langue dans son ensemble, nous sommes ainsi conduit à regrouper certaines formes de mot sous le même mot : Il en serait ainsi des exemples suivants :

{*cheval* et *chevaux*}

{*beau*, *beaux*, *belle* et *belles*}

{*donner*, *donné*, *donnant*, *donnait*, *donnera* ...}

Par contre la liste suivante :

{*accepter, acceptable, acceptation*}

ne constitue pas un ensemble de différentes formes du même mot. *accepter, acceptable, acceptation* sont trois mots différents, bien qu'ils présentent une régularité de forme comparable à celle de la liste du verbe *donner* de la liste précédente. Il appartient donc à la morphologie – définie en termes de mot – d'expliquer pourquoi ***accepter, acceptable, acceptation*** sont trois mots différents alors que ***acceptait, acceptera*** et ***acceptons*** sont trois formes du même mot. Notons qu'ils ont tous en commun une part de signification et une part de leur forme phonologique (séquence en gras).

Une forme de mot est le mot tel qu'il est employé dans la langue, dans un énoncé. La forme du mot dépend principalement de son contexte phonologique et syntaxique.

◆ nom

maison et *maisons* sont deux formes du même mot. L'un est au singulier, l'autre, au pluriel

◆ adjectif

beau, beaux, belle, belles et *bel* sont des formes différentes du même mot ; variation en genre et en nombre (contexte syntaxique) et la différence d'emploi entre *beau* et *bel* dépend du contexte phonologique : *bel* devant un mot qui commence par une voyelle (*bel appartement*) et *beau* devant un mot qui commence par une consonne (*beau décor*).

◆ pronom

j' et *je* sont deux formes du même mot ; le pronom de 1^{ère} personne du singulier. Le premier est employé devant un mot qui commence par une voyelle tandis que le second est employé devant un mot qui commence par une consonne.

je, me et *moi* sont trois formes du même mot ; le pronom de 1^{ère} personne du singulier. Le premier est sujet, le deuxième est complément, et le troisième est employé notamment comme complément d'une préposition.

◆ verbe

chanter, chantons, chantaient, chanterait, chanté, chantées, ... sont différentes formes du même mot. La variation concerne le mode, le temps et la personne. Pour les participes, la variation relève du temps (présent ou passé) du nombre et du genre.

◆ adverbe

très est l'unique forme du même mot (mot invariable)

Dès lors que la distinction entre forme de mot et mot est introduite, il est plus difficile de définir ce qu'est un mot. On peut contourner le problème en disant, par exemple, que *maisons* est une forme de mot et également un mot. Plus précisément, le mot se révèle plus complexe, plus abstrait, comme un ensemble de formes différentes qui partagent la même signification lexicale. La signification lexicale est ce que le dictionnaire donne comme définition du mot.

Résumons : *chanter, chantons, chantaient...* sont un même mot avec différentes formes car ils ont la même signification lexicale : produire sons mélodieux avec la voix....Par contre *chanter, chanteur, chanson*

sont des mots différents car ils n'ont pas la même signification lexicale, même s'ils partagent des propriétés sémantiques. *chanteur* n'a pas pour définition : produire sons mélodieux avec la voix mais plutôt : celui qui chante ou personne qui chante ...

Reste maintenant à introduire la notion de lemme. Lorsqu'on cherche dans le dictionnaire la définition d'un mot, sa signification lexicale, on ne cherche pas le mot à partir de toutes ses formes ; il n'y a pas une entrée dans le dictionnaire pour *maison* et une autre pour *maisons*, de même qu'il n'y a pas pour chaque forme verbale une entrée. Le mot est accessible dans le dictionnaire à partir d'une de ses formes : le lemme. Le lemme est donc la forme du mot qui représente l'ensemble des différentes formes du mot. La choix de la forme lemmatique repose sur un principe simple : on retient la forme non marquée (le concept de marque a été introduit par le linguiste R. Jakobson). La forme non marquée d'un mot est celle qui ne contient aucune information spécifique, forme neutre, sans morphème explicite. Le singulier est la forme non marquée d'un nom car un mot au singulier tel que *maison* ne contient aucun morphème de nombre explicite, contrairement à la forme *maisons* qui présente un morphème de pluriel (en gras). Le masculin singulier est la forme non marquée d'un adjectif car un mot au masculin singulier tel que *petit* ne contient aucun morphème de genre et de nombre explicite, contrairement à la forme *petites* qui présente un morphème de genre (souligné) et de pluriel (en gras). L'infinitif est la forme non marquée d'un verbe car cette forme ne donne aucune indication de temps (infinitif = non fini) contrairement aux formes conjuguées et aux participes. Pour les pronoms, le choix de la forme lemmatique est plus délicat ; on peut retenir la forme *moi* pour le pronom de première personne du singulier dans la mesure où cette forme est la seule qui peut s'utiliser seule, en dehors de tout contexte. Les adverbes étant invariables, le lemme est identique à l'unique forme du mot.

Dans les exemples suivants, les formes de mot sont en italiques, le mot correspond à l'ensemble des formes, et le lemme est souligné :

- ◆ nom {*maison, maisons*}
- ◆ adjectif {*beau, beaux, belle, belles, bel*}
- ◆ pronom de première personne du singulier {*je, j', me, m', moi*}
- ◆ verbe {*chanter, chantons, chantaient, chanterait, chanté, chantées, chantant...*}
- ◆ adverbe {*très*}

1.3. ÉTUDES DES MORPHÈMES

La seconde définition pose moins de problèmes du point de vue des termes impliqués : il existe une définition non-ambigüe et générale du morphème. Cette définition est la suivante : le morphème est la plus petite unité de langue qui soit dotée d'un signifiant et d'un signifié.

Je reviendrai sur la définition du morphème plus loin. Cette seconde définition ne dit rien de la relation entre les morphèmes et les mots. De fait, on peut traiter de la morphologie d'une langue sans qu'il soit

nécessaire de faire appel au préalable à une définition générale du mot et à l'organisation des morphèmes en mots.

Cette seconde définition s'accommode mieux de la diversité des langues : elle permet la comparaison des langues alors qu'une approche de la morphologie à partir du mot ne peut conduire qu'à la constatation que le mot ne recouvre pas la même chose selon les langues. J'illustrerai ce point en comparant deux énoncés sémantiquement équivalents en français et en finnois (deux langues très différentes) :

- a. *dans ses poches*
- b. *taskuissaan*
- c. *tasku+i+ssa+an* (*poche*+pluriel+"*dans*" +3e personne)

Le groupe de mots français *dans ses poches* est traduit en finnois par un seul mot : *taskuissaan*. Là où en français, nous avons 3 mots, nous en avons un seul en finnois. Par contre, si nous prenons le morphème comme élément de comparaison, les deux langues contiennent exactement les mêmes informations : l'énoncé français s'analyse comme une mise en construction syntaxique du mot *poche*, du nombre pluriel, du possessif de troisième personne, et de la préposition *dans*. Les mêmes informations se trouvent réunies dans le mot finnois, puisque celui-ci s'analyse comme une séquence de morphèmes (voir l'analyse morphologique en c). Les deux langues sont donc différentes quant à la forme des mots mais sont très proches quant au nombre de morphèmes présents dans l'énoncé (4 morphèmes dans les deux cas) et quant à leur signification. Dans cette comparaison on ne prend pas en compte l'accord en nombre entre le déterminant et le nom en français car il s'agit d'une propriété syntaxique. De même on laisse de côté ici le fait que le nom *poche* est masculin en français alors que la distinction de genre n'existe pas en finnois, car le genre grammatical n'est pas un élément de signification pour l'énoncé (le locuteur ne choisit pas le genre).

Les deux définitions sont donc à retenir, et nous les articulerons l'une par rapport à l'autre en définissant la morphologie comme l'étude des morphèmes, de leur(s) réalisations, de leurs fonctions et de leurs relations dans le cadre du mot. Le mot est ainsi conçu comme le produit de la mise en forme des morphèmes et comme l'unité la plus grande de l'analyse morphologique. De même que la phrase peut se définir comme l'unité maximale de l'analyse syntaxique. Précisons néanmoins que ce choix n'épuise pas l'analyse morphologique envisagée à partir des morphèmes. Certains morphèmes ont une fonction qui se définit non pas par rapport au mot mais par rapport aux syntagmes ou par rapport à la phrase dans sa totalité. La description de cette incidence de la morphologie sur la syntaxe s'appelle la morphosyntaxe.

Trois exemples simples suffiront à montrer que le morphème ne se définit pas seulement par rapport au mot :

- a. *chanteur* : [_N chanteur]
- b. *chants* : [_{SN} les premiers chants]
- c. *chantions* : [_{PHRASE} Nous chantions la Marseillaise tous les 14 juillet]

Dans le mot *chanteur*, le morphème *-eur* est en construction avec le verbe *chant(er)* (avec le radical du verbe *chanter* pour être plus précis ; *chanter* étant une forme verbale infinitive, et sa présence n'a pas d'autre

incidence sur le reste de l'énoncé dans lequel ce mot peut apparaître. *Chanteur*, du fait de la présence du suffixe *-eur* est un nom, et il aura par conséquent toutes les propriétés des noms simples dans la phrase. Par contre dans le mot *chants* (au pluriel) la présence du morphème de pluriel n'est pas une propriété du nom *chant* bien qu'il soit présent sur ce mot. Le pluriel est en fait une propriété du syntagme nominal tout entier. Cette propriété est explicite lorsque le nom *chant* est accompagné d'un déterminant et d'un adjectif comme dans le syntagme *les premiers chants*. Le pluriel est présent trois fois ; sur le déterminant, sur l'adjectif et sur le nom, mais il n'y a bien entendu qu'une seule notion de pluralité dans ce syntagme. La présence du pluriel sur le déterminant et sur l'adjectif relèvent d'une relation syntaxique d'accord dont la fonction est de signaler la cohésion du syntagme nominal. Considérons maintenant le mot *chantions* : ce mot contient un morphème de temps (*-i* pour l'imparfait) et la présence d'un morphème de temps dans une phrase n'est pas une propriété du verbe, même si c'est ce mot qui contient ce morphème. Le temps est une propriété de la phrase dans sa totalité, il en est d'ailleurs la principale caractéristique. C'est l'événement dans sa totalité (*chanter la Marseillaise tous les 14 juillet*) qui se situe dans le passé. Cette fois donc, le morphème de temps reçoit une interprétation dans le cadre de la phrase.

Faire les exercices 2 et 3

Ces exercices portent sur la segmentation d'énoncés (qui ne sont pas nécessairement des mots) en morphèmes. Les corpus sont empruntés à différentes langues afin de ne pas présupposer une segmentation en morphèmes ou mots sur la base d'une connaissance intuitive ou acquise. Voir le modèle d'analyse donné en annexe (document 2).

2. MORPHOLOGIE ET LEXICOLOGIE

2.1. INTUITION ET ÉTYMOLOGIE

Le premier point concerne la différence entre intuition et étymologie. S'agissant de décrire la compétence des locuteurs, il convient de souligner tout d'abord que notre intuition sur les mots complexes n'est pas nécessairement en accord avec leur étymologie. Comparons à cet effet les deux couples suivants :

déchirement / parlement

banalité / vanité

On reconnaît un même suffixe, *-ment* dans un cas, et *-ité* dans l'autre. Le premier exemple de chacun des deux couples de mots est parfaitement régulier: *déchirement* s'analyse en deux morphèmes *déchire* et *-ment*, et *banalité* s'analyse en *banal* et *-ité*. Sans qu'il soit nécessaire de faire appel à la commutation, notre intuition reconnaît en *-ment* un suffixe qui désigne le résultat d'une action, et cette analyse se retrouve dans beaucoup de mots comme les suivants :

affaiblissement (affaiblir)

peuplement (peupler)

étalement (étalement)

remplacement (remplacement)

Le suffixe *-ment* permet donc de former un nom à partir d'un verbe. Pour le deuxième couple de mots, on reconnaît en *-ité* un suffixe qui permet de former un nom à partir d'un adjectif. On retrouve cette propriété dans les mots suivants :

fragilité (fragile)

docilité (docile)

féroce (féroce)

universalité (universel)

Déchirement et *banalité* s'inscrivent donc dans un paradigme où le suffixe garde la même signification : *déchirement* appartient au même paradigme que *affaiblissement*, *peuplement*, *étalement*, *remplacement*. *Banalité* s'inscrit dans le même paradigme que *fragilité*, *docilité*, *féroce*, *universalité*.

Mais peut-on en dire autant de *parlement* et de *vanité*? S'il est vrai que *parlement* est dérivé de *parler*, nous ne reconnaissons pas aussi facilement dans ce mot le verbe *parler*, même si l'on dit qu'un *parlement* est un endroit où l'on parle. La signification de *-ment* n'a plus rien à voir avec celle du suffixe *-ment* que l'on a dans l'ensemble des mots auquel appartient notamment *déchirement*. *Parlement* n'est pas le résultat de l'action de parler. Il faut donc admettre que notre intuition sur ce mot ne s'accorde pas avec les éventuelles informations que peut fournir notre connaissance de l'étymologie du mot. *Parlement* sera donc traité de manière intuitive comme un mot simple par la majorité des locuteurs français, car on ne retrouve pas dans le suffixe *-ment* une signification connue.

Passons au deuxième exemple : l'écart qui sépare le premier exemple *banalité* du second *vanité* est moins grand qu'entre *déchirement* et *parlement*. On reconnaît dans *vanité* un concept lié également à un état, mais le morphème libre associé au segment *van-* se laisse plus difficilement appréhender que *banal* dans *banalité*. Il y a deux raisons à cela : d'une part *van-* est un allomorphe du morphème libre *vain* – entre les deux il y a une modification phonologique bien connue en français : une voyelle nasale se dénasalise au contact d'une consonne nasale, et il y a aussi dans ce cas un changement de timbre de la voyelle (autres exemples de dénasalisation dans le radical : *vin/vinifier*, *pain/panifier*). La simple dénasalisation de la voyelle [ɛ̃] aurait donné la voyelle [ɛ] comme dans la forme féminine de l'adjectif *vain*, c'est-à-dire *vaine* [vɛn]. Par ailleurs, le rapport sémantique entre l'adjectif *vain* et *vanité* n'apparaît pas aussi clairement qu'entre *banal* et *banalité*, parce que le mot *vanité* est applicable aux êtres humains et que l'adjectif *vain* ne s'emploie guère pour qualifier une personne. Pour certains, le mot *vanité* sera donc perçu comme un mot complexe, tandis que beaucoup d'autres le traiteront comme un mot simple en dépit de sa finale en *-ité*. Dans le cas de *vanité*, il est évident que la connaissance de la langue en tant qu'objet externe peut modifier notre intuition et l'usage que nous faisons des mots.

Ces deux exemples nous montrent une chose simple : il n'est pas toujours facile de trancher sur le statut des mots : mot simple ou mot complexe. Cela dépend des intuitions des locuteurs et de la finesse de

l'analyse morphologique. *Parlement* – tout comme *établissement* ou *gouvernement* – peut très bien réintégrer l'ensemble des mots du type *déchirement* mais à condition d'avoir soit une conception plus large, plus abstraite, de la signification du suffixe *-ment*, soit une analyse plus minutieuse des familles de mots, analyse qui prendrait en compte la distinction entre mots construits et mots non-construits (voir plus loin) et qui prendrait en compte également la dimension diachronique.

Faire les exercices 12 et 13

Ces exercices portent sur l'analyse morphologique de mots qui peuvent être perçus comme des mots simples mais dont l'étymologie relève qu'ils s'agit en fait de mots complexes.

2.2. MOTS POSSIBLES ET MOTS ATTESTÉS

Le second point concerne la différence entre morphologie et lexicologie. Ce sont deux domaines bien différents : la lexicologie présuppose la morphologie mais pas l'inverse. La lexicologie, c'est l'étude des mots et des morphèmes de la langue, de leurs propriétés phonologiques, morphologiques, syntaxiques et sémantiques, et également de leur formation et de leur origine. De plus, la lexicologie ne prend pas en compte la flexion des mots, c'est-à-dire la variation des mots en fonction de leur emploi dans la phrase. Cela concerne notamment la variation en nombre et en genre pour les noms et les adjectifs, et la variation en temps et personne pour les verbes.

Point important ici : la lexicologie s'intéresse uniquement aux mots existants, c'est-à-dire aux mots attestés dans l'usage de la langue. Il appartient donc à la lexicologie de se prononcer sur l'existence ou non de certaines formes complexes. Ainsi, il appartient à la lexicologie d'expliquer pourquoi les verbes dérivés *décréer* ou *désinventer* n'existent pas. Pourquoi ne les trouve-t-on pas dans les dictionnaires ? Ces verbes sont parfaitement réguliers sur le plan morphologique : ils sont formés à partir des verbes *créer* et *inventer*, sur le même modèle que *défaire* (*faire*) ou *désarmer* (*armer*). Il s'agit donc là d'un problème typique de lexicologie. En l'occurrence, c'est un problème sémantique qui relève du sens commun ; *créer* et *inventer* décrivent des processus irréversibles. Le verbe *désinventer* existe cependant car il existe au moins une attestation dans l'usage de la langue ; un scientifique a dit un jour à la radio « *On ne peut pas désinventer la bombe atomique* ». Son propos est compréhensible. La morphologie permet de rendre compte de ce néologisme (point de vue du locuteur) et de l'interprétation que nous faisons de ce mot dans l'énoncé (point de vue de l'auditeur).

La morphologie n'a pas à se prononcer sur l'existence ou non des mots dans les dictionnaires ou ailleurs. La morphologie est de nature formelle, elle se contente de dégager les règles de formation des mots complexes notamment. Dans le cas de *décréer* et *désinventer*, elle se contente de dire que ces mots sont possibles. C'est seulement l'usage qui décide si ces mots doivent ou non être pris en compte par la lexicologie. Le mot *décidabilité* (formé à partir de l'adjectif *décidable*) par exemple n'est pas attesté dans la plupart des dictionnaires et pourtant il est d'un usage courant en science et plus particulièrement en épistémologie.

La principale conséquence de cette distinction entre lexicologie et morphologie, est que la morphologie est à même de rendre compte des mots ou des morphèmes qui n'ont pas d'attestation dans la langue. C'est le cas notamment des verbes – possibles mais non attestés – **prisonner* dans *emprisonner* et **ratiser* dans *dératiser*. Ce qui compte donc pour la morphologie, ce sont les règles et non les éléments sur lesquels elles opèrent. Il en sera de même en syntaxe. Ce qui est pertinent pour la syntaxe ce sont les catégories et les fonctions et non pas les éléments eux-mêmes qui forment la phrase et qui relèvent d'un choix de la part du locuteur. Ce qui sera pertinent pour l'analyse syntaxique sera la grammaticalité de la phrase et non son acceptabilité, c'est-à-dire son interprétation sémantique.

En résumé, la morphologie énonce les règles qui expliquent la forme des mots (ce qui est possible) alors que la lexicologie rend compte de l'usage des mots possibles, des mots attestés ; mots fréquents, rares ou très rares.

3. LE MOT

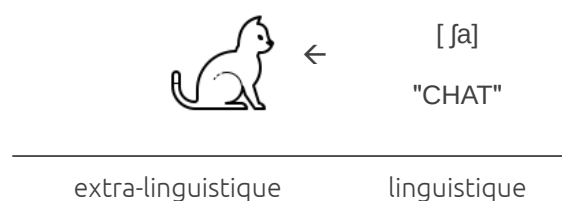
La définition retenue pour le mot est la suivante : un mot est un signe segmental autonome qui ne soit pas divisible en mots. (Mel'cuk, 1993)

Je rappelle tout d'abord qu'une telle définition n'est possible que dans le cadre d'une langue particulière.

Trois points doivent être développés dans cette définition : 1) la notion de signe segmental, 2) la notion d'autonomie du signe, et 3) la notion subsidiaire de divisibilité.

3.1. LE SIGNE LINGUISTIQUE, SENS, SIGNIFICATION ET VALEUR GRAMMATICALE

Depuis Ferdinand de Saussure, fondateur de la linguistique moderne, il est d'usage d'analyser le signe linguistique – on peut à titre provisoire assimiler le signe au mot – comme une unité à double face comportant d'une part un signifiant et d'autre part un signifié. Dans la langue orale, le signifiant correspond à la forme phonique du signe et dans la langue écrite, il correspond à sa forme graphique. Le signifié représente le concept associé au signifiant. À ces deux notions indissociables, il convient d'en ajouter une autre – le référent – qui représente ce à quoi renvoie le signe dans le monde. Il y a donc une ligne de partage entre ce qui est proprement linguistique – le signe, et ce qui est non linguistique : le référent. Cette conception du signe est illustré par le schéma suivant :



Le référent est ici représenté par un dessin qu'il faut interpréter comme un objet réel ou imaginaire, un chat particulier – le chat de ma concierge par exemple – et le signe se présente sous la forme d'une transcription phonétique pour le signifiant, et sous la forme conventionnelle d'une notation en majuscule et entre guillemets pour le signifié. Pour être plus précis, nous devrions remplacer la notation "CHAT" par l'ensemble des propriétés descriptives nécessaires à la caractérisation du concept. On peut à cet effet utiliser la définition du dictionnaire : *animal domestique de la famille des félidés*.

La notion de signifié doit être précisée : dans l'exemple précédent, le signifié renvoie à deux notions distinctes, le sens et la signification, qu'il convient d'articuler correctement par rapport au référent.

Précisons tout d'abord qu'un signe linguistique ne se réduit pas au mot. C'est simplement par commodité que le signe est souvent illustré par un mot, comme dans l'exemple précédent. Tout constituant de la langue fonctionne comme signe. Considérons les exemples suivants :

- a. *Le chat de la concierge était dans l'escalier hier*
- b. *Le chat de la concierge*
- c. *La concierge*
- d. *concierge*

Les quatre énoncés a, b, c et d sont des signes. Le premier est une phrase, le second et le troisième sont des syntagmes nominaux, le dernier est simplement un mot. Le référent de la phrase – ce à quoi elle renvoie dans le monde – est un événement situable dans le temps et dans l'espace. Les référents de b et de c sont respectivement un animal et une personne qu'il est possible également de situer dans l'espace, dans le temps, dans la réalité ou dans une fiction. Par contre le dernier exemple – *concierge* –, ne peut pas être associé directement à un référent. Le signifié des trois premiers énoncés met en jeu une signification et un sens. Le sens, c'est ce qui permet à une entité du monde (événement, objet, être vivant...) d'être représentée par un signe. Nous dirons que *le chat de la concierge* et *la concierge* sont des constituants avec un sens parce qu'il est possible de leur associer un référent. Par contre les simples mots *chat* et *concierge*, n'ont pas de sens, car tels quels, on ne peut pas leur associer de référent ; ils n'ont qu'une signification. On peut seulement associer aux signifiants *chat* et *concierge* un concept ; ce qu'est un ou une concierge, ce qu'est un chat. Par conséquent, les noms communs isolés n'ont pas de sens par eux-mêmes mais seulement une signification. C'est seulement lorsqu'ils sont employés dans la langue qu'ils acquièrent un sens au moyen notamment de la détermination.

Par ailleurs, le déterminant *le* est également un signe, mais son signifié n'est pas assimilable à une signification qui puisse faire l'objet d'une définition dans les mêmes conditions que les mots *chat* et *concierge*. Il n'a ni sens ni signification. Par contre il permet de construire le sens d'un mot à partir de sa signification, en lui donnant une interprétation particulière. L'article *le* permet d'isoler un élément parmi un ensemble d'éléments dont la nature est définie par la signification du nom. On dira alors que *le* à une fonction ou une valeur grammaticale dans l'énoncé – il permet de construire le sens – et le signifié de cet élément correspond à cette valeur grammaticale.

De ce qui vient d'être dit à propos du signe linguistique en général, on peut en déduire pour le mot qu'il a soit un sens, soit une signification, soit une valeur grammaticale. Les trois cas de figure sont illustrés par les exemples suivants :

- a. *hier, Paul, je* (sens)
- b. *chat, manger, joli* : (signification)
- c. *le, de, à* : (valeur grammaticale)

hier, Paul, je, sont des mots qui ont un sens mais pas de signification. Les mots *chat, manger, et joli* pris isolément ont une signification mais pas de sens, et les mots *le, de, à*, ont une valeur grammaticale mais ni sens ni signification. Il faut surtout retenir ici que la plupart des noms communs n'ont par eux-mêmes qu'une signification. Notons au passage qu'un mot grammatical telle qu'une préposition peut combiner à la fois une valeur grammaticale et une signification. Si l'on considère les trois prépositions *de, par* et *dans* ; il y a une gradation entre ces trois prépositions qui va de la simple valeur grammaticale à une combinaison d'une valeur grammaticale avec une signification spatiale. La préposition *par* occupe une position intermédiaire : on peut considérer que sa signification est plus abstraite que celle de *dans* mais qu'elle n'est pas nulle comme dans le cas de la préposition *de*.

Le mot est donc un signe, et la définition donnée précédemment précise qu'il s'agit d'un signe segmental, c'est-à-dire qu'il se présente dans la langue nécessairement comme une unité isolable dans un énoncé et qu'il peut être analysé comme une succession de sons, de syllabes et de morphèmes. Ce point est important car le morphème quant à lui n'est pas nécessairement de nature segmentale (voir plus loin).

Le mot n'est pas la plus petite unité qui fonctionne comme signe dans la langue ; c'est le morphème. Un mot tel que *chanteur* s'analyse en deux morphèmes *chant* et *eur* et chacun d'eux forme un signe. Les mots qui contiennent plusieurs éléments de signification sont des mots complexes, par opposition aux mots simples comme *chat, escalier, hier* qui ne sont pas analysables en signes segmentaux.

3.2. L'AUTONOMIE DU MOT

La notion d'autonomie comporte deux volets. Un mot est autonome parce qu'il est libre et séparable. Dans l'exemple *le chat de la concierge* repris ici, *chat* est un signe segmental libre parce qu'il n'est pas lié par son contexte immédiat ; la présence du segment *chat* n'implique pas la présence du segment *le* à gauche, ni la présence du segment *de* à droite. Comme le montre le découpage, on peut remplacer *le* et *de* par autre chose :

| | | | |
|------------------|--------------------|----------------|---------------------|
| <i>le</i> | <i>chat</i> | <i>de</i> | <i>la concierge</i> |
| <i>plusieurs</i> | | <i>sauvage</i> | |

Si nous appliquons le même raisonnement pour le mot *escalier*, on ne peut pas dire, par exemple, que *escalier* est constitué de deux segments *escal+ier*, car la présence du segment "*escal*" implique celle de "*ier*"

pour former un signe. De plus, il est impossible de traiter ce mot comme un mot complexe formé de deux morphèmes sur le modèle de *vitrier* par exemple, car la relation entre le tout que forme *escalier* et les deux segments "*escal*" et "*ier*" n'est pas assimilable à la relation qui existe entre *vitrier* et les segments *vitri-* (*vitre*) et *-ier*. Pour qu'une segmentation du mot *escalier* soit possible sur le modèle de *vitrier*, il faudrait dans ce cas que l'élément *-ier* ait la même signification (formation d'un nom de métier). *Escalier* est donc un mot simple qui se trouve avoir la même terminaison phonétique et graphique que le mot complexe *vitrier*.

Le mot défini à partir du critère de liberté ne correspond pas nécessairement au mot tel qui se présente à l'écrit. Considérons les deux exemples suivants :

- a. *au fur et à mesure*
- b. *d'ores et déjà*

Dans ces deux exemples *fur* et *ores* ne peuvent pas être considérés comme des mots dans la mesure où *fur* et *ores* ne peuvent être envisagés indépendamment des contextes respectifs *au ___ et à mesure et d' ___ et déjà*. Je reviendrai sur ce problème lorsqu'il sera question de la divisibilité du mot.

Considérons maintenant le cas de l'article, des pronoms personnels associés au verbe et des prépositions dans les contextes suivants :

- a. **le** *chat*
- b. **je** *parle*
- c. *des poils* **de** *chat*

L'article *le*, le pronom personnel *je* et la préposition *de* ne sont pas libres puisque leur présence dans l'énoncé est liée à la présence d'un nom pour l'article et pour la préposition, et à la présence d'un verbe pour le pronom personnel. De plus, aucun de ces mots ne peut apparaître seul dans la langue : l'article et la préposition sont toujours associés à un nom ou à un syntagme nominal pour les prépositions, tandis que les pronoms personnels tels que *je*, *tu*, *il(s)*, *elle(s)*... placés devant le verbe – pronoms clitiques – sont toujours associés à un verbe. Par contre les mots *chat* et *parle* sont bien des éléments libres car ils peuvent apparaître dans un autre contexte ; contexte sans article pour le mot *chat* comme dans l'exemple *des poils de chat*, contexte sans pronom pour le verbe comme dans l'exemple *Pierre parle trop*.

Si nous devions nous en tenir simplement au critère de liberté, appliqué à l'article, au pronom et à la préposition, nous devrions traiter la séquence article + nom, pronom + verbe et préposition + nom comme un seul mot, alors que le nom et le verbe sont traités comme des segments libres par ailleurs. Ce qui est effectivement le cas en roumain pour l'article défini, en espagnol pour les pronoms objet d'une forme verbale impérative, gérondive ou infinitive, et en estonien pour l'équivalent de la préposition *avec* :

- a. *lupul* (roumain) = **le** *loup* (*lup* = *loup*, *-ul* = *le*)
- b. *dádomelo* (espagnol) = (*en*) **me le** *donnant*
- c. *pojaga* (estonien) = **avec** (*le*) *garçon* (*poja* = *garçon*, *-ga* = *avec*)

Lupul correspond à *le loup*, *lup* étant le nom *loup* en roumain. De même en espagnol, ce qui en français correspond à trois mots dans *en me le donnant* est traité dans cette langue comme un seul mot à l'écrit. En

estonien – langue apparentée au finnois – l'équivalent de la préposition *avec* – le segment *-ga* – est incorporé au nom *poja* (*garçon*).

En français, l'article, la préposition et les pronoms clitiques sont bien des éléments autonomes, et il nous faut donc faire appel à un critère supplémentaire pour pouvoir les traiter comme mots à part entière.

Ce critère supplémentaire est celui de la séparabilité. Deux segments quelconques forment deux mots distincts s'il est possible d'insérer entre les deux, un autre mot. Entre l'article et le nom il est possible d'insérer un adjectif : entre *le* et *chat* (exemple a), il est possible d'insérer le mot *petit*. Cette opération étant impossible en roumain, on comprend pourquoi l'article est soudé au nom. Entre le pronom clitique et le verbe il est possible d'insérer la négation (exemple b). Entre la préposition et le nom il est possible d'insérer un déterminant possessif (exemple c) :

- a. *le **petit** chat*
- b. *je **ne** parle pas*
- c. *de **son** chat*

3.3. LES MOTS COMPOSÉS

Les seuls éléments qui posent problème pour une définition du mot fondée sur son caractère autonome sont les mots composés tels que :

- a. *pomme de terre* (**pomme nouvelles de terre*)
- b. *porte-bagages*
- c. *portemanteau*
- d. *ivre mort*

Ces quatre mots illustrent quatre types de mots composés. Ceux qui, sur le modèle de *pomme de terre*, se présentent comme des constructions syntaxiques ; les deux mots sont réunis par le biais d'une préposition et *de terre* est ici un complément de nom ordinaire. Et ceux qui, comme *porte-bagages*, *portemanteau* ou *ivre mort* se présentent sous la forme d'une simple juxtaposition, avec éventuellement une différence de traitement orthographique : certains comme *porte-bagages* ont un trait d'union, certains ne forment qu'un seul segment écrit comme *portemanteau*, et d'autres enfin se présentent simplement comme deux mots indépendants.

Chacun de ces mots composés a une signification particulière et les mots n'y sont pas libres. Lorsque nous parlons du légume, la présence de *pomme* implique la présence de *de terre* et inversement. De plus, ces composants ne sont pas séparables. Il n'est pas possible d'insérer un adjectif ou un article entre *de* et *terre*. On dit : des *pommes de terre nouvelles* et non des **pommes nouvelles de terre*. N'étant ni libres ni séparables les mots composés doivent donc être traités comme des mots uniques. Cette approche est correcte si l'on s'intéresse aux unités de la langue qui sont dotées chacune d'une signification spécifique. D'un autre côté, il est clair que *pomme de terre* met en jeu trois mots, et deux mots pour *porte-bagages*, *portemanteau* et *ivre*

mort. Si l'on s'en tient à la [définition de Mel'cuk donnée en 3](#), les quatre exemples ci-dessus sont bien des formes complexes dans la mesure où chacun d'eux est divisible en mots. Pour chacun d'eux, on reconnaît dans le mot composé plusieurs mots qui existent indépendamment dans la langue. *Pomme* est un mot du français, *de* également et *terre* également. La même chose pour les trois autres exemples. Il y a donc une analyse possible en plusieurs mots bien que les critères de liberté et de séparabilité ne s'appliquent pas pour les mots composés.

Loin d'être une difficulté pour la définition du mot, ce problème montre simplement que nous devons nous accommoder d'une certaine souplesse dans la conception du mot. En fait il faut considérer que la définition du mot varie aussi selon le point de vue adopté. En ce qui concerne les mots composés des quatre exemples ci-dessus, deux points de vue sont en concurrence pour le moment : le point de vue lexical et le point de vue syntaxique. Du point de vue lexical, *pomme de terre* est un mot unique du fait qu'il a une signification propre qui ne se réduit pas à une combinaison des significations des mots qui le composent. Si nous devons dans un texte quelconque relever tous les mots pour constituer un glossaire, il est clair que *pomme de terre* serait traité comme un mot unique (ce mot est également traité comme une entrée lexicale indépendante dans les dictionnaires). Du point de vue syntaxique, *pomme de terre* est ambigu : d'un côté il fonctionne comme une unité simple susceptible de recevoir un déterminant et d'autres expansions propres au nom comme l'adjectif *nouvelles* dans *pommes de terre nouvelles*, d'un autre côté, *de terre* se présente comme un complément de *pomme*. Point important : si le mot composé *pomme de terre* est au pluriel, la marque de pluriel apparaît sur *pomme* et non à la fin du mot composé. Ce type de mot composé est donc du point de vue syntaxique et morphologique une construction complexe.

3.4. LES CLITIQUES

Le problème qui vient d'être évoqué à propos des mots composés et sa conséquence pour la définition du mot n'est pas spécifique à ce type de mots. Il en va de même pour les pronoms clitiques. Les pronoms clitiques sont les pronoms qui apparaissent dans le contexte immédiat du verbe : *je, me, tu, te* etc... Du point de vue syntaxique ce sont des mots à part entière – c'est pour cette raison qu'ils sont traités comme tels à l'écrit en français – mais du point de vue phonologique ce ne sont pas des éléments indépendants. Du fait de cette propriété phonologique, on les appelle aussi pronoms atones. Par définition les pronoms clitiques sont toujours adjoints à un autre mot qui leur sert de support. C'est pourquoi ils ne peuvent jamais apparaître seuls. Ainsi, en réponse à une question qui porte sur le sujet de la phrase, on ne dira pas *je*, mais *moi*. En outre, les pronoms clitiques ne peuvent pas être coordonnés :

a. - *Qui est venu hier ?*

- ***je** / *moi*

b. ***Je et tu** sommes venus hier.

Toi et moi sommes venus hier.

Les pronoms clitiques ne sont pas une particularité du français, on les retrouve dans d'autres langues à commencer par les langues romanes. La notion de clitique est plus générale et ne vaut pas seulement pour les pronoms. Par exemple, en latin, il existe une conjonction de coordination clitique équivalente à *et* :

*Terra mari**que***

(sur terre **et** sur mer)

Un clitique est donc un mot syntaxique mais pas un mot phonologique. La conséquence la plus immédiate est qu'ils ne peuvent pas avoir d'accent.

3.5. LE MOT ET LES NIVEAUX D'ANALYSE

En conclusion, ce qui correspond à un mot varie d'une langue à une autre, et pour une langue donnée, la notion de mot varie également en fonction du niveau d'analyse. Il convient donc de distinguer le mot phonologique (qui se définit notamment par la distribution de l'accent de mot), le mot lexical (possédant une signification propre) et le mot syntaxique qui définit un constituant de la phrase. À ces trois points de vue, on ajoutera le mot morphologique défini comme une unité contenant un ou plusieurs morphèmes :

- a. mot phonologique
- b. mot lexical
- c. mot syntaxique
- d. mot morphologique

Comme on vient de le voir à propos des mots composés et des clitiques, l'analyse en unités plus petites ne coïncide pas d'un niveau à un autre. J'illustrerai à nouveau ce point par deux exemples. Le premier exemple concerne les mots complexes :

- a. mot phonologique : *chanteur* /ʃātœr/
- b. mot morphologique : *chanteur* /ʃāt +œr/
- c. mot syntaxique : *chanteurs*

Un mot complexe tel que *chanteur* se découpe au niveau phonologique en deux syllabes *chan* + *teur* alors qu'au niveau morphologique son découpage en morphèmes est différent ; *chant* + *eur*. Sur le plan syntaxique, *chanteur* est une unité, il aura par conséquent exactement les mêmes propriétés qu'un nom simple. La marque du pluriel (en gras) se placera à la fin du mot.

Le second exemple met en opposition le niveau phonologique et le niveau syntaxique :

- a. mot syntaxique : *(les) enfants*
- b. mot phonologique : *les enfants* /lɛ+zã+fã/

Le groupe *les enfants* contient deux constituants syntaxiques qui s'accordent en genre et en nombre. Ces deux mots sont séparables. Au niveau phonologique, ce groupe ne forme qu'une seule unité accentuelle dans laquelle le découpage syllabique montre que la consonne de liaison appartient à la syllabe initiale du

mot *enfant* et non à la syllabe précédente de l'article. Ce fait est mis en évidence par la liaison sans enchaînement : on dira plus volontiers "*le-zenfants*" que "*lez-enfants*".

3.6. LA FORME ÉCRITE DES MOTS

La forme écrite des mots fournit une assez bonne indication sur la segmentation des énoncés en mots. Mais comme pour le rapport entre l'écriture et la prononciation, il convient de prendre garde cependant à une certaine part d'arbitraire dans la langue écrite. Pour s'en rendre compte il suffit de considérer comment sont réalisés à l'écrit les mots composés – on a déjà vu quelques exemples en [3.3. Les mots composés](#). Ce qui vaut pour l'analyse phonologique, vaut pour l'analyse morphologique ; de même que la forme écrite reflète plus ou moins bien la prononciation des mots, la forme écrite des mots n'est pas toujours en accord avec l'analyse morphologique. L'adéquation est cependant bien meilleure.

Un exemple particulièrement intéressant pour illustrer le fait que l'écriture peut néanmoins être un bon indicateur pour l'analyse morphologique est la différence écrite entre les verbes *s'enfuir* et *s'en aller*. Si on laisse de côté le pronom réfléchi, le premier verbe s'écrit en un seul mot alors que le second s'écrit en deux mots. Ce qui est en jeu ici c'est précisément le critère de la séparabilité. *S'en aller* s'écrit en deux mots bien qu'il soit comparable à *s'enfuir* sur le plan de la signification (aspect inchoatif des verbes *aller* et *fuir*), parce qu'il existe une construction où il est encore d'usage d'insérer un mot entre *en* et *aller*. On a ainsi dans la langue deux formes concurrentes dont l'une appartient à la langue soutenue alors que la seconde est fréquemment réalisée à l'oral :

- a. Il s'en **est** allé de bonne heure.
- b. Il s'**est** en allé de bonne heure.

4. LES MORPHÈMES

Revenons maintenant à la seconde définition de la morphologie. Contrairement au mot, le morphème ne pose pas de problème de définition et cette définition est universelle. Cette définition est donnée en 5. On peut très bien faire de la morphologie en utilisant uniquement la notion de morphème et sans faire appel à la notion de mot. Un énoncé quelconque peut donc s'analyser non pas en une succession de mots mais en une succession de morphèmes. Pour dégager les morphèmes contenus dans la phrase *le chat mangeait lentement* nous devons procéder à la segmentation et à sa justification par commutation :

| | | | | | | | |
|---------------|-----------|--------------|---------------|-------------|----------------|--------------|------------------------------------|
| | <i>le</i> | <i>chat</i> | <i>mange-</i> | <i>-ait</i> | <i>lente-</i> | <i>-ment</i> | |
| <i>le</i> | <i>un</i> | | | | | | <i>un chat mangeait lentement</i> |
| <i>chat</i> | | <i>chien</i> | | | | | <i>le chien mangeait lentement</i> |
| <i>mange-</i> | | | <i>march-</i> | | | | <i>le chat marchait lentement</i> |
| <i>-ait</i> | | | | <i>-ra</i> | | | <i>le chat mangera lentement</i> |
| <i>lente-</i> | | | | | <i>rapide-</i> | | <i>le chat mangeait rapidement</i> |
| <i>-ment</i> | | | | | | ∅ | |

La segmentation vise à isoler et à identifier les différents morphèmes de l'énoncé : phrase, morceau de phrase, ou mot. Dans cette phrase, nous comptons 4 mots et 6 morphèmes. Pour isoler les morphèmes nous utilisons l'opération de commutation pratiquée en phonologie pour l'identification des phonèmes. J'en rappelle ici le principe appliqué à l'analyse morphologique : il s'agit d'isoler les segments afin que chaque segment corresponde à un morphème (segment ayant un signifiant et un signifié), et que chaque segment puisse être remplacé par un autre segment appartenant au même paradigme. Il faut que le nouvel énoncé obtenu donne lieu à une séquence bien formée qui ne se différencie du premier énoncé que par le segment testé. Dans le tableau ci-dessus, chacun des morphèmes de la phrase est remplacé par un autre ayant une signification différente. Seul le dernier – le morphème adverbial – ne peut pas être remplacé par un autre car il n'existe pas d'autre morphème adverbial en français. Le fait qu'il s'agit bien néanmoins d'un segment indépendant est justifié par la première partie du mot – l'adjectif *lente* – qui entre en commutation avec *rapide*.

4.1. LES MOTS ET LES MORPHÈMES

L'opération de segmentation doit être justifiée de la même manière dans le mot. Quelques exemples :

a. *refaire*

| | | | |
|---------------|------------|----------------|-----------------|
| | <i>re-</i> | <i>-faire</i> | |
| <i>re-</i> | dé- | | défaire |
| <i>-faire</i> | | -coudre | recoudre |

b. *souhaitable*

| | | | |
|-----------------|-----------------|--------------|--------------------|
| | <i>souhait-</i> | <i>-able</i> | |
| <i>souhait-</i> | néglige- | | négligeable |
| <i>-able</i> | | -er | souhaiter |

c. *chantons*

| | | | |
|---------------|----------------------|-------------------|------------------------|
| | <i>chant-</i> | <i>-ons</i> | |
| <i>chant-</i> | <i>mange-</i> | | <i>mangeons</i> |
| <i>-ons</i> | | <i>-ez</i> | <i>chantez</i> |

Le mot *refaire* s'analyse en deux morphèmes *re-* et *faire* et ce découpage se justifie par le fait qu'on peut remplacer chacun des deux morphèmes par un autre ayant une signification différente : *re-* commute avec *dé-* pour donner *défaire*, et *faire* commute avec *coudre* pour donner *recoudre*.

Le mot *souhaitable* s'analyse en deux morphèmes *souhait-* et *-able* : *souhait-* commute avec *néglige-* pour donner *négligeable*, et *-able* commute avec le morphème d'infinitif *-er* pour donner *souhaiter*.

Le mot *chantons* s'analyse en deux morphèmes *chant-* et *-ons* : *chant-* commute avec *mange-* pour donner *mangeons*, et *-ons* commute avec *-ez* pour donner *chantez*.

Pour des raisons de commodité – pour ne pas alourdir la description – l'opération de segmentation et de commutation proposée dans ces exemples se fait à partir de la forme écrite des mots. Si l'on voulait être rigoureux, il faudrait la faire à partir de la transcription phonologique.

Dans ces trois mots, les deux morphèmes qui composent chacun des mots ne sont pas à mettre sur le même plan. Nous allons donc distinguer dans un premier temps les morphèmes libres et les morphèmes liés, puis les radicaux, les bases et les affixes.

4.2. MORPHÈMES LIBRES ET MORPHÈMES LIÉS

On appelle morphèmes liés, les morphèmes qui n'ont pas d'autonomie dans la langue. Soit le mot complexe *chanteur* : ce mot est formé de deux morphèmes *chant-* et *-eur*. *-eur* est un morphème lié car il ne peut pas apparaître seul dans la langue, il doit obligatoirement être associé à un autre morphème, un morphème verbal en l'occurrence. Un morphème qui n'est pas lié est un morphème libre. Les morphèmes libres sont autonomes dans les énoncés et peuvent être employés seuls. *chant-* dans *chanteur* est un morphème libre puisqu'il peut apparaître tel quel dans un énoncé où il fonctionne alors comme forme verbale conjuguée = *chante* (présent, 1ère personne et 3ème personne du singulier). C'est sa forme phonologique (/ʃɑ̃t/) qui compte ici et non sa représentation morphologique tronquée (*chant-*). La segmentation en morphèmes devant opérer à partir d'une représentation phonologique des mots, c'est bien le morphème verbal *chant-* qui est identifié, non le morphème nominal *chant* comme le suggère le découpage sur la forme orthographique du mot. La même chose pour le mot *beauté*. Ce mot est formé d'un morphème libre, *beau*, et d'un morphème lié *-té*.

| | | | | | |
|----------------------|------------------|---------------|--------------------|------------------|---------------|
| a. <i>chanteur</i> = | chant- /ʃɑ̃/ | -eur /tœr/ | b. <i>beauté</i> = | beau- /bo/ | -té /te/ |
| | libre radical | lié affixe | | libre radical | lié affixe |

Faire l'exercice 1

Cet exercice porte sur la segmentation de mots en morphèmes.

4.3. RADICAL, BASE ET AFFIXES

Dans les exemples [4.1. Les mots et les morphèmes](#), nous avons des mots dérivés (*refaire* et *souhaitable*) ; c'est-à-dire des unités lexicales obtenues à partir d'autres mots ; respectivement *faire* et *souhait(er)*. Ces morphèmes libres sont des radicaux dans les mots dérivés et ce qui a été ajouté pour former un autre mot est un affixe. De même pour la forme conjuguée *chantons*. Elle est obtenue par affixation au radical du verbe *chant(er)*, du morphème de conjugaison *-ons* (1^{ère} personne du pluriel).

Considérons maintenant les mots dérivés suivants qui contiennent plus de deux morphèmes :

- a. *nationaliser* → *nation+al+is+er*
- b. *nationalisation* → *nation+al+is+ation*
- c. *dénationalisation* → *dé+nation+al+is+ation*

Le radical commun est le nom *nation*. Les affixes sont les suivants :

- ◆ *-al* est un affixe adjectival qui permet de former un adjectif à partir d'un nom.
(*nation* → *national*)
- ◆ *-is* est un affixe verbal qui permet de former un verbe à partir d'un adjectif.
(*national* → *nationalis(er)*)
- ◆ *-er* est un affixe infinitival qui permet de former un infinitif à partir d'un verbe.
(*nationalis* → *nationaliser*)
- ◆ *-ation* est un affixe nominal qui permet de faire un nom à partir d'un verbe.
(*nationalis* → *nationalisation*)
- ◆ *dé-* est un affixe qui permet de faire un verbe à partir d'un autre verbe.
(*nationaliser* → *dénationalis(er)*)

Chacun de ces affixes peut entrer en relation avec une forme de départ qui est soit un mot simple soit un mot complexe. Un affixe se caractérise par la nature catégorielle de l'unité – simple ou complexe – qui entre en construction avec lui. Cette unité est appelée une base. On peut ainsi reformuler les descriptions précédentes en utilisant la notion de base :

- ◆ *-al* est un affixe adjectival qui a pour base un nom.
- ◆ *-is* est un affixe verbal qui a pour base un adjectif.
- ◆ *-er* est un affixe infinitival qui a pour base un verbe.
- ◆ *-ation* est un affixe nominal qui a pour base un verbe
- ◆ *dé-* est un affixe qui a pour base un autre verbe.

Pour un affixe donné, sa base peut être simple ou complexe. Quelques exemples :

| | BASE SIMPLE | BASE COMPLEXE |
|---------------|---|--|
| <i>-al</i> | <i>national</i> = <i>nation</i> + <i>al</i> | <i>gouvernemental</i> = [<i>gouverne</i> + <i>ment</i>]+ <i>al</i> |
| <i>-is</i> | <i>fragilis(er)</i> = <i>fragil</i> + <i>is(er)</i> | <i>nationaliser</i> = [<i>nation</i> + <i>al</i>]+ <i>is(er)</i> |
| <i>-er</i> | <i>créer</i> = <i>cré</i> + <i>er</i> | <i>nationaliser</i> = [<i>nation</i> + <i>al</i> + <i>is</i>]+ <i>er</i> |
| <i>-ation</i> | <i>création</i> = <i>cré</i> + <i>ation</i> | <i>nationalisation</i> = [<i>nation</i> + <i>al</i> + <i>is</i>]+ <i>ation</i> |
| <i>dé-</i> | <i>défaire</i> = <i>dé</i> + <i>faire</i> | <i>dénationaliser</i> = <i>dé</i> + <i>[nation+is+er]</i> |

À la différence de la base, un radical est toujours un morphème unique. Par ailleurs, un mot complexe – hormis les mots composés – a un seul radical, mais peut avoir plusieurs bases ; autant qu'il y a d'affixes. Le mot complexe *dénationalisation* a pour radical *nation* et chaque affixe est en construction avec une base différente. Ces bases sont *nation* (= radical) pour l'affixe adjectival *-al*, *national*, pour l'affixe verbal *-is*, *nationalis-* pour l'affixe nominal *-ation*, et *nationalisation* pour l'affixe *dé-*.

Faire les exercices 5, 6, 7, 8, 9, 10 et 11

Ces exercices portent sur l'analyse de mots dérivés ; un même suffixe peut avoir des significations différentes selon la nature catégorielle de la base avec laquelle ils entrent en construction.

4.4. LES AFFIXES

Les affixes sont de plusieurs types selon la place qu'ils occupent par rapport à leur base :

- ◆ préfixes
- ◆ suffixes
- ◆ infixes
- ◆ circonfixes
- ◆ suprafixes

4.4.1. LES PRÉFIXES

Les préfixes se placent devant la base :

| | | |
|-----------------|--------------|---|
| FRANÇAIS | <i>dé-</i> | <i>défaire</i> (<i>dé+faire</i>) |
| | <i>in-</i> | <i>inavouable</i> (<i>in+avouable</i>) |
| | <i>anti-</i> | <i>anticorps</i> (<i>anti+corps</i>) |
| | <i>post-</i> | <i>postopératoire</i> (<i>post+opératoire</i>) |
| ANGLAIS | <i>un-</i> | <i>unhappy</i> (<i>un+happy</i>) |
| | <i>pre-</i> | <i>preview</i> (<i>pre+view</i>) |
| | <i>re-</i> | <i>reanalysis</i> (<i>re+analysis</i>) |
| GALLOIS | <i>ech-</i> | <i>echnos</i> = "nuit d'avant" : <i>ech</i> "avant" + <i>nos</i> "nuit" |
| GÉORGIEN | <i>v-</i> | <i>vc'er</i> = "j'écris" : <i>v-</i> "1 ^{ère} personne" + <i>c'er</i> "écrire" |
| TZOTZIL | <i>x-</i> | <i>xng</i> = "ma maison" : <i>x-</i> "1 ^{ère} personne du singulier" + <i>na</i> "maison" |

4.4.2. LES SUFFIXES

Les suffixes se placent derrière la base :

| | | |
|-----------------|--------------|--|
| FRANÇAIS | <i>-able</i> | <i>acceptable</i> (<i>accept+able</i>) |
| | <i>-ons</i> | <i>chantons</i> (<i>chant+ons</i>) |
| | <i>-ment</i> | <i>lentement</i> (<i>lente+ment</i>) |
| | <i>-iste</i> | <i>gréviste</i> (<i>grév+iste</i>) |

| | | |
|-----------------|------|--|
| ANGLAIS | -s | <i>books</i> (<i>book</i> +s) |
| ITALIEN | -ate | <i>parlate</i> (<i>parl</i> + <i>ate</i>) "vous parlez" |
| | -eto | <i>uliveto</i> = "oliveraie" : <i>ulivo</i> "olive"+ <i>ate</i> (collectif) |
| TURC | -luk | <i>oburluk</i> = "gourmandise" : <i>obur</i> "gourmand(e)" + <i>luk</i> (propriété) |
| | -cik | <i>evcik</i> = "maisonnette" : <i>ev</i> "maison" + <i>cik</i> (diminutif) |
| | -mek | <i>kesmek</i> = "couper" : <i>kes</i> - "coup(er)" + <i>mek</i> (infinitif) |
| | -im | <i>seçim</i> = "élection" : <i>seç</i> - "choisir" + <i>im</i> (action) |
| JAPONAIS | -te | <i>kakite</i> = "écrivain" : <i>kak(i)</i> - "écrire" + <i>te</i> (agent) |
| | -ru | <i>taberu</i> = "mange" : <i>tabe</i> - "manger" + <i>ru</i> (présent) |
| YOLOF | -u | <i>uatu</i> = "se raser" : <i>uat</i> "raser" + <i>u</i> (réfléchi) "se raser" |

4.4.3. LES INFIXES

Les infixes prennent place dans la base. En arabe, le radical verbal est identifié par trois consonnes (KTB = *écrire*, KKM = *parler*, FTH = *ouvrir*, HBB = *aimer*...). La dérivation verbale et la variation aspectuelle se fait par infixation. Quelques exemples à partir du radical KTB = *écrire* :

KiTaBun : *livre*
KuTuBun : *livres*
KāTiBun : *écrivain*
maKTaBun : *bureau*
KāTiBun : *écrivain*
KaTaB : *écrire*
KuTiB : *a été écrit*

Dans les exemples qui suivent, les infixes ne se présentent pas sous une forme discontinue comme en arabe, mais comme une syllabe supplémentaire (ou parfois uniquement une consonne comme en lituanien) qui vient s'insérer dans la structure syllabique du radical :

| | | |
|----------------|-----------|---|
| TAGALOG | <i>um</i> | <i>tumakbo</i> = "courir" : <i>takbo</i> radical de "écrire" + <i>um</i> (infinitif) |
|----------------|-----------|---|

| | | |
|------------------|-----------|--|
| BALANGAO | <i>in</i> | <i>?inib?a</i> = "compagnons" : <i>ib?a</i> "compagnon" + <i>in</i> (pluriel) |
| LITUANIEN | <i>n</i> | <i>sninga</i> = "il neige" : <i>snig-</i> "neiger" + <i>n</i> (présent) |
| KAMHMU | <i>sr</i> | <i>srnee</i> = "vrille" : <i>see</i> "perforer" + <i>sr</i> (instrument) |

4.4.4. LES CIRCONFIXES

Les circonfixes ont un statut très controversé dans la plupart des théories linguistiques en morphologie. A priori, un circonfixe est un affixe qui se place de part et d'autre de la base. Il s'agit par conséquent d'un morphème discontinu. Le statut des circonfixes est problématique car il est toujours possible d'analyser ce type de segment discontinu comme la conjonction d'un préfixe et d'un suffixe (on parle alors de dérivation parasynthétique). Exemples possibles mais discutables :

| | | |
|-----------------|----------------|--|
| ALLEMAND | <i>ge...t</i> | <i>gemacht</i> : <i>mach-</i> "faire" + <i>ge...t</i> (participe passé) "fait" |
| FRANÇAIS | <i>en...er</i> | <i>emprisonner</i> (<i>prison</i> + <i>em...er</i>) |
| | <i>a...ir</i> | <i>assouplir</i> (<i>souple</i> + <i>a...ir</i>) |

Voyons comment les verbes *emprisonner* et *assouplir*, formés à partir des radicaux *prison* et *souple* peuvent être avantageusement analysés comme faisant intervenir une préfixation et une suffixation plutôt que de faire appel à la notion de circonfixe.

Dans le verbe *emprisonner* le préfixe *em-* et le suffixe d'infinitif *-er* peuvent être traités comme un circonfixe dans la mesure où la présence du premier segment implique la présence du second, et inversement : le mot **emprison* n'existe pas, de même que le verbe **prisonner*. La même chose pour *assouplir*, l'adjectif **assouple* n'existe pas, ni le verbe **souplir*. Ce qui apparaît comme étant ici un circonfixe est en réalité une présence simultanée d'un préfixe et d'un suffixe que l'on retrouve dans la langue indépendamment l'un de l'autre.

Considérons les deux autres exemples suivants comparables aux précédents :

| | | |
|-----------------|----------------|--|
| FRANÇAIS | <i>en...er</i> | <i>ensabler</i> (<i>sable</i> + <i>en...er</i>) |
| | <i>a...ir</i> | <i>amaigrir</i> (<i>maigrir</i> + <i>a...ir</i>) |

Ensabler contrairement à *emprisonner* ne pose pas de problème ; on obtient le verbe *sabler* à partir du nom *sable* et on forme ensuite le verbe *ensabler* à partir du verbe *sabler*. De même pour *amaigrir* : *maigrir* est formé à partir de l'adjectif *maigre*, puis *amaigrir* est formé à partir de *maigrir*. La particularité d'*emprisonner* et d'*assouplir* provient du fait que la dérivation fait intervenir un mot qui n'est pas attesté dans la langue.

L'absence des verbes **prisonner* et **souplir* doit être considérée comme un accident dans le lexique du français (le verbe **prisonner* pourrait très bien exister et avoir pour signification "être en prison". Toutes les possibilités offertes par la morphologie ne sont évidemment pas systématiquement exploitées par le lexique. Par conséquent, le cas des verbes *emprisonner* et *assouplir* en français ne doit pas être traité comme dérivés au moyen d'un circonfixe. Lorsque la formation d'un mot met en jeu simultanément un préfixe et un suffixe on parlera de dérivation parasynthétique. Cela montre par ailleurs que l'analyse morphologique d'un mot doit prendre en considération les autres contextes dans lesquels apparaît l'affixe problématique, et qu'il est nécessaire parfois de restituer dans la formation d'un mot des formes morphologiquement possibles mais non attestées dans la langue. Un autre exemple qui cette fois repose sur une attestation : le mot *covoiturage* pourrait également être analysé a priori au moyen d'un morphème discontinu (= *voiture* + *co...age*) ou d'une dérivation parasynthétique (préfixe + suffixe). En fait, ce mot est tout à fait régulier dans sa formation : à partir du nom *voiture*, on obtient le verbe *voiturer*, puis le nom *voiturage* par suffixation, et enfin *covoiturage* par préfixation. Ce qui donne à penser que le mot *covoiturage* comporterait un circonfixe ou serait du type parasynthétique tient au fait que le verbe *voiturer* et le nom *voiturage* ne sont pas d'un usage courant dans la langue. Mais ces mots sont cependant bien attestés dans la langue.

En conclusion, un véritable circonfixe – si cette notion est pertinente – se devrait d'obéir à une définition stricte de constituant discontinu. Cette définition est la suivante : la séquence A ... B constitue un constituant discontinu si, et seulement si, la présence de A implique la présence de B, et la présence de B implique la présence de A.

Dans les exemples du tagalog donnés précédemment, nous avons bel et bien un authentique morphème discontinu de chaque côté de l'infixe (radical verbal = *su...lat* "écrire"). Mais – point important – il s'agissait là d'un radical et non d'un affixe.

Faire l'exercice 4

Cet exercice porte sur un possible morphème discontinu.

4.4.5. LES SUPRAPHIXES

Les suprafixes sont des affixes prosodiques. Ils se placent "au dessus" dans la représentation phonologique. En phonologie, la prosodie couvre l'intensité, la hauteur et la durée. Les langues qui font un emploi distinctif de la prosodie peuvent utiliser l'accent (intensité), les registres ou les tons (hauteur) et l'opposition voyelle brève/voyelle longue ou consonne simple/consonne géminée (durée) également pour la dérivation et pour la flexion des mots. En fait on ne trouve attestés dans les langues que des exemples de flexion avec des suprafixes qui font varier la hauteur. En ngbaka, langue de la République démocratique du Congo à trois registres (haut, moyen et bas), les registres interviennent dans la conjugaison des verbes. Dans l'exemple suivant un registre haut sur la voyelle du radical correspond à un futur, alors qu'un registre bas correspond à un parfait :

| | | |
|--------------------|----------|---|
| NGBAKA | ' = haut | <i>wá</i> = tirer (futur) <i>wa</i> radical de "tirer" + ' (futur) |
| | ` = bas | <i>wà</i> = tirer (parfait) <i>wa</i> radical de "tirer" + ` (parfait) |
| KINYARWANDA | ' = haut | <i>nagíye</i> = "je suis allé" (passé immédiat) <i>nagiye</i> = "je suis allé" (passé) |

4.5. LES RADICAUX COMME MORPHÈMES LIÉS

Il faut prendre garde à ne pas confondre morphème lié et affixe. Un affixe est toujours un morphème lié, tandis qu'un morphème lié n'est pas nécessairement un affixe. Dans les mots dérivés, il y a un radical et un ou plusieurs affixes. Le radical peut être un morphème libre comme dans les exemples donnés précédemment (*chanteur* et *beauté* notamment), mais le radical peut aussi se présenter sous une forme liée. Quelques exemples :

| | | | | |
|-----------------------|---------------|--------------|---------------|----------------|
| a. <i>populaire</i> = | <i>popul-</i> | <i>-aire</i> | <i>peuple</i> | morphème libre |
| | RADICAL | SUFFIXE | <i>popul-</i> | morphème lié |
| b. <i>charnel</i> = | <i>charn-</i> | <i>-el</i> | <i>chair</i> | morphème libre |
| | RADICAL | SUFFIXE | <i>charn-</i> | morphème lié |
| c. <i>anguleux</i> = | <i>angul-</i> | <i>-eux</i> | <i>angle</i> | morphème libre |
| | RADICAL | SUFFIXE | <i>angul-</i> | morphème lié |

Dans les mots, *populaire*, *charnel* et *anguleux*, on ne retrouve pas les morphèmes libres, que sont respectivement *peuple*, *chair* et *angle*. En lieu et place du morphème libre, on a un allomorphe, c'est-à-dire une forme liée, associée dans le lexique à une forme libre. L'allomorphe est une notion comparable à celle d'allophone ou de variante combinatoire en phonologie. *Popul-* est un allomorphe de *peuple* dans *populaire*, *charn-* est un allomorphe de *chair* dans *charnel*, *angul-* est un allomorphe de *angle* dans *angulaire*. Certains mots ont même plusieurs allomorphes. C'est le cas notamment du radical du verbe *aller* :

allomorphes du radical de *aller* : {*al-*, *ir-*, *v-* et *aill-*}

Une précision pour finir cette présentation des différents affixes – précision qui n'est pas inutile car il y a souvent une confusion à propos de la définition même des préfixes et des suffixes (confusion présente même dans des manuels de linguistique). La nature exacte des affixes est déterminée par leur place par

rapport au radical ou par rapport à leur base. Dans le mot *nationalisation* (*nation+al+is+ation*) dont le radical est le morphème libre *nation*, il y a trois affixes : *al*, *-is* et *-ation*. Tous les trois sont suffixes dans la mesure où ils sont placés après le radical ou après leur base. Les morphèmes *-al-* et *-is* ne sont pas des infixes en vertu du fait qu'ils apparaissent à l'intérieur du mot. De même, *-ation* n'est pas un suffixe en vertu de sa place dans le mot. C'est un suffixe parce qu'il est placé après le radical *nation* ou mieux encore – parce qu'il est placé après sa base verbale *nationalis(er)*.

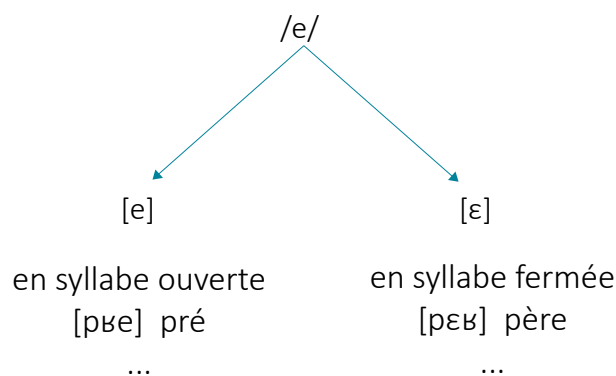
4.6. MORPHÈMES ET ALLOMORPHES

Les mots *populaire*, *charnel* et *anguleux* ont pour radical une forme liée. *popul-*, *charn-* et *angul-* sont des allomorphes d'un morphème libre dans le contexte d'un suffixe particulier. Les morphèmes libres correspondants sont *peuple*, *chair* et *angle*. Le morphème libre *peuple* prend la forme *popul-* lorsqu'il entre en construction avec le suffixe adjectival *-aire*. On retrouve ici la notion d'allophone ou de variante combinatoire de la phonologie, puisqu'un même élément se réalise différemment selon le contexte dans lequel il apparaît. Comme pour les allophones ou variantes combinatoires, il convient de déterminer ce qui est pertinent dans le contexte pour expliquer la présence de l'allomorphe. Dans ce qui suit nous nous intéresserons donc à la distribution des allomorphes.

Je rappellerai tout d'abord le principe même de la variation contextuelle en phonologie et les deux types d'explication des variantes combinatoires.

4.6.1. LA VARIATION CONTEXTUELLE EN PHONOLOGIE

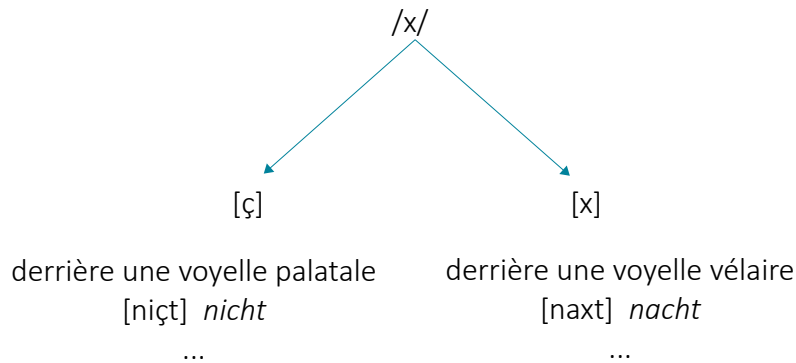
Nous savons qu'en français méridional, il existe un phonème /e/ qui se réalise phonétiquement [e] en syllabe ouverte comme dans *pré* et qui se réalise [ɛ] en syllabe fermée comme dans *père*. Cet exemple de variation contextuelle est représentée par le schéma suivant :



Dans cet exemple, la distribution des deux sons ne repose sur aucune propriété phonétique particulière. Il n'y a pas de relation naturelle entre la différence phonétique d'aperture entre les deux voyelles [e] et [ɛ] et la nature ouverte ou fermée de la syllabe. Il s'agit là d'une distribution strictement phonologique qui

s'appuie sur les particularités du système phonologique de la langue française. Il n'y a donc pas d'explication phonétique pour cette distribution.

Dans l'exemple allemand suivant, c'est différent :



Il existe en allemand un phonème consonantique /x/, qui selon le contexte se réalise phonétiquement soit comme une fricative palatale [ç], soit comme une fricative vélaire [x]. La fricative palatale apparaît dans le contexte droit d'une voyelle palatale, comme dans le mot de négation *nicht* et la fricative vélaire apparaît dans le contexte droit d'une voyelle vélaire, comme dans le mot *nacht* "nuit". Cette fois, la variation contextuelle est d'ordre phonétique, puisque le point d'articulation de la consonne fricative est conditionné par son contexte gauche. La consonne est palatale si la voyelle est palatale et la consonne est vélaire si la voyelle est vélaire. Autrement dit, il s'agit d'une assimilation progressive ; on garde pour la consonne finale le point d'articulation mis en place pour la voyelle. Il y a bien cette fois une explication phonétique.

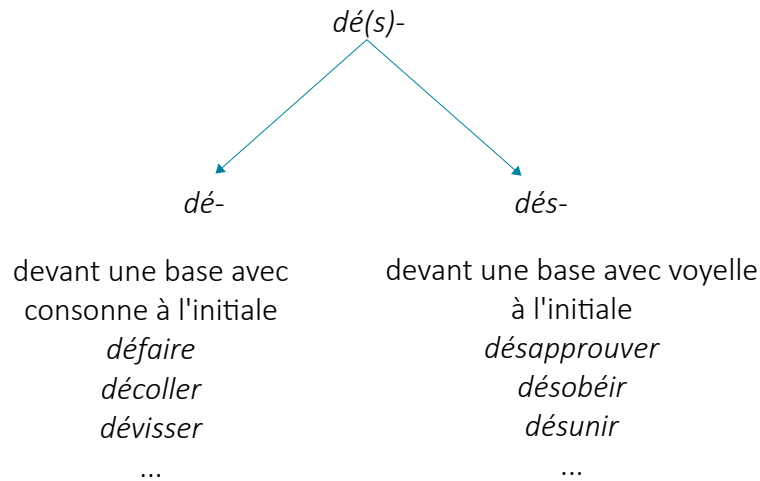
On a ainsi deux types de variantes combinatoires ou contextuelles en phonologie, celles qui à l'instar de l'exemple allemand sont conditionnées par le contexte phonétique, et celles qui à l'instar de l'exemple français sont conditionnées uniquement par la structure phonologique de la langue.

4.6.2. LA VARIATION CONTEXTUELLE EN MORPHOLOGIE

De façon comparable, on doit distinguer en morphologie, les allomorphes qui sont conditionnés par le contexte phonologique et ceux qui sont conditionnés par le contexte strictement morphologique.

Voyons quelques exemples. Les allomorphes se rencontrent aussi bien dans le domaine des radicaux, que dans celui des affixes. On aura donc des allomorphes de radicaux, des allomorphes de préfixes et des allomorphes de suffixes (pour le français). Dans les exemples qui suivent, on a trois types de contextes pour expliquer la présence d'un allomorphe : les deux premiers s'appuient sur des propriétés phonologiques tandis que le troisième est d'ordre strictement morphologique. 1) La présence d'un allomorphe peut être déclenchée par le contexte phonologique. 2) Un allomorphe peut apparaître dans un contexte phonologique particulier sans qu'il y ait pour autant une explication phonologique. 3) La présence d'un allomorphe s'explique uniquement par la présence d'un autre morphème particulier sans que la forme phonologique de celui-ci soit en cause. Ces trois cas de figure sont illustrés respectivement par des exemples avec préfixes, avec suffixes et avec allomorphie du radical.

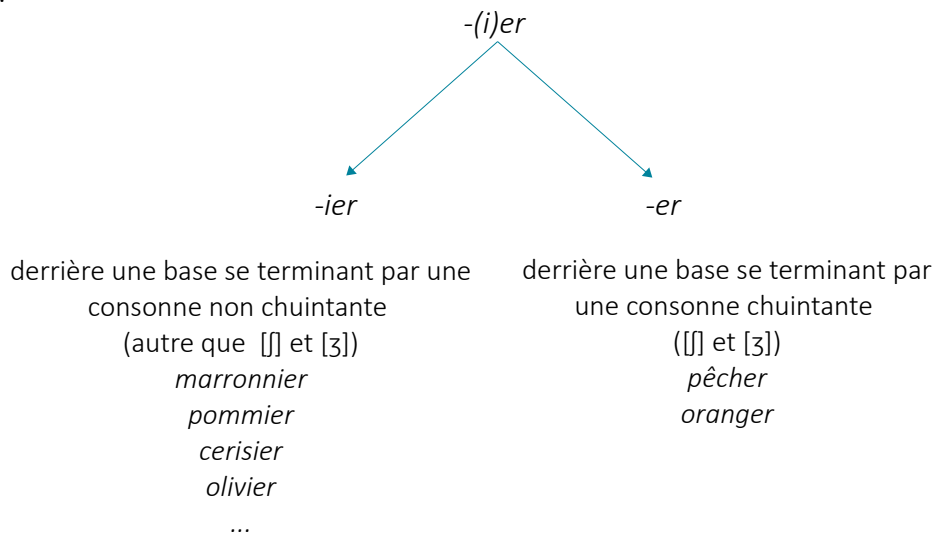
Le premier exemple d'allomorphie concerne le préfixe antonyme *dé-*. Comme le montre le schéma ci-dessous ce préfixe peut se réaliser *dé-* ou *dés-* (on ne tient pas compte ici de la variante orthographique *des-* (*dessaler, dessorganiser*) qui se ramène à *dé-* sur le plan phonologique) :



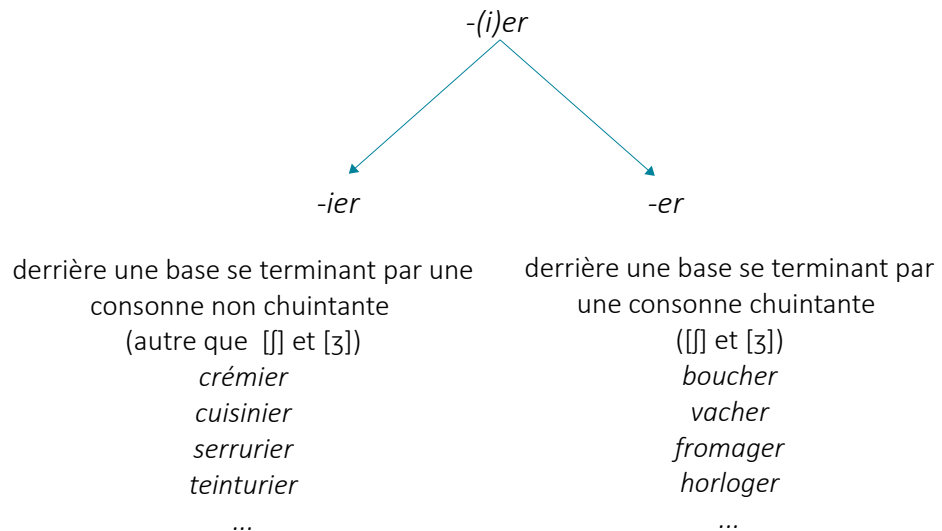
Le morphème *dé(s)-* a deux allomorphes dont la distribution est réglée par ce qui se trouve à l'initiale de la base lexicale. Ce premier exemple simple de variation repose sur une propriété phonologique (distinction entre voyelles et consonnes).

Le deuxième exemple concerne le suffixe *-(i)er* qui permet d'obtenir un nom d'arbre à partir du nom du fruit. Ce suffixe se réalise soit comme *-ier* [je] : *prunier, marronnier, pommier, poirier, olivier...* soit comme *-er* [e] : *oranger* et *pêcher*. Bien que cela ne soit pas évident de prime abord, la forme du suffixe est déterminée par la consonne finale de la base lexicale. Le suffixe *-er* est avec une base qui se termine par une consonne chuintante [ʃ] ou [ʒ]. Dans ce cas la forme du suffixe est liée au contexte phonologique gauche mais contrairement à l'exemple précédent, il n'y a pas d'explication phonétique ou phonologique à cette distribution. La semi-consonne [j] du suffixe *-ier* peut très bien apparaître dans le contexte d'une chuintante comme dans les mots *rangiez* ou *marchiez*. La répartition entre les deux contextes n'est pas évidente dans la mesure où il n'y a pas beaucoup de fruits dont le nom se termine en [ʃ] ou [ʒ]. On pourrait très bien considérer les terminaisons en *-er* pour les noms d'arbre comme des formes irrégulières et non comme une contrainte imposée par la consonne chuintante. Mais cette répartition se vérifie également pour les noms de métier qui se terminent également en *-(i)er*. Si la base lexicale se termine par une chuintante, on a le suffixe *-er*, comme dans *boucher, vacher, fromager, horloger...* alors que dans tous les autres contextes, on a le suffixe *-ier*, comme dans *crémier, cuisinier, serrurier, teinturier...* Bien entendu, il n'y a pas de relation entre le suffixe qui permet de former un nom d'arbre et le suffixe qui permet de faire un nom de métier. Il s'agit donc de deux suffixes homophones. Le point important ici est que ces deux suffixes obéissent à la même contrainte phonologique : tous les deux ont des allomorphes dont la distribution est réglée par la consonne finale de la base lexicale :

Les noms d'arbres :



Les noms de métiers :



Le troisième exemple de variation contextuelle concerne cette fois le radical. C'est dans ce domaine que l'on rencontre le plus d'allomorphe. Revenons sur l'exemple de *peuple* et *populaire*. On a deux séries de mots dérivés, ceux qui sont obtenus à partir du morphème *peuple* (*peuplade*, *peuplement*, *peupler*, *dépeupler*) et ceux qui sont formés à partir du morphème lié *popul-* (*populaire*, *population*, *populace*, *popularité*, *populisme*). Dans ce cas on dira que le morphème lié *popul-* est un allomorphe du morphème libre *peuple*. Mais contrairement aux deux exemples précédents, il n'y a pas d'explication phonologique à cette distribution entre les deux radicaux. On doit se contenter d'enregistrer que le morphème libre prend la forme *popul-* lorsqu'il est en construction avec le suffixe adjectival *-aire*, et avec les suffixes nominaux *-ation*, *-ace*, et *-isme*. C'est donc le suffixe lui-même qui gouverne la forme du radical. La variation est ici strictement morphologique (l'explication doit faire intervenir des considérations historiques sur la formation des mots qui sortent du cadre stricte de la morphologie).

Un autre exemple caractéristique d'allomorphie est présentée dans le tableau ci-dessous. Dans la conjugaison des verbes, certains radicaux comportent plusieurs allomorphes. Il en va ainsi du verbe *aller* qui a

quatre allomorphes, *all-*, *v-*, *i-* et *aill-*. Comme le montrent les différents exemples extraits de la conjugaison du verbe, l'allomorphe *all-* est requis par la présence du morphème d'infinitif, par la présence du morphème d'imparfait, par la présence du morphème de première personne du pluriel au présent, et par la présence du morphème de passé simple. L'allomorphe *v-* est requis par le morphème de présent (seulement à la 1^e, 2^e et 3^e personne du singulier et à la 3^e personne du pluriel), l'allomorphe *ir-* est propre au futur et au conditionnel, et l'allomorphe *aill-* appartient au subjonctif :

| <i>all-</i> | <i>v-</i> | <i>i-</i> | <i>aill-</i> |
|-----------------|-------------|---------------|-----------------|
| <i>aittait</i> | <i>vais</i> | <i>irai</i> | <i>aille</i> |
| <i>aittions</i> | <i>va</i> | <i>iras</i> | <i>aittions</i> |
| <i>aittons</i> | <i>vont</i> | <i>irions</i> | <i>aittiez</i> |
| <i>aittâmes</i> | ... | ... | ... |
| ... | | | |

Faire les exercices 14, 15, 16 et 17

Ces exercices portent sur les allomorphes : allomorphie de l'affixe (préfixe ou suffixe) ou allomorphie du radical.

5. LA STRUCTURE INTERNE DES MOTS COMPLEXES

Il sera question dans ce qui suit de la segmentation des mots complexes dérivés en morphèmes, de la justification de la segmentation et de la représentation formelle des mots complexes.

Nous verrons notamment que les mots complexes doivent s'analyser non pas comme une simple concaténation de morphèmes – c'est-à-dire une simple suite de morphèmes – mais comme une structure hiérarchisée de morphèmes.

Commençons par une comparaison simple entre le niveau phonologique et le niveau morphologique. Dans la description phonologique d'un mot complexe tel que *nationalisation*, on est amené à dégager trois types d'unité : le mot, la syllabe et les phonèmes. Le mot peut être décrit à ce niveau comme étant une suite de syllabes – 6 syllabes pour le mot *nationalisation* :

nationalisation = na+sjo+na+li+sa+sʒõ

et chaque syllabe s'analyse à son tour comme une suite de phonèmes :

nationalisation = n+a+s+j+o+n+a+l+i+s+a+s+j+õ

De façon comparable, on peut analyser ce même mot au niveau morphologique comme une simple suite de 4 morphèmes :

nationalisation = *nation*+*al*+*is*+*ation*

Cette simple comparaison donne à penser que le mot est soumis à une même règle d'organisation linéaire pour les syllabes, et pour les morphèmes. Le mot se présente donc comme une suite de syllabes ou de phonèmes, et comme une suite de morphèmes. On parle dans ce cas de concaténation de syllabes ou de

phonèmes et de concaténation de morphèmes. Nous allons voir précisément que, si cette conception du mot est correcte pour le niveau phonologique, il n'en n'est rien au niveau morphologique, et qu'il convient d'ajouter à l'organisation linéaire des morphèmes, une organisation hiérarchisée.

Pour mettre en évidence la structure interne des mots complexes, une série de mots complexes du français sera utile. Les mots choisis sont des mots qui ne posent pas de problème de découpage en morphèmes, de façon à ce que nous puissions raisonner directement sur la forme écrite du mot. Les mots de cette liste sont tous des mots dérivés secondaires :

2 morphèmes dans : *lenteur* *lent-eur*

fragilité *fragil-ité*

bêtise *bêt-ise*

3 morphèmes dans : *chantions* *chant-i-ons*

nationalité *nation-al-ité*

brutalement *brut-ale-ment*

4 morphèmes dans : *nationalisation* *nation-al-is-ation*

5 morphèmes dans : *dénationalisation* *dé-nation-al-is-ation*

5.1. LA COMMUTATION EN PHONOLOGIE

Avant de voir comment justifier la segmentation en morphème, revenons tout d'abord sur le principe de l'opération de commutation en phonologie, et nous verrons ensuite comment il s'applique ensuite en morphologie. C'est donc par cette opération de commutation que l'on peut dégager en phonologie les paires minimales nécessaires à l'identification des phonèmes. La commutation consiste à substituer un son par un autre à un même endroit de la chaîne sonore. Si la substitution donne lieu à un mot qui existe dans la langue alors les deux sons permutés sont deux phonèmes. On peut représenter l'opération de commutation par un tableau à double entrée dans lequel la suite des sons situés sur l'axe horizontal est reprise sur l'axe vertical. Pour chaque ligne, on procède à une et une seule commutation. Dans l'exemple suivant, l'analyse permet d'identifier trois segments phonologiques dans le mot *bal*, chacun d'eux étant un phonème.

| | | | | |
|----------|----------|----------|----------|-------------|
| | b | a | l | |
| b | m | | | mal (mal) |
| a | | y | | byl (bulle) |
| l | | | z | baz (base) |

[b] commute avec [m] ce qui donne *mal*

[a] commute [y] ce qui donne *bulle*

[l] commute avec [z] ce qui donne *base*

Cette opération nous renseigne sur deux choses : le mot *bal* est formé de trois segments et chaque segment est un phonème puisqu'il entre en rapport d'opposition distinctive avec un autre phonème. Bien entendu, il peut arriver que, pour une position donnée dans la chaîne des phonèmes d'un mot, il ne soit pas possible de pratiquer la commutation parce qu'aucun mot de la langue ne correspond à cette séquence de phonèmes. Cela tient au fait que la langue ne peut pas réaliser toutes les combinaisons de sons possibles. Dans un tel cas, le statut phonologique du segment dégagé sera déterminé par d'autres mots.

5.2. LA COMMUTATION EN MORPHOLOGIE

Cette opération permet également de segmenter la chaîne parlée en morphèmes mais un problème apparaît lorsque nous l'appliquons rigoureusement avec certains mots complexes. Notons tout d'abord que cette opération se fait facilement pour tous les mots qui sont formés de deux morphèmes. On peut ainsi justifier les deux morphèmes dans chacun des mots de la liste précédente. Dans le premier mot *lenteur*, *lente*, la forme féminine de l'adjectif *lent* (il convient de procéder sur la forme prononcée du mot = /lät/), commute avec *grande* pour donner *grandeur*, et le suffixe *-eur* commute avec le suffixe *-ement* pour donner *lentement*.

| | | | | |
|----------------|--------------|---------------|---------------|------------------|
| <i>lenteur</i> | | <i>lent-</i> | <i>-eur</i> | |
| | <i>lent-</i> | <i>grand-</i> | | <i>grandeur</i> |
| | <i>-eur</i> | | <i>-ement</i> | <i>lentement</i> |

La même chose pour les deux autres mots *fragilité* et *bêtise* :

| | | | | |
|------------------|----------------|-------------------|-------------|--------------------|
| <i>fragilité</i> | | <i>fragil-</i> | <i>-ité</i> | |
| | <i>fragil-</i> | <i>tranquill-</i> | | <i>tranquilité</i> |
| | <i>-ité</i> | | <i>-ise</i> | <i>fragilise</i> |

| | | | | |
|---------------|-------------|----------------|--------------|------------------|
| <i>bêtise</i> | | <i>bêt-</i> | <i>-ise</i> | |
| | <i>bêt-</i> | <i>franch-</i> | | <i>franchise</i> |
| | <i>-ise</i> | | <i>-ment</i> | <i>bêtement</i> |

Considérons maintenant les mots de plus de deux morphèmes, comme *chantions*. À nouveau, chacun des morphèmes se justifie simplement par l'opération de commutation. Dans *chantions*, on a trois segments :

chantions

| | | | | |
|---------------|---------------|-------------|-------------|--------------------|
| | <i>chant-</i> | <i>-i</i> | <i>-ons</i> | |
| <i>chant-</i> | <i>lis-</i> | | | <i>lisions</i> |
| <i>-i</i> | | <i>-er-</i> | | <i>chanterions</i> |
| <i>-ons</i> | | | <i>-ez</i> | <i>chantiez</i> |

- ◆ *chant-* commute avec *lis-* pour donner *lisions*
- ◆ *-i* (le morphème de temps) commute avec *er* (le morphème de futur) pour donner la forme verbale du futur *chanterons*
- ◆ *-ons* (le morphème de personne) commute avec *-ez* pour donner la forme verbale de deuxième personne du pluriel *chantiez*.

Considérons maintenant le mot *brutalement*, qui est constitué également de trois morphèmes : *brutalement* contient les morphèmes *brut-*, *-al* et *-ment*. Contrairement à *chantions*, il n'est pas possible de justifier le découpage en trois segments, car on ne peut pas pratiquer l'opération de commutation pour le second morphème *-ale*. *Brut* commute avec *machine* pour donner *machinalement*, *-ment* commute avec *-ité* pour donner *brutalité*, mais *-al* ne commute avec rien. Aucun morphème – pas même le morphème zéro – ne peut remplacer le morphème adjectival dans cette position. Si l'on remplace *-al* par le morphème zéro, nous aurions la séquence **brutment* qui n'existe pas, car en aucun cas un adverbe est formé directement à partir d'un nom.

brutalement

| | | | | |
|--------------|--------------|-------------|--------------|----------------------|
| | <i>brut-</i> | <i>-ale</i> | <i>-ment</i> | |
| <i>brut-</i> | <i>lis-</i> | | | <i>machinalement</i> |
| <i>-ale</i> | | | | |
| <i>-ment</i> | | | <i>-ité</i> | <i>brutalité</i> |

Le problème se situe bien au niveau de la justification de la segmentation et non au niveau de la segmentation. On ne peut pas remettre en cause ici le découpage en trois segments.

Ce problème de justification se retrouve avec le mot *nationalité*, qui est également formé de trois morphèmes :

nationalité

| | | | | |
|----------------|------------------|------------|--------------|----------------------|
| | <i>nation-</i> | <i>-al</i> | <i>-ité</i> | |
| <i>nation-</i> | <i>individu-</i> | | | <i>individualité</i> |
| <i>-al</i> | | | | |
| <i>-ité</i> | | | <i>-isme</i> | <i>nationalisme</i> |

de même pour le mot *nationalisation*, qui contient 4 morphèmes :

| | | | | | |
|----------------|----------------|------------|------------|---------------|------------------------|
| | <i>nation-</i> | <i>-al</i> | <i>-is</i> | <i>-ation</i> | |
| <i>nation-</i> | <i>région-</i> | | | | <i>régionalisation</i> |
| <i>-al</i> | | | | | |
| <i>-is</i> | | | | | |
| <i>-ation</i> | | | | <i>-ait</i> | <i>nationalisait</i> |

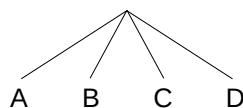
De façon générale, on observe donc qu'au delà de deux morphèmes, il devient impossible, à part quelques exceptions bien choisies comme *chantions*, de justifier la segmentation pour les morphèmes en position médiane. La justification marche pour le premier et le dernier segment, mais pas pour ceux qui sont placés entre les deux. Ainsi dans le mot de 4 morphèmes *nationalisation*, on peut pratiquer l'opération de commutation sur le segment *nation* pour obtenir le mot *régionalisation*, et sur le segment *-ation* pour obtenir la forme verbale *nationalisait*. Mais aucun des segments intermédiaires ne peut être remplacé par quoi que ce soit. Précisons que ce type de problème qu'on rencontre avec l'opération de commutation en morphologie – il en est de même en syntaxe – ne se rencontre pas en phonologie. Il y a par conséquent entre ces deux niveaux d'analyse une différence fondamentale qui est révélée par l'opération de commutation. Est-ce à dire que l'opération de commutation est inopérante en morphologie ? Non, car en réalité, elle doit être pratiquée autrement.

5.3. LES RELATIONS SYNTAGMATIQUES HIÉRARCHISÉES

Précisons tout d'abord l'idée générale qui est à l'origine de la différence entre la phonologie et la morphologie pour ce qui est de l'opération de commutation. Si l'opération de commutation fonctionne bien pour tous les segments phonologiques c'est parce que les mots se présentent et s'analysent comme une succession de phonèmes qui répond au schéma rudimentaire :

...+A+B+C+D+...

où des syllabes ou des phonèmes quelconques d'un mot sont représentés par les lettres A, B, C et D. La relation entre un phonème et un autre phonème est toujours du même type, de même pour les syllabes. Autrement dit la relation syntagmatique entre le phonème A et le phonème B est équivalente à la relation entre le phonème B et le phonème C, et ainsi de suite, de même pour les syllabes. L'arbre suivant illustre cette équivalence syntagmatique où toutes les syllabes ou tous les phonèmes sont sur le même plan et chacun d'eux est en relation uniquement avec ce qui précède et ce qui suit :

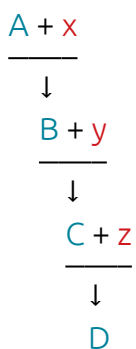


En morphologie le seul schéma possible pour les mots est le suivant (dorénavant, la base en bleu, et l'affixe en rouge) :

A+x

base + affixe

Un mot complexe se ramène toujours à la construction de deux éléments – et uniquement deux – où x représente un affixe et A, une base lexicale. C'est pour cette raison que la segmentation des mots de 2 morphèmes ne pose pas de problème. Si la forme A+x est la seule structure possible pour les mots, quelle est donc la structure interne des mots complexes de plus de deux morphèmes ? Elle est la même bien entendu, mais la structure A+x se reproduit de manière cyclique. Comme le montre le schéma suivant :

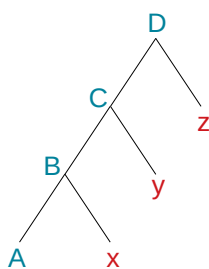


A+x forme alors une nouvelle unité morphologique B, qui va se combiner avec l'affixe y. On a alors la concaténation B+y. Cette nouvelle unité, toujours formée de deux parties va donner lieu à une autre unité C, qui se combine avec un autre affixe z pour former encore une nouvelle unité D.

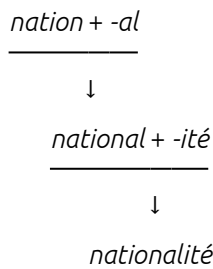
Cette construction cyclique du mot complexe est donnée habituellement sous la forme d'une représentation parenthésisée :

$[D [C [B A+x] +y] + z]$

ou sous la forme d'un arbre :

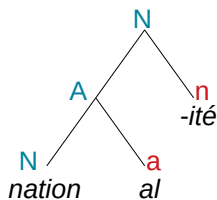


On voit donc que la structure des mots, contrairement à la structure phonologique, n'est pas une simple structure linéaire, mais une structure hiérarchisée où les relations syntagmatiques se font sur des niveaux différents, et à chaque fois il s'agit d'une relation entre deux unités. Prenons par exemple le mot *nationalité* qui est formé de trois morphèmes. Le mot *nationalité* est tout d'abord le produit de la concaténation entre le nom *nation* et le suffixe adjectival *-al* ; ce qui forme l'adjectif *national*. Ensuite cette unité – l'adjectif *national* – est associée au morphème nominal *-ité* pour former un nom *nationalité*.



La représentation parenthésisée et la représentation en arbre font apparaître la segmentation en morphèmes et également l'identité catégorielle des unités formées par l'opération de concaténation. Ce qui permet d'exprimer les propriétés lexicales des affixes en jeu ; à savoir pour le mot *nationalité*, le fait très simple que le suffixe adjectival *-al* se construit avec un nom, tandis que le suffixe nominal *-ité* se construit avec un adjectif.

$[_N [_A [_N \text{ nation}] + \text{al}] + \text{ité}]$



L'opération de suffixation ne correspond pas ainsi à l'ajout d'un suffixe derrière un morphème, mais derrière un mot (base lexicale). *-ité* est suffixé au mot *national* (base lexicale) et non au morphème *-al*. Dit autrement, *-ité* est donc en relation syntagmatique avec le mot *national* et non avec le morphème *-al*, à la différence de la phonologie où chaque phonème est en relation avec un autre phonème, celui de devant ou de derrière.

Analyse du mot complexe *nationalisation* :

segmentation : *nationalisation* = *nation*+*al*+*is*+*ation*

On a tout d'abord la formation de l'adjectif *national* à partir du nom *nation* et du suffixe adjectival *-al*, puis la formation du verbe *nationalise-* à partir de l'adjectif *national* et du suffixe verbal *-ise*, et enfin la formation du nom *nationalisation* à partir du verbe *nationalise* et du suffixe nominal *-ation*.

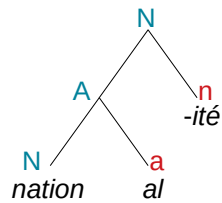
5.4. JUSTIFICATION DE LA SEGMENTATION

Passons maintenant à la segmentation et à sa justification. Si la structure des mots complexes répond au schéma général $A+x$ (base+affixe) la justification doit être opérée par étape de façon à ce qu'à chaque fois on ait seulement deux segments à justifier, et le segment qui se présente comme une unité devra – pour justifier précisément le fait qu'il s'agit d'une unité – commuter avec une unité inanalysable (mot simple).

Reprenons le cas de l'exemple *nationalité* (*nation+al+ité*)

- ◆ représentation parenthésée : $[_N [_A [_N \text{ nation}] + \text{al}] + \text{ité}]$

- ◆ représentation en arbre :



Au premier niveau (première étape de la justification), il s'agit de la construction entre le nom *nation* et le suffixe adjectival *-al*. Le segment *nation* est justifié par la commutation avec *région* pour donner *régional* et le segment *-al* est justifié par la commutation avec le suffixe pluriel *-aux* :

nation+al

| | | | |
|----------------|----------------|-------------|------------------|
| | <i>nation-</i> | <i>-al</i> | |
| <i>nation-</i> | <i>région-</i> | | <i>régional</i> |
| <i>-al</i> | | <i>-aux</i> | <i>nationaux</i> |

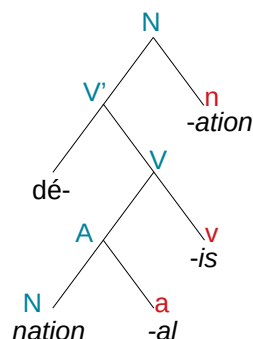
Au second niveau (seconde étape de la justification), on remarquera que la forme complexe *national*, qui fonctionne comme une unité par rapport au morphème *-ité*, commute bien avec un élément simple (pas de segmentation) *tranquille*, qui est un adjectif constitué d'un seul morphème, pour donner le mot *tranquillité*, et le suffixe *-ité* commute avec le suffixe *-isme* pour donner *nationalisme*.

national+ité

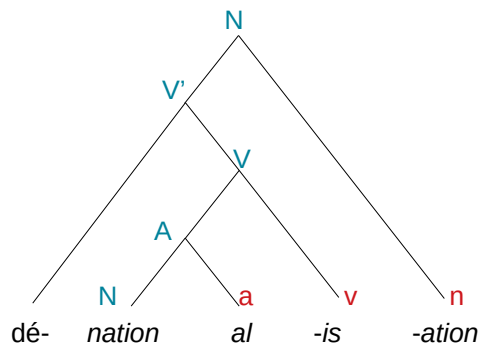
| | | | |
|------------------|-------------------|--------------|---------------------|
| | <i>national-</i> | <i>-ité</i> | |
| <i>national-</i> | <i>tranquill-</i> | | <i>tranquillité</i> |
| <i>-ité</i> | | <i>-isme</i> | <i>nationalisme</i> |

Analyse du mot *dénationalisation*.

- ◆ représentation parenthésée : $[_N [_{V'} \text{dé}+ [_V [_A [_N \text{nation}] +al]] +is]] +ation$
- ◆ représentation en arbre :



Autre représentation plus lisible du mot (avec alignement des morphèmes) :



Justification de la segmentation :

Premier niveau :

nation+al

| | | | |
|----------------|----------------|-------------|------------------|
| | <i>nation-</i> | <i>-al</i> | |
| <i>nation-</i> | <i>région-</i> | | <i>régional</i> |
| <i>-al</i> | | <i>-aux</i> | <i>nationaux</i> |

Le radical *nation* se construit avec le suffixe adjectival *-al* pour former l'adjectif *national*. *Nation* commute avec *région*, et le suffixe *-al* commute avec le morphème pluriel *-aux*.

Deuxième niveau :

national+is

| | | | |
|------------------|-------------------|-------------|---------------------|
| | <i>national-</i> | <i>-is</i> | |
| <i>national-</i> | <i>tranquill-</i> | | <i>tranquillise</i> |
| <i>-is</i> | | <i>-ité</i> | <i>nationalité</i> |

La base lexicale *national* se construit avec le suffixe verbal *-ise* pour former le verbe *nationalise*. La forme complexe *national* commute avec la forme simple *tranquille* pour former le verbe *tranquillise*, et le suffixe *-ise* commute avec le suffixe *-ité* pour former le nom *nationalité*.

Troisième niveau :

dé+nationalis

| | | | |
|--------------------|------------|--------------------|----------------------|
| | <i>-dé</i> | <i>-nationalis</i> | |
| <i>nationalis-</i> | <i>re-</i> | | <i>renationalise</i> |
| <i>-ation</i> | | <i>-faire</i> | <i>refaire</i> |

La base lexicale *nationalise* se construit avec le préfixe verbal *dé-* pour former le verbe *dénationalise*. La forme complexe *nationalise* commute avec la forme simple *faire* pour former le verbe *défaire*, et le préfixe *dé-* commute avec le préfixe *re-* pour former le verbe *renationalise*.

dénationalis+ation

| | | | |
|----------------------|----------------------|---------------|------------------------|
| | <i>dénationalis-</i> | <i>-ation</i> | |
| <i>dénationalis-</i> | <i>libér-</i> | | <i>libération</i> |
| <i>-ation</i> | | <i>-ait</i> | <i>dénationalisait</i> |

La base lexicale *dénationalise* se construit avec le suffixe nominal *-ation* pour former le nom *dénationalisation*. La forme complexe *dénationalise* commute avec la forme simple *libér-* (*libérer*) pour former le nom *libération*, et le suffixe *-ation*, commute avec le suffixe verbal *-ait* pour former la forme verbale conjuguée *dénationalisait*. *Nationalisation* et *libération* sont donc des mots du même type à la différence près que *nationalisation* est dérivée à partir d'une forme complexe, alors que *libération* est dérivée à partir d'une forme simple.

On remarquera qu'il y a autant de niveaux d'analyse qu'il y a d'affixes dans le mot complexe.

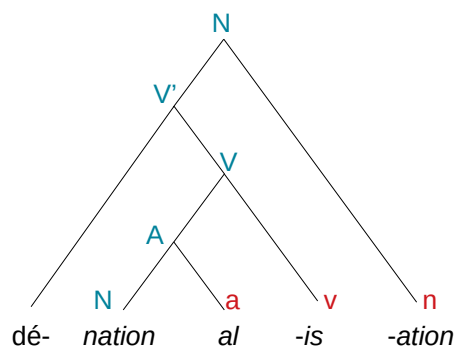
La seule difficulté avec cette nouvelle segmentation réside dans la place du préfixe dans la dérivation. À quel moment, ou plutôt à quel niveau intervient la préfixation au moyen du préfixe *dé-*? Nous savons que ce préfixe permet de créer un verbe à partir d'un autre verbe et que si ce préfixe apparaît avec des noms complexes, comme *dénationalisation* c'est parce qu'il s'agit en fait d'un nom dérivé d'un verbe. Par ailleurs, sachant que la structure des mots complexes est toujours de type A+x, il suffit de tester avec quelle unité morphologique il entre en construction. Comme le montre la série d'exemples suivants :

dé+nation*dé+national**dé+nationalise**dénationalis+ation*

On n'a pas de nom **dénation*, ni d'adjectif **dénational*, par contre on a le verbe *dénationalise*. À ce niveau, *nationalise*, forme complexe, commute avec une forme simple comme *faire* pour donner *dé+faire*. Le préfixe *dé-* se construit toujours avec un verbe.

Nature catégorielle des affixes et des mots complexes

Une représentation en arbre telle que celle du mot *dénationalisation* contient plusieurs informations :



1. Elle indique l'ordre linéaire des différents morphèmes, il y a 5 morphèmes dans ce mot :

dé+nation+al+is+ation.

2. Elle indique la nature catégorielle des morphèmes : *nation* est un nom, *-al* est un suffixe adjectival, *-ise* est un suffixe verbal et *-ation* est un suffixe nominal. La nature catégorielle des affixes est notée au moyen des lettres minuscules. On remarquera que le préfixe *dé-* dans cette représentation n'a pas d'étiquette catégorielle. En fait, ce type d'élément n'a pas de catégorie puisque ce n'est pas lui qui détermine la nature catégorielle du produit A+x.

Considérons les exemples suivants :



Si *lenteur* est un nom c'est parce que *-eur* est un nom (suffixe nominal), si *souhaitable* est un adjectif, c'est parce que *-able* est un adjectif (suffixe adjectival).



Par contre, si *refaire* est un verbe, c'est parce que *faire* est un verbe. Si *impur* est un adjectif, c'est parce que *pur* est un adjectif.

On a donc ici une différence importante entre préfixe et suffixe en français. Lorsqu'il y a un suffixe dans un mot complexe, c'est lui qui détermine la catégorie du mot complexe, alors qu'avec un préfixe c'est la base lexicale qui détermine la catégorie du mot complexe.

Revenons aux informations contenues dans la représentation en arbre. Outre l'ordre des morphèmes et leur nature catégorielle, lorsqu'ils en ont une, on trouve également :

3. Une information sur la nature catégorielle du produit de chaque affixation. Cette information est représentée par les lettres majuscules qui dominent les deux parties du mot à chaque niveau.

Ainsi dans *dénationalisation*, en procédant niveau par niveau, *national* est un adjectif parce que *-al* est un adjectif (suffixe adjectival), *nationalise* est un verbe parce que *-ise* est un verbe (suffixe verbal), *dénationalise* est un verbe parce que cette fois *nationalise* est un verbe et *nationalisation* est un nom parce que *-ation* est un nom (suffixe nominal).

Faire les exercices 18 et 19

Ces exercices portent sur l'analyse morphologique de mots dérivés. Les étapes de l'analyse sont :

1. segmentation du mot en morphèmes
2. restitution des étapes dérivationnelles
3. justification de la segmentation
4. représentation de la structure du mot sous la forme d'un arbre

5.5. BASE ET RADICAL

Dans une représentation en deux dimensions, une représentation arborescente (plus lisible qu'une représentation parenthésisée) permet de visualiser les différentes relations hiérarchisées entre les morphèmes. Le suffixe *-ation* n'est pas au même niveau que le suffixe qui le précède et ainsi de suite. On distingue ainsi plusieurs niveaux d'analyse : 5 dans le mot *dénationalisation*, et pour 4 des 5 niveaux nous avons une concaténation de type *A+x*. Nous pouvons maintenant préciser la différence entre radical et base lexicale. Si l'on considère les exemples précédents, on constate qu'une base lexicale n'est pas nécessairement une forme simple étant donné qu'elle peut contenir elle-même un ou plusieurs affixes. Dans *dénationalisation*, *-ation* est un suffixe et *dénationalise* est une base lexicale complexe par rapport à ce suffixe. De même qu'on peut définir un affixe comme un morphème lié qui entre en construction avec une base lexicale, on peut définir une base lexicale comme la partie du mot, simple ou complexe qui est en construction avec un affixe. Parmi les bases lexicales d'un mot complexe il y en a une qui ne fait l'objet d'aucune analyse, c'est le radical. Dans *dénationalisation* le radical est représenté par le morphème libre *nation*. Un radical n'est pas obligatoirement un morphème libre ; dans un mot comme *populaire*, le radical est représenté par un morphème lié *popul-*.

En résumé, la base se définit par rapport à une forme complexe qui contient un affixe, alors que le radical se définit par rapport au mot dans sa totalité. Le mot *dénationalisation* contient un seul radical (en gras dans l'exemple qui précède), mais il contient plusieurs bases ; une pour chaque affixe. Base et radical coïncident seulement lorsqu'il n'y a qu'un seul affixe. Dans *national*, *nation* est à la fois le radical du mot et la base lexicale du suffixe adjectival *-al*.

6. CONCATÉNATION ET CONCOMITANCE

6.1. PARADIGME MORPHOLOGIQUE ET PARADIGME DE MOTS

Concaténation et concomitance sont les deux principaux processus qui permettent de combiner les morphèmes dans les mots. Tous les exemples de mots complexes que nous avons vus jusqu'à présent mettaient en jeu un processus d'affixation : autour d'une base ou d'un radical, un morphème se présente sous la forme d'un segment qui peut être isolé par l'opération de segmentation. Mais les mots complexes ne sont pas tous de ce type. En français, et de manière générale, on distingue deux processus : la concaténation et la concomitance, les deux pouvant également se combiner. Pour comprendre ce qui sépare ces deux processus, je procéderai à une comparaison de plusieurs paradigmes. Ces différents paradigmes sont à trouver ci-dessous et sont numérotés de 1 à 6. À chaque fois, il s'agit d'un paradigme dans la mesure où tous les mots de chaque paradigme partagent une propriété commune ; il y a un élément de signification commun à l'ensemble des mots du paradigme.

| 1 | 2 | 3 | 4 | 5 | 6 |
|--|-------------------|---------------------|--|--------------|------------|
| <i>normalement</i> | <i>immoral</i> | <i>chanteur</i> | <i>cheval</i> | <i>beau</i> | <i>le</i> |
| <i>fortement</i> | <i>illégal</i> | <i>conducteur</i> | <i>chevaux</i> | <i>belle</i> | <i>la</i> |
| <i>sûrement</i> | <i>irréel</i> | <i>directeur</i> | | | <i>les</i> |
| <i>exactement</i> | <i>inavoué</i> | <i>organisateur</i> | | | |
| ... | <i>impossible</i> | ... | | | |
| | ... | | | | |
| paradigmes morphologiques (ensembles ouverts) | | | paradigmes de mots (ensembles fermés) | | |

- ◆ Dans le premier paradigme, tous les mots contiennent le suffixe adverbial *-ment*.
- ◆ Dans le second paradigme, il s'agit du préfixe négatif *in-*.
- ◆ Dans le troisième paradigme, il s'agit du suffixe *-eur*.
- ◆ Dans le quatrième paradigme l'élément de signification commun est le concept de "cheval".
- ◆ Dans le cinquième, c'est la propriété qui se rapporte à la "beauté".
- ◆ Dans le dernier, c'est la définitude : les trois articles *le*, *la* et *les* expriment la catégorie du défini.

Deux types de paradigmes se dégagent : ceux qui sont construits à partir d'un affixe : ce sont les paradigmes 1, 2 et 3 et ceux qui sont construits à partir d'une propriété commune qui n'est pas identifiable par un segment dans le mot. On parlera de paradigme morphologique pour le premier type et de paradigme de mots pour le second. Notons qu'il y a une différence très nette entre les deux types de paradigmes. Les paradigmes 1, 2 et 3 sont des paradigmes ouverts – les points de suspension dans le tableau signalent cette propriété. Les mots de ces paradigmes sont très nombreux et il serait vain de vouloir en donner une liste exhaustive. Les paradigmes 4, 5 et 6 sont pour leur part des ensembles fermés. Il n'est pas possible d'ajouter d'autres mots sans dénaturer le paradigme. Comme on peut s'en douter, ce qui différencie ces différents paradigmes c'est la segmentation et la forme des morphèmes.

Il y a une gradation du paradigme 1 au paradigme 6. Dans le paradigme 1, la segmentation en morphèmes ne pose aucune difficulté et l'élément commun – le suffixe adverbial *-ment* – est invariablement le même. À l'autre bout, le paradigme 6 contient des mots qui ne peuvent pas faire l'objet d'une segmentation ; il serait absurde de procéder à un découpage aboutissant à isoler la consonne « l » de l'article *le*, *la* et *les* comme le segment correspondant au morphème défini. En fait, aucune segmentation n'est possible dans ce type de mots ; chaque article est un tout non segmentable.

Entre ces deux extrêmes, on a des cas intermédiaires qui relèvent soit d'un paradigme morphologique, soit d'un paradigme de mots. Dans le paradigme 2, la segmentation n'est pas problématique. Mais la forme

du préfixe varie en fonction de la base (préfixe ayant pour allomorphes : *in-*, *im-*, *il-*, *ir-*). Je laisse de côté pour le moment le paradigme 3, j'y reviendrai plus tard. Notons simplement que la segmentation des mots avec le suffixe agentif (*-eur*) est également sans problème. Du côté des paradigmes de mots, les mots contenus dans le paradigme 4 ne peuvent pas être segmentés en deux morphèmes ; le mot *cheval* aurait alors les morphèmes *chev-* et *-al* et *chev-* et *-aux* pour *chevaux*. L'identification du morphème de nombre (*-al* pour le singulier et *-aux* pour le pluriel) n'est pas une difficulté en soi (on retrouve ces morphèmes de nombre dans les mots *normal/normaux*, *artisanal/artisanaux*, *continental/continentaux*... dans lesquels l'adjectif est formé à partir d'un nom *norme*, *artisan*, *continent*...) mais l'autre segment, la partie restante, *chev-* n'est pas identifiable à un morphème car on ne le retrouve pas dans d'autres mots complexes. Dans les mots *chevalin* et *chevalier*, la suffixation se fait sur la base *cheval* et non sur l'hypothétique base *chev-*. Néanmoins, il y a quelque chose de régulier pour le nombre, puisque l'opposition entre les terminaisons *-al* et *-aux* se rencontre dans d'autres mots : *bocal/bocaux*, *amiral/amiraux*, *oral/oraux*. Ce sont là des mots complexes non-construits.

La situation est pire en ce qui concerne le paradigme 5. Ce qui vaut pour le couple *cheval/chevaux* vaut également pour le couple *beau/belle* (la variation porte cette fois sur le genre et non sur le nombre), mais cette fois, il paraît encore plus difficile de considérer la simple consonne "b" de *beau(x)* et *belle(s)* comme un morphème dont la signification serait la part commune aux deux ou quatre mots (selon que l'on prend en compte ou pas les formes pluriels). De plus, l'alternance *-eau/-elle* comme marque du genre grammatical est beaucoup moins productive dans la langue que l'opposition de nombre entre *-al* et *-aux*. En ce qui concerne le dernier paradigme, si la segmentation était possible nous devrions avoir trois segments : un pour la définitude, un pour le genre grammatical, et un pour le nombre, car chacun des mots du paradigme contient ces trois informations. Ces trois informations sont représentées ci-dessous par une matrice de traits morphologiques :

français

$$le = \begin{bmatrix} +\text{DÉFINI} \\ -\text{FÉMININ} \\ -\text{PLURIEL} \end{bmatrix}$$

$$la = \begin{bmatrix} +\text{DÉFINI} \\ +\text{FÉMININ} \\ -\text{PLURIEL} \end{bmatrix}$$

$$les = \begin{bmatrix} +\text{DÉFINI} \\ \pm\text{FÉMININ} \\ +\text{PLURIEL} \end{bmatrix}$$

Avant de commenter les propriétés morphologiques des articles français, considérons tout d'abord le cas plus simple, car plus réguliers, de ces mêmes articles en espagnol :

$$el = \begin{bmatrix} +\text{DÉFINI} \\ -\text{FÉMININ} \\ -\text{PLURIEL} \end{bmatrix}$$

$$la = \begin{bmatrix} +\text{DÉFINI} \\ +\text{FÉMININ} \\ -\text{PLURIEL} \end{bmatrix}$$

$$los = \begin{bmatrix} +\text{DÉFINI} \\ -\text{FÉMININ} \\ +\text{PLURIEL} \end{bmatrix}$$

$$las = \begin{bmatrix} +\text{DÉFINI} \\ +\text{FÉMININ} \\ +\text{PLURIEL} \end{bmatrix}$$

Les quatre articles définis de l'espagnol $\{el, la, los, las\}$ forment un système régulier ; deux articles singuliers $\{el, la\}$ et deux articles pluriels $\{los, las\}$, deux articles masculins $\{el, los\}$ et deux articles féminins $\{la, las\}$. Ce système est représenté par un tableau à double entrée où chaque entrée définit une opposition :

espagnol

| | | GENRE | |
|--------|----------|------------|------------|
| | | -FÉMININ | +FÉMININ |
| NOMBRE | -PLURIEL | <i>el</i> | <i>la</i> |
| | +PLURIEL | <i>los</i> | <i>las</i> |

On laissera de côté ici la question de savoir s'il convient de décrire le genre grammatical au moyen du trait $[\pm FÉMININ]$ ou $[\pm MASCULIN]$, de même que le choix entre les traits $[\pm PLURIEL]$ et $[\pm SINGULIER]$ peut être considéré comme non pertinent ici.

La description des articles de l'espagnol repose sur un système de traits morphologiques dont le nombre est déterminé par 1) la propriété commune à l'ensemble des quatre articles. Ici il s'agit du trait de définitude $[+DÉFINI]$ (le trait $[-DÉFINI]$ caractérise le système des articles indéfinis $\{un, una, unos, unas\}$). 2) les traits oppositifs qui caractérisent les différences entre les quatre articles. C'est un système régulier car toutes les combinaisons des traits oppositifs $[\pm FÉMININ]$ et $[\pm PLURIEL]$ sont représentées dans la langue (la même chose pour les trois traits $[\pm FÉMININ]$, $[\pm PLURIEL]$ et $[\pm DÉFINI]$ et si l'on prend en compte cette fois l'ensemble des articles).

Ce type d'analyse n'est pas sans rappeler la description des phonèmes au moyen de traits phonologiques.

Revenons au français maintenant. La différence entre l'espagnol et le français, est qu'en français il n'y a pas de distinction de genre pour les articles au pluriel. Il y a donc neutralisation de l'opposition de genre (signalée par l'emploi du signe \pm dans la matrice morphologique des traits du français). La description de cet article contient un trait oppositif neutre quant à l'opposition de genre : $[\pm FÉMININ]$. Le système correspondant aux matrices morphologiques des articles définis du français est le suivant :

français

| | | GENRE | |
|--------|----------|------------|-----------|
| | | -FÉMININ | +FÉMININ |
| NOMBRE | -PLURIEL | <i>le</i> | <i>la</i> |
| | +PLURIEL | <i>les</i> | |

6.2. MORPHÈMES ET SEGMENTS

Revenons maintenant plus précisément à la relation entre morphèmes et segments. Plusieurs cas de figure sont à considérer :

- a. 1 morphème = 1 segment (*invisibilité* = *in+vis+ibil+ité*)
- b. 1 morphème = 2 segments (*emprisonner* = *em+prison(n)+er*) ??

- c. plusieurs morphèmes = 1 segment (*le, la, les*)
 d. 2 mots = 1 segment (*au = à+le, du = de+le*)

Dans le premier cas, un morphème correspond à un segment et inversement, un segment correspond à un morphème. Tous les exemples d'analyse de mots complexes que nous avons vus reposaient plus ou moins sur cette relation. Dans le mot *invisibilité*, chaque morphème correspond à un segment et chacun des 4 segments correspond à un morphème.

La discussion à propos des morphèmes discontinus nous a amenés à rejeter qu'il puisse y avoir en français plusieurs segments pour un seul morphème. L'exemple *emprisonner* est donné ici simplement pour évoquer cette problématique et pour laisser ouverte la question de la pertinence des morphèmes discontinus.

L'analyse de l'article défini proposé dans les matrices des traits morphologiques du français et de l'espagnol montre qu'il est possible d'avoir un ensemble de morphèmes associé à un seul segment. Ce type d'analyse n'est nullement incompatible avec la définition du morphème rappelée ici :

Le morphème est la plus petite unité de langue qui soit dotée d'un signifiant et d'un signifié.

Cette définition du morphème implique seulement qu'il y ait une association entre un signifiant et un signifié sans préciser si cette relation doit être une relation bi-univoque, (un élément de signifié unique nécessairement associé à un segment signifiant unique). Cela étant, on débouche sur deux conceptions de la notion de morphème : soit on définit le morphème comme un segment, comme un signe par conséquent, soit on le définit à partir de son signifié sans se préoccuper de sa réalisation dans la langue. Pour éviter l'ambiguïté entre les deux approches je garderai la notion de morphème pour les éléments qui associent un signifié et un segment, et je parlerai de trait morphologique lorsque l'élément de signifié fait partie d'un ensemble de traits associé à un segment.

Le dernier cas de figure est un peu particulier mais mérite néanmoins d'être mentionné dans la mesure où il est assez bien représenté dans les langues : un segment correspond à plusieurs mots. C'est le cas pour les articles contractés *au* et *du* qui sont analysables comme un seul mot sans segmentation en morphèmes possible, mais cet unique segment renvoie sur le plan syntaxique à une séquence de deux mots (ou deux morphèmes libres). Dans ce cas on parle de mot amalgamé.

Revenons aux différents [paradigmes en 6.1](#). Deux types de processus pour l'analyse des mots complexes se dégagent : la concaténation et la concomitance.

Concaténation : Il y a concaténation de morphèmes lorsque ceux-ci se présentent sous la forme d'un enchaînement de segments A+B+C+... (chaque segment renvoie à un seul morphème).

Concomitance : Il y a concomitance de morphèmes lorsque ceux-ci se présentent sous la forme d'un seul segment (un segment pour plusieurs morphèmes). On a alors un segment analysable au moyen d'une matrice de traits morphologiques.

Ces deux processus visent à préciser comment sont réalisés les morphèmes dans la langue. Les paradigmes 1 et 2 reposent sur une concaténation de morphèmes, tandis que le paradigme 6 repose sur une concomitance de traits morphologiques.

Si l'on considère maintenant le paradigme 3, que nous avons laissé de côté précédemment, celui-ci relève à la fois de la concaténation et de la concomitance. Il y a concaténation car il y a bien segmentation en deux segments *chant-* et *-eur*, mais le second segment contient en fait deux morphèmes, deux traits morphologiques. Le segment *-eur* doit être analysé comme ci-dessous sous la forme d'une matrice de deux traits morphologiques : un trait de signification agentive noté [+AGENTIF] et un trait de genre grammatical, en l'occurrence [-FÉMININ]. Étant donné que le suffixe agentif se réalise soit sous la forme masculine *-eur*, soit sous la forme féminine *-euse*, il serait tentant de considérer ces deux formes du suffixe comme des allomorphes. Mais c'est tout à fait impossible, car l'allomorphie est une variation de la forme des morphèmes déterminée par le contexte. Hors dans le cas de la formation du nom d'agent à partir du verbe, rien dans le contexte immédiat du suffixe – c'est-à-dire le verbe lui-même – ne permet de fixer le choix de la forme du suffixe agentif. Il y a donc deux segments *-eur* et *-euse* en concurrence parce qu'il y a deux traits morphologiques associés à chacun d'eux.

$$\textit{chant+ -eur} = \begin{bmatrix} + \text{AGENTIF} \\ - \text{FÉMININ} \end{bmatrix}$$

$$\textit{chant+ -euse} = \begin{bmatrix} + \text{AGENTIF} \\ + \text{FÉMININ} \end{bmatrix}$$

Quant aux paradigmes 4 et 5, ils comportent des mots qui se présentent comme la concomitance d'une signification lexicale et d'un trait morphologique de nombre pour le paradigme 4 et d'un trait de genre pour le paradigme 5. Exemples (le symbole $\sqrt{\quad}$ représente la signification lexicale du mot) :

$$\textit{cheval} = \begin{bmatrix} \sqrt{\text{CHEVAL}} \\ - \text{PLURIEL} \end{bmatrix}$$

$$\textit{belle} = \begin{bmatrix} \sqrt{\text{BEAU}} \\ + \text{FÉMININ} \end{bmatrix}$$

6.3. L'ANALYSE EN TRAITS MORPHOLOGIQUES

Nous allons voir maintenant comment il convient de procéder pour justifier les traits morphologiques. Les exemples que nous analyserons sont ceux de la catégorie de l'article, du pronom personnel de première personne et de l'adjectif possessif de première personne :

| | | | | | |
|----|--------------------------------------|----|--------------------------------------|----|---------------------------------------|
| a. | article défini | b. | pronom personnel | c. | adjectif possessif |
| | <i>le</i> <i>la</i> <i>les</i> | | <i>je</i> <i>me</i> <i>moi</i> | | <i>mon</i> <i>ma</i> <i>mes</i> |

Commençons par l'article défini. On a vu précédemment que la matrice des traits morphologiques était composée de trois traits : un trait de définitude [\pm DÉFINI], un trait de genre [\pm FÉMININ] et un trait de nombre [\pm PLURIEL].

$$le = \begin{bmatrix} +\text{DÉFINI} \\ -\text{FÉMININ} \\ -\text{PLURIEL} \end{bmatrix}$$

$$la = \begin{bmatrix} +\text{DÉFINI} \\ +\text{FÉMININ} \\ -\text{PLURIEL} \end{bmatrix}$$

$$les = \begin{bmatrix} +\text{DÉFINI} \\ \pm\text{FÉMININ} \\ +\text{PLURIEL} \end{bmatrix}$$

Le trait commun à l'ensemble des trois termes est le trait de définitude [\pm DÉFINI], et les traits de nombre et de genre sont des traits oppositifs. Cette répartition systématique entre un trait commun et des traits différentiels se retrouve dans les deux autres paradigmes.

Chacun des traits se justifie par le fait que le mot correspondant entre en relation d'opposition avec un autre mot qui ne se différencie du précédent que par ce trait. La démarche suivie est ainsi exactement la même que celle suivie en phonologie pour déterminer les traits distinctifs des phonèmes.

Dans le cas de l'article *le*, le trait de définitude – trait commun aux trois termes – se justifie par l'existence d'un article qui ne se différencie de *le* que par le trait [+DÉFINI] et qui est par ailleurs identique par les deux autres traits ([-FÉMININ], [-PLURIEL]). Cet article qui a pour matrice morphologique les traits : [-DÉFINI], [-FÉMININ] et [-PLURIEL] est l'article indéfini *un*. On remarquera que le trait commun aux trois termes se justifie par le recours à un terme qui n'appartient pas au paradigme. Les deux autres traits – les traits différentiels – trouvent leur justification dans le paradigme lui-même. Le trait [-FÉMININ] est justifié par l'existence d'un terme qui est [+DÉFINI], [+FÉMININ] et [-PLURIEL], c'est l'article *la*. Le trait [-PLURIEL], se justifie au regard de l'existence de l'article *les* qui est [+DÉFINI], [\pm FÉMININ] et [+PLURIEL].

La justification des traits morphologiques des deux articles *la* et *les* se fait de la même manière ; chaque trait doit entrer en opposition avec un autre trait qui appartient à une matrice équivalente dont les autres traits restent inchangés.

Voyons maintenant le cas du pronom personnel de première personne. Le trait commun est bien entendu le trait de personne. On notera ce trait [1SG] sans signe \pm car ce trait n'entre pas dans le cadre d'une

opposition binaire ; il appartient à l'opposition à 6 termes ([1SG], [2SG], [3SG], [1PL], [2PL] et [3PL]). Pour déterminer quels sont les traits différentiels, il convient de considérer la différence entre *je* et *me*, d'une part et entre *je* et *moi* d'autre part. La différence entre *je* et *me* est une différence de fonction ; le premier est un pronom sujet tandis que le second est un pronom objet (direct ou indirect). On utilisera par conséquent le trait [±SUJET]. La différence entre *je* et *moi*, est une différence entre pronom clitique (dépendant d'un verbe) et pronom non clitique. On codera donc cette différence au moyen du trait [±CLITIQUE]. Les matrices des pronoms de première personne du singulier sont les suivantes :

$$je = \begin{bmatrix} 1SG \\ +SUJET \\ +CLITIQUE \end{bmatrix} \quad me = \begin{bmatrix} 1SG \\ -SUJET \\ +CLITIQUE \end{bmatrix} \quad moi = \begin{bmatrix} 1SG \\ \pm SUJET \\ -CLITIQUE \end{bmatrix}$$

Justifions maintenant la présence des traits du pronom *je* :

Le trait de personne [1SG] se justifie par l'existence du pronom *tu* qui est [2SG], [+SUJET] et [+CLITIQUE] (on aurait pu prendre tout autre pronom ayant les traits [+SUJET] et [+CLITIQUE]). Le trait de fonction [+SUJET] se justifie par l'existence du pronom *me* qui est [1SG], [-SUJET] et [+CLITIQUE]. Le trait [+CLITIQUE] se justifie par l'existence du pronom *moi* qui est [1SG], [±SUJET] et [-CLITIQUE]. Ici il convient de remplacer le trait [-SUJET] par le trait [±SUJET] car le pronom non-clitique *moi* couvre toutes les fonctions (sujet : *Toi et moi viendrons demain*, objet : *Regarde moi!* complément indirect : *Il parle de moi*). Il est donc neutre quant à l'opposition [±SUJET].

Pour l'adjectif possessif de première personne du singulier, je me contenterai de donner la liste des traits morphologiques. Le trait commun est à nouveau le trait de personne [1SG]. Le second trait est le trait de genre [±FÉMININ] et le troisième est le trait de nombre [±PLURIEL]. Pour les adjectifs possessifs, le trait de genre et de nombre correspond, non pas à la personne du possesseur mais au genre et au nombre du nom possédé avec lequel l'adjectif possessif est en relation d'accord.

$$mon = \begin{bmatrix} 1SG \\ -FÉMININ \\ -PLURIEL \end{bmatrix} \quad ma = \begin{bmatrix} 1SG \\ +FÉMININ \\ -PLURIEL \end{bmatrix} \quad mes = \begin{bmatrix} 1SG \\ \pm FÉMININ \\ +PLURIEL \end{bmatrix}$$

Pour la justification des traits morphologiques, je vous renvoie au corrigé de l'exercice correspondant.

Faire les exercices 20 et 21

Ces exercices portent sur l'analyse en traits morphologiques de mots. Chaque trait morphologique doit être justifié par une opposition avec un autre mot du même paradigme.

7. LES TYPES DE MOTS

Deux classifications sont proposées ici :

- ◆ les types de mots lexicaux,
- ◆ les mots construits et non construits.

7.1. LES TYPES DE MOTS LEXICAUX

Pour classer les mots du lexique en différents types, plusieurs notions sont utiles ; certaines ont été définies dans les chapitres précédents, d'autres, sont nouvelles :

- ◆ morphème : distinction entre morphème libre et morphème lié
- ◆ segment : partie du mot qui peut être isolée par l'analyse (un segment peut être un morphème ou non)
- ◆ étymologie : origine des mots

Les mots du lexique se répartissent tout d'abord en mots simples et mots complexes. Les mots simples sont constitués d'un seul morphème libre, exemple :

- ◆ *maison, riche, trop, ne, ...*

Les mots complexes contiennent plusieurs morphèmes. Les mots complexes se répartissent en :

- ◆ mots composés
- ◆ mots dérivés
- ◆ mots savants
- ◆ mots-valises
- ◆ mots amalgamés

7.1.1. LES MOTS COMPOSÉS

Les mots composés sont formés de plusieurs morphèmes libres. Les mots composés font l'objet d'une entrée lexicale dans les dictionnaires. Ils sont formés : 1) par construction syntaxique (*pomme de terre, chemin de fer, pied de biche, ...*), 2) par concaténation de deux mots séparés par un trait d'union (*sourd-muet, libre-service, contre-indiquer, ...*) ou d'une apostrophe (*presqu'île, entr'ouvert, prud'homme, entr'acte, ...*), 3) par concaténation directe (*portemanteau, contredire, ...*). Les mots composés formés par construction syntaxique se distinguent des constructions syntaxiques ordinaires (syntagme) par le fait qu'ils ont une signification propre qui ne se réduit pas à une mise en composition sémantique de chacun des mots. *Livre de cuisine* et *pomme de terre* ont la même construction (N de N) mais le premier est un syntagme tandis que le second est un mot composé. Ces composés sont également appelés synapsie (Benveniste) ou synthème (Martinet). Un *livre de cuisine* est un livre, tandis qu'une *pomme de terre* n'est pas une pomme. Il n'est pas nécessaire d'avoir

une entrée lexicale dans le dictionnaire pour savoir ce qu'est un *livre de cuisine*, alors que *pomme de terre* aura une définition particulière (tubercule...) sans rapport avec la signification de pomme qui est un fruit. Certains de ces composés syntaxiques sont écrits avec des tirets, parfois les deux orthographes sont possibles.

Certains mots, perçus parfois comme des mots simples, sont d'anciens mots composés : *vinaigre* (*vin aigre*), *pissenlit* (*pisse en lit*), *longtemps* (*long temps*, orthographié *long-temps* encore au XIX^e siècle), *cependant* (*ce pendant*), *puisque* (*puis que*)...

7.1.2. LES MOTS DÉRIVÉS

Les mots dérivés sont formés à partir d'une base à laquelle est ajouté un affixe (préfixe ou suffixe). La base lexicale d'un mot dérivé peut être simple ou complexe ; *chanteur* est un mot dérivé à partir de la base verbale simple *chant-* (suffixe *-eur*), *inacceptable* est un mot dérivé à partir de la base adjectivale complexe *acceptable* (préfixe *in-*), *acceptable* est un mot dérivé à partir de la base verbale simple *accept-* (suffixe *-able*). Rappel : dans le mot *inacceptable*, il y a deux bases lexicales (*acceptable* et *accept-*) car il y a deux constructions ; l'un avec un préfixe (*in-*) et l'autre avec un suffixe (*-able*). Le mot *inacceptable* a pour radical le verbe *accept(er)*. Dans le mot *chanteur*, qui ne contient qu'une seule base (car un seul affixe), base et radical sont identiques : verbe *chant(er)*. Dans tous les exemples suivants, on ne retiendra comme exemples de mots dérivés que des mots dérivés avec une base simple ; la base est alors réduite au radical du mot.

Pour les mots dérivés, il peut être utile de faire la distinction entre mots dérivés primaires et mots dérivés secondaires.

Les mots dérivés secondaires sont formés sur le modèle du mot *secondaire* : la base lexicale est un morphème libre. La prononciation de la base *second-* [ʒəɡɔ̃d] est identique à la prononciation de la forme féminine du mot *second/seconde* [ʒəɡɔ̃d]. Ne pas oublier que la dérivation se fait à partir des formes phonologiques des mots et non à partir de leur forme écrite. Autres exemples de ce type : *légitime*, *quatrième*, *doucement*, *séparable*...

Les mots dérivés primaires sont formés sur le modèle du mot *primaire* : la base lexicale est un morphème lié (morphème qui ne s'emploie pas comme mot isolé). La base lexicale de ce mot est *prim-* qui est un allomorphe du morphème libre *premier* (*prime* est en fait une forme ancienne de *premier*, que l'on trouve encore dans l'expression *de prime abord* (= *au premier abord*)).

Exemples de ce type :

- ◆ *populaire* : *popul-* est un allomorphe du nom peuple
- ◆ *charnel* : *charn-* est un allomorphe du nom chair
- ◆ *véridique* : *vérid-* est un morphème lié qui provient d'une racine latine (*verum/veridicus* = vrai/véridique)
- ◆ *déballer* : **baller* est un mot non attesté dans le lexique construit à partir du nom *balle* (= paquet).

7.1.3. LES MOTS SAVANTS

Les mots savants qui relèvent de la néologie (création de nouveaux mots) sont formés par concaténation de deux morphèmes liés qui ont une origine latine ou grecque. Les mots savants n'ont pas de radicaux. Exemples :

- ◆ *télescope* = *télé*+*scope*
- ◆ *géologie* = *géo*+*logie*
- ◆ *polychromie* = *poly*+*chromie*
- ◆ *homéopathie* = *homéo*+*pathie*

télescope est un mot savant mais *télévision* et *télétravail* sont des mots dérivés car le morphème lié *télé* est dans ces deux mots un préfixe ; *vision* et *travail* sont des mots du français (bases lexicales respectives des mots *télévision* et *télétravail*). Le mot *télé*, employé pour *télévision* est un mot simple, forme apocopée.

7.1.4. LES MOTS-VALISES

Les mots-valises sont des mots complexes formés par concaténation de deux parties de mots. La partie tronquée retenue n'est pas un morphème. Ce mode de formation relève également de la néologie.

Exemples :

- ◆ *foultitude* = *fou*~~te~~ + ~~m~~*ultitude*
- ◆ *progiciel* = *pro*~~gramme~~ + *logiciel*
- ◆ *quasar* = *quasi*+ ~~s~~*tar*
- ◆ *cultivar* = *cult*~~ure~~ + *variété*
- ◆ *franglais* = *fran*~~çais~~ + *anglais*

Certains mots-valises présentent une voyelle, une consonne ou une syllabe commune aux deux parties :

- ◆ *motel* = *mo*~~te~~r + ~~h~~*otel*
- ◆ *avionique* = *avio*~~n~~ + *électro*~~n~~ique
- ◆ *cadonner* = *cadeau* + *donner* (africanisme)

D'autres sont formés à partir de deux mots dont un seul est tronqué :

- ◆ *publireportage* = *public*~~ation~~ + *reportage*
- ◆ *distribanque* = *distrib*~~uteur~~ + *banque*

D'autres encore présentent une voyelle épenthétique qui permet de lier les syllabes :

- ◆ *domotique* = *domi*~~cile~~ + *in*~~forma~~tique
- ◆ *cubytainer* = *cube* + ~~con~~tainer

Les mots-valises sont parfois des créations fantaisistes d'écrivains :

- ◆ *parlementeur* = *parlement*~~aire~~ + *menteur*
- ◆ *nauséabondance* = *nauséabond* + *abondance*
- ◆ *caméléopard* = *caméléon* + *léopard*

Dans les exemples tels que *cad**donner*** ou *nausé**abondance***, il n'y a pas de partie tronquée mais plutôt une partie de mot commune aux deux mots, comme si les deux mots à l'origine du mot-valise étaient en partie superposés (partie mise en gras).

7.1.5. LES MOTS AMALGAMÉS

Les mots amalgamés sont peu nombreux. Ils se présentent comme des mots simples – aucune segmentation possible – mais leur analyse repose sur la fusion de deux mots simples. C'est le contexte syntaxique qui détermine leur emploi dans la langue. Exemples :

- ◆ *du* = *de* + *le*
- ◆ *des* = *de* + *les*
- ◆ *au* = *à* + *le*
- ◆ *aux* = *à* + *les*
- ◆ *es* = *en* + *les*

7.2. MOTS CONSTRUITS ET MOTS NON CONSTRUITS

Mot construit : Un mot construit est un mot complexe (plusieurs morphèmes) dont la signification globale est obtenue à partir de la signification de ses morphèmes.

Exemple :

- ◆ **orangeade** = **orange**+ade
-ade = préparation à base de ...
signification du mot dérivé = *préparation à base d'**orange***
- ◆ **persillade** = **persil(l)**+ade
-ade = préparation à base de ...
signification du mot dérivé = *préparation à base de **persil***

Par contre le mot *salade* qui est également un mot dérivé au moyen du suffixe *-ade* (base lexicale = *sel*, *sal* étant un allomorphe de *sel* que l'on retrouve dans *saler*, *salin*, *salpêtre*...), n'est pas un mot construit, car une **salade** n'est pas une préparation à base de sel. Dans la conscience des locuteurs, *salade* est un mot simple dont la signification n'a plus rien à voir avec l'origine du mot (définition du dictionnaire Hachette : *plante potagère*...). Selon le dictionnaire Littré, *salade* a pour définition : « *Mets composé de certaines herbes ou de certains légumes assaisonnés avec du **sel**, du poivre, du vinaigre et de l'huile* ». Deux points de vue se présentent : le premier prend en compte la compétence des locuteurs pour lesquels *salade* est un mot simple ; et donc, un mot non construit. Le second se fonde sur l'origine du mot et dans ce cas, *salade* est un mot construit. Afin de concilier les deux approches, le mot *salade* peut être analysé comme un mot complexe dérivé mais non construit.

7.2.1. MOTS DÉRIVÉS CONSTRUITS ET NON CONSTRUITS

Considérons maintenant l'ensemble des mots suivants : {*chanteur, voleur, réparateur, conducteur, électeur, lugeur, routeur, aviateur, facteur*}. Tous ces mots sont a priori formés par dérivation au moyen du suffixe agentif *-(at)eur*. Tous partagent la signification agentive : *celui qui...* Les exemples *chanteur, voleur* et *réparateur* s'analysent facilement comme des mots construits avec une base verbale :

- ◆ ***chanteur*** = ***chant***+*eur*
-(at)eur = celui qui...
 signification du mot dérivé = *celui qui chante*

De même pour les exemples *conducteur* et *électeur*, pour lesquels il faut néanmoins préciser que la base se présente sous la forme d'un allomorphe :

- ◆ ***conducteur*** = ***conduct***+*eur*
-(at)eur = celui qui...
conduct = allomorphe de la base verbale *cond-* (*conduire*)
 signification du mot dérivé = *celui qui conduit*

L'analyse est moins évidente pour les mots *lugeur* et *routeur* car il convient de restituer une base verbale peu usitée : *luger* = faire de luge et *router* = tracer la route. Ces deux verbes sont effectivement attestés en français (consulter le TLF (Trésor de la Langue Française) : <http://atilf.atilf.fr/tlf.htm>). *Routeur* a pour signification : personne qui conseille au skipper la route à suivre dans une course à la voile. *Routeur* a aussi une signification instrumentale : ce qui permet de tracer la route ; propriété générale du suffixe *-(at)eur* qui s'applique aussi bien à des personnes (suffixe agentif) qu'à des objets (suffixe instrumental) : *photocopieur*. *Lugeur* et *routeur* sont donc parfaitement réguliers : ce sont des mots dérivés construits.

Il en va différemment des mots *aviateur* et *facteur* pour lesquels il n'est pas possible de reconstruire la signification du mot dérivé à partir de la base lexicale et de son suffixe. Si l'on décrit la signification du mot *aviateur* comme étant : celui qui pilote un avion, et celle de *facteur* comme : celui qui distribue le courrier (une des significations du mot *facteur*), il n'est pas possible de restituer une base verbale en rapport avec la signification globale du mot. On dira donc que ces mots sont non construits.

Tous les mots de l'ensemble précédent sont donc des mots dérivés – reconnaissance d'un suffixe commun *-(at)eur* qui a une signification agentive ou instrumentale – mais certains sont construits et d'autres non.

Autre exemple du même type : soit l'ensemble des mots {*peureux, épineux, aventureux, avantageux, capricieux, astucieux, joyeux, ennuyeux, crayeux, anguleux, comateux, heureux, aqueux, onctueux, visqueux, sinueux*}. Comme pour l'ensemble précédent, cette liste non-exhaustive présente une gradation, allant des mots dont l'analyse est évidente aux mots dont l'analyse pose problème. A priori, tous sont analysables comme des adjectifs dérivés au moyen du suffixe *-eux/euse* qui exprime une qualité.

Peureux, épineux, aventureux, avantageux s'analysent aisément comme des adjectifs dérivés de noms (à noter qu'il n'est pas facile de décrire la signification du suffixe avec une phrase qui convienne pour les exemples) :

- ◆ *peureux* = **peur**+*eur*
 -*eux* = *qualité/état en rapport avec...*
 signification du mot dérivé = *qualité/état en rapport avec la peur*
 = *qui a peur*

Capricieux, astucieux s'analysent de même à la différence près qu'il faut prendre en compte la présence d'un allomorphe *caprici-* et *astuci-* pour *caprice* et *astuce*.

La même chose pour *joyeux, ennuyeux, crayeux* qui présentent également une base nominale allomorphique : *joy-*, *ennuy-* et *cray-* pour *joie, ennui* et *craie*.

De même pour *anguleux* et *comateux* : *angul-* pour *angle* et *comat-* pour *coma*.

Les problèmes d'analyse apparaissent avec les mots *heureux* et *aqueux*. Il est cependant possible d'établir une relation sémantique entre l'adjectif et un nom, en faisant appel à l'étymologie. *Heureux* est dérivé du nom *heur* qui signifie chance (*heureux* = qui a de la chance, qui procure de la chance). *Aqueux* est à mettre en rapport avec le mot latin *aqua* (*eau*) ; *aqueux* = qui a la nature de l'eau). On peut donc les considérer comme des mots construits à condition de faire appel à une base nominale étymologique.

Par contre les mots *onctueux, visqueux, sinueux*, qui sont également des adjectifs qui expriment une qualité, ne peuvent pas être mis facilement en relation avec un nom en rapport avec la qualité. *onctueux* est en rapport étymologique avec *unctum* (*ce qui est oint* (verbe *oindre*), *visqueux* avec le mot latin *viscum* (*glu*), et *sinueux* avec le mot latin *sinus* (*sein*). Il est donc très difficile de considérer ces mots comme des mots construits. Ce sont donc bien des mots dérivés – reconnaissance d'un suffixe commun *-eux/euse* qui exprime une qualité – mais certains sont construits et d'autres non.

Faire les exercices 22 et 23

Ces exercices portent sur la classification de mots en différents types (voir le document 4). Si un mot est perçu comme étant un mot complexe, il doit être possible de rétablir la signification du mot à partir de la signification de ses morphèmes. Ce qui présuppose évidemment de connaître la signification du mot. Dans le cas contraire, le recours au dictionnaire est nécessaire.

7.2.2. MOTS NON DÉRIVÉS ET NON CONSTRUITS

Considérons maintenant l'ensemble des mots suivants : {*neutron, proton, électron, boson, fermion, hadron, gluon, lepton, photon, positon, muon, graviton...*}. Tous ces mots se terminent en *-on* et ont la propriété commune de désigner une particule élémentaire du domaine de la physique des particules. Il est donc tentant de considérer la finale *-on* comme un suffixe ayant pour signification : particule élémentaire... Mais un suffixe est un morphème lié qui a pour caractéristique essentielle de sélectionner la catégorie de sa base lexicale. Dans les exemples précédents, *-eur* est un suffixe nominal agentif qui se construit avec un verbe,

-*eux*-**euse** est un suffixe adjectival qui se construit avec un nom. L'analyse de ces mots montre qu'il est impossible de définir une catégorie unique car la règle sémantique qui relie le tout à une de ses parties (la base) varie d'un mot à l'autre :

- ◆ **neutron** = particule élémentaire **neutre**
- ◆ **hadron** = particule élémentaire régie par l'interaction forte (grec : **hadrós** = fort)
- ◆ **électron** = particule élémentaire chargée **électriquement**
- ◆ **boson** = particule élémentaire découverte par Satyendranath **Bose**
- ◆ **fermion** = particule élémentaire découverte par Enrico **Fermi**
- ◆ **gluon** = particule élémentaire de transmission de l'interaction forte (anglais : **glue** = colle)
- ◆ **positon** (ou *positron*) = particule élémentaire de charge **positive**
- ◆ **photon** = particule élémentaire de lumière (grec : **phôtós** = lumière)
- ◆ **muon** = particule élémentaire notée μ (lettre grecque : μ = **mu**)

Ce sont des mots complexes dans la mesure où ils présentent une analyse en deux morphèmes mais le morphème lié *-on* n'est pas un véritable suffixe dérivationnel car sa signification (particule élémentaire) n'inclut pas une sélection de la catégorie de la base lexicale. Par ailleurs ce ne sont pas des mots construits car la formulation de la relation sémantique entre le tout et la base lexicale (relation entre *neutre* et *neutron*, entre *hadrós* et *hadron*, entre *Bose* et *boson*...) n'est pas constante.

La même analyse s'applique aux unités de l'analyse linguistique {*phonème, prosodème, chérème, morphème, lexème, grammatème, sémantème, sème, synthème, monème*...} et également à l'ensemble réduit des agrumes de petite taille {*mandarine, clémentine, tangerine*}.

Faire les exercices 24 et 25

Ces exercices portent sur la distinction entre mots dérivés construits et mots dérivés non construits.

7.2.3. SYNTHÈSE

À partir des oppositions suivantes :

- ◆ mots complexes et non complexes
- ◆ mots dérivés et non dérivés
- ◆ mots construits et non construits

il est possible d'envisager tous les cas de figure :

- ◆ mots complexes, dérivés et construits (exemples : *chanteur, acceptable, lentement*...)
- ◆ mots complexes, dérivés et non construits (exemples : *salade, aqueux, parlement*...)
- ◆ mots complexes, non dérivés et construits (exemples : *clairvoyant, géologie, informatique*...)
- ◆ mots complexes, non dérivés et non construits (exemples : *neutron, morphème, clémentine*...)
- ◆ mots non complexes, non dérivés et non construits (mots simples) (exemples : *maison, trop, devant*...).

Cette typologie et ces exemples sont repris dans un schéma dans le document 5.

8. FLEXION ET DÉRIVATION

La flexion est la partie de la morphologie qui traite de la variation des mots en syntaxe en phonologie.

Un mot est variable (différentes formes selon sa catégorie syntaxique) ou invariable en syntaxe :

- ◆ nom : variation en nombre (singulier et pluriel)
- ◆ adjectif : variation en nombre et genre (singulier et pluriel, masculin et féminin)
- ◆ pronom : variation en personne, nombre et genre (seulement pour la 3^e personne)
- ◆ verbe : variation en mode, temps et personne
- ◆ adverbe : invariable

Exemples :

- ◆ *maison* et *maisons* sont deux formes du même mot (lemme = *maison*)
- ◆ *beau*, *beaux*, *belle*, *belles* et *bel* sont cinq formes différentes du même mot (lemme = *beau*)
- ◆ *je*, *j'*, *me*, *m'*, *moi* sont cinq formes différentes du même mot (lemme = *je*)
- ◆ *chanter*, *chanté*, *chante*, *chantes*, *chantaient*, *chanterions...* sont différentes formes du même mot

(lemme = *chanter*). Un verbe a en moyenne 56 formes différentes.

- ◆ *trop* est l'unique forme de l'adverbe dont le lemme est *trop*.

Un mot peut avoir également une forme différente selon son contexte phonétique ; l'article défini *le* ou *la* ou pour variante *l'* devant un mot qui commence par une voyelle.

En dehors de la syntaxe et de la phonologie, un mot peut avoir une forme différente en composition (*franco-* est la forme du mot *français* dans le composé *franco-provençal*), selon la période (*leu* est une forme ancienne du mot *loup*) ou selon le registre de langue ; *puis* est une forme soutenue de *peux* (lemme = *pouvoir*), *fastoche* est une forme populaire du mot *facile*.

Un morphème peut avoir également des formes différentes :

- ◆ *-eur* et *euse* (*chanteur* et *chanteuse*) sont deux formes différentes du même morphème (variation en genre),
- ◆ *-al* et *aux* (*original* et *originaux*) sont deux formes différentes du même morphème (variation en nombre),
- ◆ *-s* et *-x* (*maisons* et *poux*) sont deux formes différentes du morphème de pluriel.

La dérivation est la partie de la morphologie qui traite de la formation des mots dans le lexique.

Chanteur, *chanson*, *chant*, *chantage*, *enchantement...* sont des mots différents, formés par dérivation à partir du même mot *chanter*, dont le radical est *chant-*.