

Morphologie

lundi 12 janvier 16h55 – 18h55

Denis Creissels

Université Lumière (Lyon 2)

denis.creissels@univ-lyon2.fr

<http://deniscreissels.fr>

Plan

1. Introduction
2. Morphologie et syntaxe
3. Lexème et mot-forme, paradigme, sens lexical, mots pleins et mots grammaticaux
 - 3.1. Lexèmes et mots-formes
 - 3.2. Mots pleins et mots grammaticaux
 - 3.3. Le sens lexical des mots
4. Morphologie constructionnelle (composition et dérivation) et morphologie flexionnelle
5. La morphologie concaténative : morphèmes et allomorphes
 - 5.1. Morphèmes, bases et affixes
 - 5.2. Suffixes, préfixes, infixes, circumfixes et transfixes
 - 5.3. Allomorphes d'un morphème
 - 5.4. Allomorphes phonologiques et allomorphes supplétifs
 - 5.5. Affixes réductifs
 - 5.6. Morphèmes amalgamés (ou portemanteaux)
 - 5.7. Morphèmes zéro, morphèmes vides, exponence multiple
 - 5.7.1. Morphèmes zéro
 - 5.7.2. Morphèmes vides (ou éléments thématiques)
 - 5.7.3. Exponence multiple
6. Les limites de la notion de morphème et la morphologie non concaténative
7. Spécificités de la morphologie tonale
 - 7.1. Remarques introductives
 - 7.2. Morphotonèmes additifs
 - 7.3. Morphotonèmes substitutifs
 - 7.4. Propagation tonale, glissement tonal et l'analyse du rôle morphologique des tons
8. Formes libres et formes liées, cliticisation et composition
 - 8.1. La distinction entre formes libres et formes liées
 - 8.2. Mots grammaticaux, clitiques et affixes

1. Introduction

La morphologie, au sens couramment donné à ce terme en linguistique, étudie la structure interne des mots. Un mot est, soit un segment significatif minimal qui n'est ni décomposable en unités significatives plus petites, ni attaché à ce qui le précède ou le suit immédiatement (mot simple), soit un segment significatif analysable comme une combinaison d'unités significatives plus petites mais qui manifeste à la fois un fort degré de cohésion interne et une relative autonomie par rapport à son environnement (mot complexe).

L'objectif de cet exposé est une présentation élémentaire de la morphologie qui ne se limite pas a priori à une langue particulière ou à un groupe de langues particulier, et qui privilégie la perspective de la linguistique de terrain : à partir de données collectées par le linguiste auprès de locuteurs de la langue, décrire la morphologie d'une langue peu documentée ou pas documentée du tout.

2. Morphologie et syntaxe

La division du travail entre morphologie et syntaxe est l'un des points sur lesquels il y a de manière évidente des contrastes considérables entre les langues. D'une part, il n'existe aucun type de signification qui s'exprimerait universellement au moyen de formes liées dont la combinaison avec d'autres unités relèverait nécessairement de la morphologie. Mais d'autre part toutes les relations exprimées par des combinaisons de mots (c'est-à-dire dans le cadre de la syntaxe) dans certaines langues peuvent dans d'autres langues se trouver exprimées par des combinaisons d'unités relevant de la morphologie.

Par exemple en soninké, la combinaison formée par un nom et un adjectif qui le modifie a toutes les caractéristiques d'un mot composé, et on ne ferait que compliquer inutilement la description en voulant à tout prix la décrire comme relevant de la syntaxe. Toujours en soninké – ex. (1), la relation entre un verbe comme **kárá** 'casser' et son deuxième argument (ici **súwà** 'bois') peut se concrétiser comme une relation syntaxique entre un verbe et un groupe nominal objet, mais elle peut aussi donner lieu à la formation d'un verbe composé. Outre le fait qu'en (b), le verbe est marqué formellement comme détransitivisé, la preuve la plus nette de la formation d'un verbe composé vient du fonctionnement de la négation, qui en soninké provoque une modification de la tonalité du verbe : dans la négation de (a), la modification tonale affecte seulement **káráná**, alors que dans la négation de (b), c'est toute la séquence **súwákárené** qui est affectée¹ :

- (1) a. **Ŋá** **súwàn** **kárá-ná.**
1SG COPLOC bois.D casser-GER
'Je casse le bois.'
- b. **Ŋá** **súwá-káré-né.**
1SG COPLOC bois-casser.DETR-GER
'Je casse du bois.'
- c. **Ŋá** **ntá** **súwàn** **kàrà-nà.**
1SG COPLOC.NEG bois.D casser-GER
'Je ne casse pas le bois.'
- d. **Ŋá** **ntá** **sùwà-kàré-nè.**
1SG COPLOC.NEG bois-casser.DETR-GER
'Je ne casse pas de bois.'

¹ Abréviations utilisées dans les gloses : COPLOC = copule de localisation, D = marque de détermination nominale, DETR = marque de détransitivisation, NEG = négation, SG = singulier.

Par ailleurs, selon les langues, la division entre combinaisons d'unités relevant de la morphologie et combinaisons d'unités relevant de la syntaxe peut s'avérer plus ou moins problématique. Dans les cas favorables, on peut s'appuyer sur plusieurs types de critères qui convergent pour caractériser la limite entre deux segments significatifs qui se succèdent, soit comme une limite entre deux mots, soit comme une limite entre deux formatifs d'un même mot. Les conventions orthographiques créent l'illusion qu'il s'agit là de la situation normale, mais en réalité, il y a relativement peu de langues dans lesquelles on peut sans difficulté majeure proposer une segmentation des phrases en mots qui évite toute décision arbitraire, car les critères sur lesquels on peut s'appuyer sont parfois inopérants, et pire encore, il peut arriver qu'ils donnent des indications contradictoires².

Dans l'enquête sur une langue peu documentée, il ne faut pas s'attendre à disposer dès le début du travail des critères permettant de délimiter les mots de la façon la mieux adaptée au système de la langue qu'on décrit. On peut espérer que la progression du travail permette d'affiner peu à peu les critères, mais même à un stade avancé de travail il n'y a rien d'anormal à ce qu'il subsiste un certain nombre d'hésitations, car des processus de grammaticalisation impliquant notamment la perte progressive d'autonomie de formes initialement libres sont constamment à l'œuvre dans les langues, ce qui veut dire que dans la description synchronique on a toutes les chances de rencontrer des unités pour lesquelles le processus est engagé mais n'a pas encore abouti.

Dans une enquête de terrain, il est difficile de ne pas procéder à un découpage en mots dès les premiers relevés de phrases, mais il faut être conscient qu'en dehors des frontières d'unités auxquelles l'informateur marque régulièrement une pause de façon spontanée, ce découpage a forcément un caractère provisoire et doit nécessairement être, soit confirmé, soit rectifié, au fur et à mesure qu'apparaîtront les données pertinentes. En début d'enquête, il faut surtout s'efforcer d'être cohérent, c'est-à-dire de toujours segmenter de la même façon ce que l'on peut penser être les manifestations d'une même construction, tout en ayant conscience du fait que c'est seulement à un stade plus avancé du travail qu'on pourra réellement s'interroger sur la validité des choix initiaux.

Pour les langues qui ont déjà une tradition d'écriture plus ou moins établie, il faut être attentif au fait que les conventions orthographiques donnent souvent une vision faussée de ce qu'est réellement le découpage des phrases en mots. Il est particulièrement commun que les conventions orthographiques fassent apparaître comme mots à part des formes qui ont de façon plus ou moins nette des propriétés de formes liées, mais la situation inverse existe aussi. Et même pour des langues qui n'ont pas de tradition ancienne d'écriture, on voit chez les linguistes travaillant à la description de langues de tel ou tel domaine se constituer très rapidement des traditions de découpage des phrases qui, si on les examine d'un œil critique, s'avèrent souvent comporter une forte dose d'arbitraire sinon même d'incohérence.

Nous reviendrons sur les critères qui peuvent servir à délimiter la mot à la fin de cet exposé, après avoir introduit et discuté les notions fondamentales de l'analyse

² Pour une discussion récente de cette question, cf. Haspelmath (2011).

morphologique. Dans l’immédiat, nous allons simplement examiner un exemple tiré du mandinka.

En mandinka, il existe deux constructions de sens voisin mais tout de même différent dans lesquelles un lexème verbal apparaît combiné à la copule locative **bè**, présentées ici telles qu’elles sont analysées dans Creissels & Sambou (2013)³ :

(2) a. **Fàatú bè m àan-óo t ùw-ó là.**
Fatou COPLOC riz-D piler-D POSTP
‘Fatou est en train de piler le riz.’

b. **Fàatú bè m àan-óo t ùu-là.**
Fatou COPLOC riz-D piler-INF
‘Fatou va piler le riz.’

Dans la mesure où, en mandinka, (a) tout lexème verbal peut s’utiliser comme nom d’événement sans avoir à subir une quelconque opération morphologique et (b) un nom en fonction de génitif précède le nom qu’il détermine, il est clair que la construction de la première de ces deux phrases, où la marque de détermination nominale suffixée à **t ùu** ‘piler’ indique sans ambiguïté que ce lexème verbal est utilisé nominalement, correspond littéralement à ‘Fatou est au pilage du riz’. Quant à la deuxième, il est probable qu’un linguiste abordant l’analyse du mandinka ne verrait aucune raison évidente de la traiter différemment de la première, et notamment d’analyser **t ùu-là** comme une forme suffixée du verbe ‘piler’ plutôt que de voir là la combinaison d’un verbe utilisé comme nom à la forme nue et d’une postposition. Mais une meilleure connaissance du système de détermination nominale du mandinka montre que cette analyse contredit les principes généraux de fonctionnement du marqueur de détermination nominale glosé D, qui en règle générale ne peut être omis que dans des conditions précises et limitées. En outre, dans la phrase (a), des particules discursives peuvent s’insérer entre **t ùw-ó** et **là**, alors qu’aucune insertion n’est possible entre **t ùu** et **là** dans la phrase (b). La conclusion est que, même si (comme cela est plus que probable) la construction illustrée en (b) est issue par grammaticalisation d’une construction qui initialement devait être du type illustré en (a), **là** est dans la phrase (a) une postposition qui a une relative autonomie par rapport au nom auquel elle succède, alors qu’en (b), **là** est un suffixe verbal.

3. Lexème et mot-forme, paradigme, sens lexical, mots pleins et mots grammaticaux

3.1. Lexèmes et mots-formes

Le terme de mot recouvre deux notions distinctes, lorsque nous disons par exemple que le mot *dormez* comprend deux syllabes, ou que *dormez* et *dort* sont deux formes du même mot. Pour éviter cette ambiguïté, on peut utiliser *lexème* pour se référer spécifiquement à ce qui est commun à un ensemble de mots regroupés sous une

³ Abréviations utilisées dans les gloses : COPLOC = copule de localisation, D = marque de détermination nominale, INF = infinitif, POSTP = postposition.

même entrée lexicale⁴, et *mot-forme* pour se référer au mot en tant que séquence de phonèmes ou de lettres. On désigne fréquemment comme *paradigme* l'ensemble des mots-formes appartenant à un même lexème.

La question de savoir ce qui peut justifier de regrouper les mots-formes en paradigmes est dans l'immédiat laissée de côté ; elle sera reprise à la section 4. Pour l'instant, limitons-nous à observer que l'appartenance à un même paradigme telle qu'elle vient d'être définie est une notion plus restreinte que le simple fait pour deux mots d'avoir une parenté morphologique : *dormez* et *dormeur* ont manifestement un formatif en commun, mais ne sont pas regroupés à une même entrée lexicale : à la différence de *dormez*, *dormeur* n'est pas une forme fléchie du verbe *dormir*, mais un nom dérivé du verbe *dormir*.

3.2. Mots pleins et mots grammaticaux

Comme beaucoup des notions de base de l'analyse linguistique, la distinction entre mots pleins et mots grammaticaux, bien qu'unanimement admise, s'avère difficile à préciser au moyen de tests à la fois opératoires et théoriquement justifiés permettant de distinguer les uns des autres mots pleins et mots grammaticaux⁵. Le test le plus généralement invoqué est que les mots pleins constituent des classes ouvertes et peu stables, alors que les mots grammaticaux constituent des classes fermées et relativement stables, mais l'application de ce critère est parfois problématique, et la question de sa justification théorique reste entière.

Une définition particulièrement intéressante, bien que rarement envisagée, consiste à poser que les classes de mots pleins sont celles où vont se ranger les lexèmes (au sens défini en 3.1) qui ont un sens dénotatif indépendant de toute référence à l'énonciation et au discours (par exemple, un objet est reconnaissable comme *table* ou *chaise* quelles que soient les conditions dans lesquelles on en parle), alors que le sens des mots grammaticaux est relatif à une situation d'énonciation et à la construction d'un discours (par exemple, il n'y a pas d'objet qui soit intrinsèquement *celui-ci* ou *le mien*). Cette définition ne résout pas toutes les difficultés, mais elle permet au moins d'éviter quelques faux problèmes et de résoudre quelques difficultés auxquelles se heurtent les autres approches possibles.

⁴ Ceci n'est malheureusement pas la seule acception avec laquelle on peut rencontrer *lexème*. Chez certains auteurs, *lexème* est synonyme de mot plein, et la notion de lexème exclut les mots grammaticaux, alors que selon la définition retenue ici elle les englobe. Chez d'autres, *lexème* s'applique à des unités significatives minimales (ou morphèmes lexicaux, par opposition aux morphèmes grammaticaux).

⁵ Ces termes sont repris ici faute de mieux pour une distinction dont il semble impossible de se passer, mais pour laquelle n'existe aucune solution terminologique pleinement satisfaisante. Il convient de s'attacher à leur définition plus qu'à la signification que suggère leur étymologie. Le terme de *mot plein* correspond dans la terminologie anglaise à *content word*. Il a bien sûr l'inconvénient de suggérer que les mots qui ne sont pas 'pleins' seraient d'une manière ou d'une autre 'vides', ce qu'on souhaiterait éviter ici. On peut être tenté de préférer *mot lexical*, mais ce terme implique une opposition entre lexical et grammatical difficile à articuler avec la notion courante de lexique comme ensemble des mots d'une langue, ainsi qu'avec la décision d'utiliser de manière générale le terme de lexème pour se référer à ce qui est commun aux mots-formes regroupés sous une même entrée lexicale. En ce qui concerne *mot grammatical*, le terme anglais correspondant est *function word*, qu'on traduit parfois en français par *mot fonctionnel* ; on trouve parfois aussi le terme de *mot-outil*.

Par exemple, quels que soient les critères invoqués pour distinguer les mots pleins des mots grammaticaux, les prépositions apparaissent difficiles à classer selon cette distinction. Or, la définition ci-dessus justifie de reconnaître par exemple que la préposition *sur* du français, même si on peut dans certaines constructions lui reconnaître des emplois grammaticalisés, a fondamentalement le statut de mot plein : dans *Le journal est sur la table*, *sur* signifie entre deux entités désignées comme *le journal* et *la table* une relation dont la reconnaissance est purement un problème de conceptualisation d'une situation de référence.

Beaucoup de difficultés auxquelles on se heurte dans la description grammaticale des langues tiennent toutefois au fait que sont constamment à l'œuvre des processus de grammaticalisation par lesquels les mots pleins présentant certaines caractéristiques sémantiques tendent à devenir des mots grammaticaux, et les mots grammaticaux tendent à perdre leur autonomie et à devenir des affixes.

3.3. *Le sens lexical des mots*

Le *sens lexical* d'un mot est ce qui est sémantiquement commun à l'ensemble des mots-formes qui constituent le paradigme auquel il appartient. Dans la mesure où *lexical* est un terme fortement polysémique, qui s'oppose dans certaines de ses acceptions à *grammatical*, il importe de souligner que selon cette définition, il n'y a aucune contradiction à parler de sens lexical d'un mot grammatical. Par exemple, le possessif *mien* est un mot grammatical ; il fait partie d'un paradigme dont les autres membres sont *mienne*, *miens* et *miennes*, et a donc comme sens lexical 'appartenance à la sphère personnelle de l'énonciateur' ; pour chacun des quatre mots-formes qui constituent le paradigme, à ce sens lexical s'ajoutent une valeur de nombre (*singulier / pluriel*) et de genre (*masculin / féminin*).

4. Morphologie constructionnelle (composition et dérivation) et morphologie flexionnelle

On peut désigner du terme général de *morphologie constructionnelle* l'ensemble des relations que peuvent entretenir des lexèmes différents mais apparentés (appartenant selon la terminologie traditionnelle à la même 'famille de mots'), un lexème construit étant un lexème dont la formation peut se décrire à partir d'un ou plusieurs autres lexèmes. Les grammaires traditionnelles parlent de 'formation des mots', où 'mot' doit être compris au sens de 'lexème'.

La morphologie *constructionnelle* est à distinguer de la morphologie *flexionnelle*, ensemble des relations que peuvent entretenir les différentes formes d'un même lexème (par exemple selon les langues, le singulier et le pluriel d'un nom, le présent et le futur d'un verbe, le masculin et le féminin d'un adjectif).

A l'intérieur de la morphologie constructionnelle, on distingue la *composition* de la *dérivation*. Dans la composition, la formation d'un lexème met en jeu deux éléments ou plus qui peuvent chacun de leur côté se rencontrer dans des mots pleins dont ils encodent le sens lexical. Par exemple en mandinka, **màani-késé** 'grain de riz' est un mot unique dont les deux formatifs s'identifient comme **màaní** 'riz' et **késé** 'grain'. Dans la dérivation par contre interviennent des éléments, les *dérivatifs*, qui n'ont pas

cette propriété. Par exemple en soninké, on peut rendre ‘siège’ de deux façons : **táaxú-hô** ou **táax-áadè**. Dans les deux cas, on reconnaît comme premier formatif **táaxú** ‘s’asseoir’, mais **táaxú-hô** relève de la composition, car on peut reconnaître comme deuxième formatif le nom **hô** ‘chaise’, tandis que **táax-áadè** relève de la dérivation, car **-aade** est un suffixe formateur de noms d’instruments qui n’a pas d’existence indépendante.

Morphologie dérivationnelle et morphologie flexionnelle ne se distinguent pas par le type d’opérations formelles qu’elles peuvent mettre en jeu, mais uniquement par le statut relatif des formes qu’elles mettent en relation : deux lexèmes différents dans le cas de la morphologie dérivationnelle, un lexème et ses formes fléchies dans le cas de la morphologie flexionnelle. Cette distinction, qui est cruciale pour l’organisation d’une grammaire descriptive, oblige donc à poser la question du regroupement des mots-formes en paradigmes, à la fois du point de vue théorique et du point de vue des critères utilisables pour établir ce regroupement.

Dans la pratique, il n’est pas difficile de reconnaître comme relevant de la dérivation les processus morphologiques qui créent des lexèmes appartenant à une autre classe de mots (partie du discours) que le lexème dont ils sont issus (*chanter* → *chanteur*). Par contre la question des critères permettant de caractériser comme flexionnels ou dérivationnels les processus morphologiques qui n’ont pas cet effet est une question délicate. De tous les critères qui ont pu être proposés, le plus intéressant est celui qui fait référence à l’interaction entre la structure du mot et son comportement syntaxique. Selon ce critère, seules les caractéristiques flexionnelles du mot peuvent intervenir dans la formulation de règles syntaxiques, et il ne doit pas exister de règle syntaxique spécifique à tel ou tel type de mot dérivé. Autrement dit, la dérivation peut modifier les propriétés syntaxiques du mot, mais le comportement syntaxique d’un mot dérivé doit pouvoir se décrire en le rattachant simplement à une catégorie à laquelle appartiennent aussi des mots non dérivés. Par contre, une règle de syntaxe peut comporter des conditions nécessaires sur les caractéristiques flexionnelles des mots auxquels elle s’applique.

Par exemple, selon ce critère, l’infinitif du français est une formation flexionnelle, car les propriétés syntaxiques des infinitifs ne se retrouvent dans aucun mot qui ne présente pas la terminaison caractéristique des infinitifs ; par contre, le nom d’action est une formation dérivationnelle, car les noms d’action ont les mêmes propriétés syntaxiques que les noms qui ne sont pas dérivés de verbe. Le temps verbal est une formation flexionnelle, car aucun mot dépourvu de variation en temps ne peut se substituer au verbe conjugué. Le pluriel des noms est aussi en français une formation flexionnelle, du fait de l’existence de règles d’accord en nombre. Par contre, dans une langue avec une formation morphologique de pluriel des noms ne donnant lieu à aucun phénomène d’accord, on devrait envisager la possibilité de reconnaître le pluriel comme dérivationnel.

5. La morphologie concaténative : morphèmes et allomorphes

5.1. Morphèmes, bases et affixes

La structure des mots se manifeste par des ressemblances partielles de forme régulièrement associées à des ressemblances partielles de signification. Ce sont ces relations qui permettent de considérer que deux mots-formes appartiennent au même lexème, ou que deux lexèmes appartiennent à la même famille. Dans les cas simples, on peut comme en (3) rendre compte de la relation entre mots apparentés à la fois par la forme et le sens en les présentant comme la concaténation de segments qui véhiculent chacun une partie de la signification.

(3)	aim-er	aim-ons	—	aim-able	—
	trouv-er	trouv-ons	—	trouv-able	in-trouv-able
	lav-er	lav-ons	lav-eur	lav-able	—
	cass-er	cass-ons	cass-eur	cass-able	in-cass-able

A partir de là, il est tentant de poser comme principe général que, de même que la phrase est une chaîne de mots, le mot est une chaîne de segments significatifs minimaux, ou *morphèmes*, le morphème étant à son tour une chaîne de phonèmes (ou de lettres, si on prend en considération la forme écrite de la langue).

Nous examinerons à la section 6 les difficultés auxquelles peut se heurter ce type d'analyse et les solutions possibles. Dans l'immédiat nous examinons les notions de base de la morphologie concaténative, qui décrit la dérivation et la flexion en termes d'adjonction d'*affixes* à une *base*.

Dans les cas simples, la base peut coïncider avec l'une des formes dont on décrit la relation, ou ne pas exister elle-même comme mot mais pouvoir se définir comme ce qui est commun à un ensemble de formes morphologiquement apparentées. La relation entre singulier et pluriel des noms en tswana illustre ces deux cas de figure :

(4) Singulier et pluriel des noms en tswana

a.	q^hòmó	'vache'	dì-q^hòmó	'vaches'
	pàtà	'poche'	dì-pàtà	'poches'
	t^hépé	'robinet'	dì-t^hépé	'robinets'
b.	lì-χòdù	'voleur'	mà-χòdù	'voleurs'
	lì-ηólé	'genou'	mà-ηólé	'genoux'
	lì-bèlé	'mil'	mà-bèlé	'mils'

5.2. Suffixes, préfixes, infixes, circumfixes et transfixes

Les *suffixes* (qui succèdent à la base) constituent de loin le type de plus courant d'affixes dans les langues du monde. Les *préfixes*, bien que moins répandus, sont eux aussi communs. Les *infixes* (qui se placent à l'intérieur de la base) et les *circumfixes* (qui se placent de part et d'autre de la base) sont par contre infiniment moins

répandus, et il faut être attentif aux emplois abusifs des termes d'infixe et de circumfixe. On appelle parfois abusivement 'infixes' des préfixes qui se placent entre un autre préfixe et la base, et on analyse souvent abusivement comme 'circumfixe' la combinaison d'un préfixe et d'un suffixe qui vont souvent de pair mais dont une analyse fine montre qu'ils ne sont pas vraiment dans une relation d'implication mutuelle.

Avec la *transfixation*, on est vraiment à la limite de ce qu'on peut considérer comme relevant encore de la morphologie concaténative. Cette notion peut être utilisée pour rendre compte de formations qu'on trouve notamment dans les langues sémitiques, où on peut analyser certains mots comme résultant de la combinaison de deux morphèmes tous deux discontinus dont les fragments s'intercalent. Par exemple en arabe, **kutiba** 'être écrit, accompli, 3ème pers. mas. sg.' peut s'analyser comme la combinaison d'une base **k-t-b**, d'un morphème de passif **u-i** et d'un suffixe **-a**.

5.3. Allomorphes d'un morphème

Souvent, l'analyse des mots comme concaténations de morphèmes ne peut fonctionner qu'à condition d'admettre que chacun des morphèmes dont la concaténation constitue les mots peut avoir plusieurs variantes (ou *allomorphes*) dont le choix peut donner lieu à divers types de conditionnement par le contexte. L'allomorphie peut concerner aussi bien les bases – ex. (5) – que les affixes – ex. (6).

(5) Les verbes de l'espagnol à la 1ère personne du singulier et du pluriel du présent

tom-o	'je prends'	tom-amos	'nous prenons'
cant-o	'je chante'	cant-amos	'nous chantons'
llev-o	'j'emporte'	llev-amos	'nous emportons'
cierr-o	'je ferme'	cerr-amos	'nous fermons'
cuent-o	'je compte'	cont-amos	'nous comptons'
jueg-o	'je joue'	jug-amos	'nous jouons'

(6) Le suffixe 'possesseur de 1ère personne du singulier' en turc

ev	'maison'	ev-im	'ma maison'	
dil	'langue'	dil-im	'ma langue'	
köy	'village'	köy-üm	'mon village'	ö = [ø], ü = [y]
gül	'rose'	gül-üm	'ma rose'	
yol	'route'	yol-um	'ma route'	
tuz	'sel'	tuz-um	'mon sel'	
ad	'nom'	ad-ım	'mon nom'	ı = [w]
kız	'fille'	kız-ım	'ma fille'	
para	'argent'	para-m	'mon argent'	

5.4. Allomorphes phonologiques et allomorphes supplétifs

Les exemples en 5.3 illustrent des allomorphes *phonologiques*, qui présentent une ressemblance de forme telle qu'on peut envisager de les décrire au moyen de règles phonologiques ou morphophonologiques spéciales dont l'input est, ou bien l'un des allomorphes à partir duquel il apparaît relativement facile de prédire les autres, ou bien une représentation phonologique abstraite.

Par exemple, il est traditionnel de décrire l'alternance illustrée en (5) comme une modification des voyelles (diphthongaison) qui se produit automatiquement dans une partie des verbes espagnols lorsque la voyelle susceptible d'être modifiée porte l'accent de mot. Par contre en turc, on a plutôt l'habitude de présenter les variations illustrées en (6) à partir d'une forme abstraite **-(I)m** où **I** note une voyelle sous-spécifiée, définie seulement comme 'fermée', et qui emprunte ses autres traits (antérieure / postérieure et arrondie / non-arrondie) à la voyelle précédente. Le choix entre ces deux types de traitement doit s'appuyer sur l'observation de la nature phonétique de la différence entre les allomorphes ainsi que de leur distribution, mais le choix est loin d'être toujours évident.

Mais à côté des *allomorphes phonologiques*, il peut exister des *allomorphes supplétifs*, sans ressemblance de forme suffisante pour qu'on puisse envisager raisonnablement de les dériver phonologiquement l'un de l'autre ou d'une même représentation phonologique abstraite. C'est par exemple le cas du morphème d'imparfait dans la morphologie verbale de l'espagnol.

(7) L'imparfait des verbes espagnols

cant-aba	'je chantais'	com-ía	'je mangeais'
cant-aba-s	'tu chantais'	com-ía-s	'tu mangeais'
cant-aba	'il/elle chantait'	com-ía	'je mangeais'
cant-ába-mos	'nous chantions'	com-ía-mos	'nous mangions'
cant-aba-is	'vous chantiez'	com-ía-is	'vous mangiez'
cant-aba-n	'ils/elles chantaient'	com-ía-n	'ils/elles mangeaient'

Pour des raisons qui tiennent de manière évidente aux changements historiques qui peuvent brouiller progressivement les éléments formels initialement communs à deux mots apparentés, il n'est pas toujours aisé de tracer la limite entre ce qu'on peut considérer synchroniquement comme allomorphie phonologique et allomorphie supplétive.

5.5. Affixes réduplicatifs

L'ex. (6) ci-dessus illustre la possibilité qu'un affixe comporte des éléments variables dont les variations se ramènent à la copie d'un élément de la base. Le cas-limite d'une telle situation est constitué par les *affixes réduplicatifs*, entièrement constitués de copies d'éléments de la base. La formation de verbes causatifs du balant ganja illustrée en (8) met en jeu un préfixe CVC- dans lequel C représente une copie de la consonne initiale de la base, et V une copie de la première voyelle de la base.

(8) Verbes causatifs formés par affixation réduplicative en balant ganja

baay	‘jouer’	bab-baay	‘faire jouer’
duul	‘devenir petit’	dud-duul	‘rendre petit’
juul	‘grossir’	juj-juul	‘faire grossir’
мөөн	‘devenir noir’	мом-мөөн	‘rendre noir’

5.6. Morphèmes amalgamés (ou portemanteaux)

La comparaison d'un mot avec d'autres qui lui sont apparentés peut conduire à analyser certains segments formellement indécomposables comme chargés d'informations multiples mais amalgamées. La reconnaissance d'un amalgame peut se justifier par le fait que les significations concernées ont par ailleurs une expression morphologique autonome (cf. en français **au** [o], amalgame de la préposition **à** et de l'article défini masculin singulier **le**, comme le montre la comparaison avec la séquence **à la**), mais peut aussi découler de l'analyse des relations sémantiques entre membres d'un paradigme donné. En effet, l'analyse des distinctions exprimées par un paradigme peut permettre de séparer des valeurs qui ne sont jamais exprimées de manière indépendante, comme le nombre et le cas dans le paradigme nominal du serbo-croate illustré en (9) avec le nom **knjiga** ‘livre’.

(9) Impossibilité de séparer l'expression du cas de l'expression du nombre dans les paradigmes nominaux du serbo-croate ⁶

	singulier	pluriel	
nominatif	knjig-a	knjig-e	‘livre(s)’
accusatif	knjig-u	knjig-e	
génitif	knjig-ē	knjig-ā	
datif	knjiz-i	knjig-ama	
instrumental	knjig-om	knjig-ama	

Ici encore, on peut se demander dans quelle mesure l'hypothèse de formes sous-jacentes abstraites permettrait d'expliquer ce type de situation comme le résultat de processus phonologiques. Dans l'ex. (9), on n'imagine pas bien comment cela serait possible, mais les verbes causatifs de l'akhvakh – ex. (10) permettent d'illustrer cette possibilité. Dans une partie des formes fléchies des verbes causatifs de l'akhvakh, il n'y a pas de segmentation évidente entre un suffixe dérivatif et un suffixe flexionnel, mais en s'appuyant sur les formes où une telle segmentation est possible, on peut poser que les formes causatives incluent toutes un suffixe causatif **-a(j)-** dont le **j** est maintenu dans une partie de la flexion seulement, et qu'en cas de chute du **j**, le **a** s'amalgame à la voyelle initiale du suffixe flexionnel.

⁶ Le macron indique ici la longueur vocalique, qui n'est pas notée dans l'orthographe courante.

(10) Amalgames dans la flexion des verbes causatifs de l'akhvakh ⁷

	'courir'	'faire courir'	
impératif	$\bar{k}'^w\text{et-a}$	$\bar{k}'^w\text{et-aj-a}$	
conditionnel	$\bar{k}'^w\text{et-ala}$	$\bar{k}'^w\text{et-aj-ala}$	
explicatif	$\bar{k}'^w\text{et-erogu}$	$\bar{k}'^w\text{et-aj-erogu}$	
infinitif	$\bar{k}'^w\text{et-urula}$	$\bar{k}'^w\text{et-ōrula}$	< $\bar{k}'^w\text{et-a-urula}$
accompli	$\bar{k}'^w\text{et-ari}$	$\bar{k}'^w\text{et-āri}$	< $\bar{k}'^w\text{et-a-ari}$
accompli négatif	$\bar{k}'^w\text{et-ila}$	$\bar{k}'^w\text{et-ēla}$	< $\bar{k}'^w\text{et-a-ila}$

Mais dans ce genre de situation, la limite est souvent floue entre ce qu'on peut admettre comme relevant vraiment d'une analyse morphologique synchronique et ce qui doit plutôt être considéré comme une hypothèse de reconstruction.

5.7. Morphèmes zéro, morphèmes vides, exponence multiple

Comme cela vient d'être évoqué en 5.6, en procédant à une analyse sémantique et distributionnelle des mots-formes considérés comme constituant un paradigme flexionnel sans se préoccuper de la nature précise de leurs ressemblances et différences au niveau formel, on peut dégager les catégories flexionnelles pertinentes pour l'interface morphologie-syntaxe (par exemple, selon les langues, nombre et cas pour les noms, TAM et négation pour les verbes, etc.). Dans les cas idéalement simples, le mot fléchi comporte autant de positions affixales que de catégories flexionnelles à exprimer, et à chaque catégorie correspond de manière biunivoque l'une des positions. Mais ce n'est pas toujours le cas, et la relation entre catégories flexionnelles et affixes peut être plus complexe. Nous en avons vu un premier exemple avec les morphèmes amalgamés (ou portemanteaux). Les morphèmes zéro, les morphèmes vides et l'exponence multiple sont d'autres exemples de complexifications possibles de la relation entre les catégories flexionnelles et les affixes qui les expriment.

5.7.1. Morphèmes zéro

Il est courant que l'une des valeurs que peut prendre une catégorie flexionnelle donnée se distingue des autres par le fait qu'elle s'exprime par l'absence d'affixe qui l'exprimerait explicitement. On peut alors parler de morphème zéro. Par exemple, l'analyse des formes finies du verbe français permet de dégager trois positions affixales dont chacune peut être laissée inoccupée. Dans cette analyse, $j \sim \epsilon$ est le morphème commun aux formes traditionnellement désignées comme imparfait et conditionnel, tandis que \mathfrak{v} est le morphème commun aux formes désignées traditionnellement comme futur et conditionnel :

⁷ Dans la transcription de l'akhvakh, le macron note les voyelles longues et les consonnes fortes.

(11) Structure morphologique des formes finies du verbe français

ʃātə	ɾ	j	ɔ̃	‘chanterions’
ʃāt	-	j	ɔ̃	‘chantions’
ʃātə	ɾ	-	ɔ̃	‘chanterons’
ʃātə	ɾ	ɛ	-	‘chanterais’
ʃāt	-	-	ɔ̃	‘chantons’
ʃāt	-	ɛ	-	‘chantais’
ʃātə	-	-	-	‘chante’

5.7.2. Morphèmes vides (ou éléments thématiques)

On peut juger souhaitable qu’une description de la flexion évite de multiplier les allomorphes et de répéter inutilement des règles d’allomorphie valables avec le même conditionnement pour beaucoup de morphèmes. Il est depuis longtemps établi en linguistique indo-européenne que souvent, cet idéal ne peut être atteint qu’en admettant que des ‘éléments thématiques’ dépourvus de signification propre puissent s’intercaler entre la base du mot fléchi et les affixes flexionnels. On retrouve par exemple cette situation dans les langues caucasiennes de l’est, où on ne peut décrire simplement la flexion en cas des noms qu’en admettant à tous les cas sauf au cas zéro⁸ la présence possible d’un ‘formatif de base oblique’ entre la base nominale et le suffixe de cas proprement dit. Par exemple en akhvakh :

(12) ‘Formatifs de base oblique’ dans la flexion nominale de l’akhvakh

	‘hache’	‘faux’	‘miroir’
cas zéro	āžite	qaliča	mat’u
ergatif	āžite-de	qaliča-ro-de	mat’u-na-de
génitif	āžite-ĭi	qaliča-ro-ĭi	mat’u-na-ĭi
datif	āžite-la	qaliča-ro-la	mat’u-na-la
locatif ₁	āžite-ge	qaliča-ro-ge	mat’u-na-ge

5.7.3. Exponence multiple

L’expression d’une valeur flexionnelle donnée n’est pas toujours strictement limitée à l’un des formatifs du mot fléchi. Lorsque le phénomène est strictement ‘local’ (c’est-à-dire se limite à deux positions contiguës), il est courant de considérer qu’il s’agit tout simplement d’un phénomène d’allomorphie, l’un des deux affixes conditionnant la forme précise prise par l’autre. Mais cette solution n’est pas toujours satisfaisante, notamment lorsque le phénomène concerne plusieurs positions éloignées les unes des autres dans des langues où le mot fléchi peut comporter de longues chaînes d’affixes. Dans de tels cas, on peut juger préférable de relâcher quelque peu la notion de morphème, et de poser à la fois qu’un morphème donné peut contribuer à l’expression de plusieurs catégories, et qu’une catégorie peut se manifester à travers

⁸ Il s’agit du cas couramment désigné comme nominatif ou absolutif selon les descripteurs.

plusieurs morphèmes différents. Par exemple, la structure des formes verbales du tswana peut être décrite selon un ‘gabarit’ (en anglais : template) comportant quatre positions préfixales et cinq positions suffixales possibles. Or comme le montre l’exemple suivant, les distinctions *présent vs. parfait* et *positif vs. négatif* mettent en jeu cinq des neuf positions affixales possibles. Autrement dit, les morphèmes occupant ces cinq positions contribuent tous à divers titres à l’expression du TAM et de la polarité (et aussi à l’indexation du sujet, pour la position –3), et c’est seulement au prix d’une certaine dose de simplification (dont il importe d’être conscient, et qu’on est d’ailleurs en droit de juger excessive) qu’on peut étiqueter certains de ces morphèmes comme marques de parfait ou de négation.

(13) *présent vs. parfait* et *positif vs. négatif* en tswana (formes ‘conjointes’⁹)

-4	-3	-2	-1	0	+1	+2	+3	+4	+5	
	ú-			rék				-á/à		‘il/elle achète’
χà-	á-			rék				-í		‘il/elle n’achète pas’
	ú-			rék		-íl		-é/è		‘il/elle a acheté’
χà-	á-	à-		rék				-á		‘il/elle n’a pas acheté’

6. Les limites de la notion de morphème et la morphologie non concaténative

Il n’y a rien d’exceptionnel à ce que les relations formelles régulières entre mots ayant une partie de leur signification en commun prennent des formes qui se prêtent mal ou pas du tout à une analyse en morphèmes du type envisagé à la section 5, et on doit reconnaître la possibilité que la morphologie dérivationnelle ou flexionnelle mette en jeu d’autres opérations que l’adjonction d’affixes à une base, la question étant de déterminer pour chaque cas concret le type d’opération qui décrit de façon optimale les données.

On doit donc envisager de décrire la structure des mots complexes, non pas forcément en termes d’adjonctions d’affixes à une base (ce qui constitue d’ailleurs une pratique relativement récente dans la linguistique européenne), mais plus généralement en termes d’opérations affectant une base, sans restriction a priori sur la nature des opérations susceptibles de produire des mots à partir d’une base. La base est bien sûr choisie de façon à permettre de décrire le plus simplement possible la relation entre mots considérés comme morphologiquement reliés entre eux. Comme dans la morphologie concaténative, la base à laquelle s’appliquent les opérations morphologiques peut être l’une des formes dont on décrit la relation, mais cela n’a rien de nécessaire.

Par exemple, dans une analyse morphologique de la flexion adjectivale en français considéré sous sa forme orale, il est intéressant de comparer l’analyse fondée sur la prononciation des parlers méridionaux et l’analyse fondée sur la prononciation standard. En effet, dans la forme orale des variétés méridionales du français, il n’y a aucun problème à isoler un morphème -ə qui marque le féminin, à condition d’admettre que pour certains adjectifs, ce suffixe sélectionne un allomorphe de la

⁹ En linguistique bantoue, on désigne comme ‘conjointes’ des formes verbales qui doivent nécessairement être suivies d’un complément et ne peuvent pas figurer en fin de phrase.

base adjectivale différent de celui qui apparaît en l'absence de toute affixation.

(14) Masculin et féminin des adjectifs en français méridional

nyl	‘nul’	+ fém. →	nyl-ə	‘nulle’
nwaɁ	‘noir’		nwaɁ-ə	‘noire’
gɔ	‘gros’		gɔs-ə	‘grosse’
lɔŋ	‘long’		lɔŋg-ə	‘longue’

Par contre en français standard, sauf à postuler des formes phonologiques sous-jacentes très abstraites, on ne peut pas décrire la relation entre le masculin et le féminin des adjectifs variables en genre en isolant un suffixe qui marquerait le genre ; par contre, on peut décrire de façon simple la relation entre les deux formes des adjectifs variables en genre par une opération de troncation : si on prend comme base la forme de féminin, on peut poser que le masculin se forme en règle générale par suppression de la consonne finale, la suppression d'une nasale s'accompagnant d'une nasalisation de la voyelle finale (et parfois aussi d'un changement de timbre de cette voyelle).

(15) Masculin et féminin des adjectifs en français standard

vɛʁt	‘verte’	+ masc. →	vɛʁ	‘vert’
bavaʁd	‘bavard’		bavaʁ	‘bavarde’
lɔŋ	‘longue’		lɔ̃	‘long’
gʁos	‘grosse’		gʁo	‘gros’
øʁøz	‘heureuse’		øʁø	‘heureux’
zɑ̃tiʝ	‘gentille’		zɑ̃ti	‘gentil’
plɛn	‘pleine’		plɛ̃	‘plein’
divin	‘divine’		divɛ̃	‘divin’

L'adjonction d'affixes à une base est donc seulement un cas particulier parmi les opérations morphologiques par lesquelles se concrétise la relation entre mots apparentés à la fois par le forme et le sens. On peut citer comme opérations morphologiques possibles :

- la suppression d'un élément de la base (troncation) – cf. ci-dessus,
- différents types d'alternance qui modifient un ou plusieurs des segments phoniques qui constituent la base : palatalisation d'une consonne, voisement / dévoisement d'une consonne, antériorisation d'une voyelle, nasalisation d'une voyelle, allongement d'une voyelle brève, etc.) – ex. (16),
- les alternances prosodiques : glissement de l'accent de mot – ex. (17) – ou changement tonal – cf. section 7,
- la redoublement de la base – à distinguer de l'adjonction d'affixes réduplicatifs, en dépit de la parenté historique probable entre les deux procédés.

- (16) anglais : noms et verbes sémantiquement apparentés se distinguant par le caractère voisé ou non voisé de leur consonne finale

house [s]	house [z]
use [s]	use [z]
wreath [θ]	wreathe [ð]
thief [f]	thieve [v]

- (17) anglais : noms et verbes sémantiquement apparentés se distinguant par la place de l'accent de mot

'abstract	ab¹stract
'compound	com¹ound
'increase	in¹crease
'record	re¹cord

Du point de vue terminologique, il faut être attentif au fait que pour certains linguistes, le terme de morphème doit être réservé à des *segments* significatifs minimaux qui se concatènent pour donner les mots, alors que pour d'autres, 'morphème' peut s'appliquer à toute manifestation formelle d'une différence de sens. Dans un souci de clarté, on évite ici d'utiliser 'morphème' en dehors de sa signification étroite, mais il n'y a pas d'objection de principe à utiliser ce terme à son sens large, à condition d'en reformuler la définition.

Une question pratique qui se pose très souvent dans la description des langues est la question du choix entre plusieurs analyses possibles, à partir du moment où la relation morphologique que l'on décrit ne s'explique pas de manière évidente par une opération d'affixation. Par exemple en anglais, la relation entre des formes de présent et de passé telles que **win** [wɪn] / **won** [wɒn] est usuellement décrite en termes de modification de la voyelle d'une base **wɪn** coïncidant avec la forme de présent, mais on pourrait a priori imaginer de décrire de telles formes de passé comme résultant de l'adjonction d'un infixé vocalique qui s'amalgame à la voyelle de la base **wɪn**, ou comme résultant d'un mécanisme de transfixation avec une base discontinue **w-n** qui doit nécessairement se combiner avec un morphème vocalique intercalé entre ses deux fragments. Et quand on fait face à de telles données dans la langue qu'on décrit, il faut évaluer les avantages et les inconvénients respectifs des différents traitements imaginables.

Ainsi, on a souvent à se poser la question de savoir dans quelle mesure une opération morphologique qui ne relève pas de manière évidente de la concaténation de morphèmes ne peut pas s'expliquer à partir d'une structure sous-jacente de type concaténatif à laquelle s'appliquent des règles ayant pour effet de brouiller la structure des mots. En liaison avec ceci, on doit souvent se poser la question de savoir dans quelle mesure l'hypothèse de structures sous-jacentes abstraites ne permettrait pas d'unifier la présentation d'opérations morphologiques à première vue différentes bien que remplissant la même fonction.

Par exemple en soninké, la relation entre la forme la plus simple des verbes (utilisée notamment à l'accompli et à l'impératif) et la forme de gérondif (qui

s'utilise notamment comme forme d'inaccompli en combinaison avec la copule locative) peut se présenter de deux façons différentes :

(18) La formation du gérondif en soninké

a.	séllà	'balayer'	gér.	séllà-ná
	bónté	'sauter'	gér.	bónté-né
	bóosí	'sucér'	gér.	bóosí-ní
	dàró	'respecter'	gér.	dàró-nó
b.	wùtú	'prendre'	gér.	wùttú
	mùgú	'entendre'	gér.	mùkkú
	ɲàrí	'voir'	gér.	ɲàllí
	kátú	'frapper'	gér.	káttá

Les gérondifs du type (a) s'expliquent de façon évidente par l'adjonction d'un suffixe **-nV́** à la forme nue du verbe, **V́** notant une voyelle 'sous-spécifiée' à ton haut qui copie la voyelle finale du lexème, avec comme seule complication que les verbes qui se terminent à la forme nue par une séquence tonale BH ont au gérondif un allomorphe dont le contour tonal se termine par BB. Quant aux gérondifs du type (b), on peut les décrire comme résultant d'une opération qui transforme une consonne médiane simple en consonne géminée (et qui en plus, lorsque les deux voyelles sont différentes, assimile parfois la deuxième voyelle à la première). Mais on peut aussi ramener la totalité des gérondifs du soninké à une forme sous-jacente où un suffixe **-nV́** est ajouté au lexème. Il suffit pour cela de poser que pour une partie des verbes, la suffixation peut s'accompagner de l'élision de la voyelle finale du lexème, à la suite de quoi le groupe de consonnes ainsi créé subit une transformation qui en fait une consonne géminée :

wùtú + -nV́	→	wùtù-nú	→	wùt-nú	→	wùt-tu
mùgú + -nV́	→	mùgù-nú	→	mùg-nú	→	mùg-gú → mùk-kú
kátú + -nV́	→	kát-nV́	→	kát-ná	→	kát-tá
ɲàrí + -nV́	→	ɲàrì-ní	→	ɲàr-ní	→	ɲàr-lí → ɲàl-lí

On est là dans un domaine où il n'est pas facile de tracer une limite stricte entre la description synchronique et la reconstitution de l'histoire de la langue, car il est bien connu que les difficultés à isoler les éléments constitutifs des mots complexes ont souvent leur source dans des changements phonétiques.

7. Spécificité de la morphologie tonale

7.1. Remarques introductives

Dans les manuels élémentaires, la notion de ton est généralement introduite en référence à des langues comme le chinois, où la seule fonction du ton est de

distinguer des lexèmes de forme segmentale identique. Mais à partir du moment où le ton existe dans une langue, il peut avoir une fonction morphologique.

Dans une langue à tons, les processus morphophonologiques à la jonction de deux formatifs d'un mot peuvent comporter une interaction tonale. Par exemple en soninké, les verbes de contour tonal BH combinés au morphème de gérondif **-nŵ** ont leur contour tonal modifié en BB – ex. (18) ci-dessus.

Il peut arriver aussi que deux affixes de même forme segmentale se distinguent par leur ton (comme **-ndí** 'causatif' et **-ndi** antipassif en soninké, ou encore **lù-** 'sujet de 2ème personne du pluriel' et **lú-** 'sujet de 3ème personne classe 11' dans certains paradigmes de conjugaison du tswana).

Mais le ton peut être impliqué dans la morphologie de manière beaucoup plus fondamentale, car toutes les distinctions grammaticales qu'on peut trouver véhiculées selon les langues par des alternances affectant les consonnes ou les voyelles d'une base peuvent dans d'autres langues se trouver véhiculées par une alternance tonale. L'exploitation du ton à des fins morphologiques est particulièrement commune dans les langues d'Afrique subsaharienne.

Lorsque deux formes d'un même paradigme ou deux lexèmes apparentés se distinguent seulement par un mécanisme régulier de changement tonal, on peut parler de manière générale de *morphotonème*. Mais pour une analyse précise des faits de morphologie tonale il est important de faire la distinction entre *morphotonèmes additifs* et *morphotonèmes substitutifs*.

7.2. Morphotonèmes additifs

On peut parler de morphotonème additif lorsque dans une alternance tonale à valeur morphologique, la courbe tonale du mot prise globalement (c'est-à-dire, en faisant abstraction du nombre de syllabes successives qui peuvent être associées à un même ton ponctuel et en décomposant les modulations comme séquences de tons ponctuels) comporte systématiquement un élément de plus pour l'une des deux formes analysées. Par exemple en (19), nous voyons que le contour global de la forme déterminée des noms en malinké de Kita se présente invariablement comme le contour de la forme non-déterminée plus un élément B final. L'opération morphologique qui rend compte de cette forme est donc l'adjonction d'un morphotonème B final suivie de règles d'association de ce ton bas qu'il n'est pas utile de discuter en détail ici.

(19) Formes non-déterminée et déterminée du nom en malinké de Kita¹⁰

sú	H	sû	HB	'maison'
sù ~ sũ	B(H)	sũ	BHB	'cheval'
básá	H	básà	HBH	'lézard'

¹⁰Dans cet exemple, chaque fois qu'une modulation descendante est notée, elle n'apparaît telle quelle que devant pause. A l'intérieur d'un groupe prosodique, elle laisse la place à un ton haut, mais l'élément bas de la modulation continue à se manifester par un abaissement tonal (downstep) si le mot suivant commence par un ton haut. Dans les mêmes conditions, la double modulation montante-descendante laisse la place à une modulation montante suivie d'un downstep.

kúlún	H	kúlûn	HB	‘bateau’
bànbá	BH	bànbâ	BHB	‘crocodile’
kùlùn ~ kùlún	B(H)	kùlûn	BHB	‘bateau’
húgúlá	H	húgúlà	HB	‘bonnet’
nógólán	H	nógólân	HB	‘feuille (sp.)’
jàngúmá	BH	jàngúmà	BHB	‘bateau’
sùntúkún	BH	sùntúkûn	BHB	‘dépotoir’
dèrègè ~ dèrègé	B(H)	dèrègê	BHB	‘bateau’
tùbàbù ~ tùbàbú	B(H)	tùbábù	BHB	‘européen’
bàràrà ~ bàràrà	HB(H)	bàràrà	HBHB	‘marmite’

Donc, si la morphologie tonale n’est jamais à strictement parler concaténative, avec les morphotonèmes additifs on n’en est tout de même pas très loin, puisque dans un cas comme celui de la forme déterminée des noms en malinké de Kita, l’adjonction d’un morphotonème B qui s’associe à la dernière voyelle du nom ressemble beaucoup à une suffixation. D’ailleurs très souvent, lorsqu’on dispose de données comparatives, on peut voir qu’historiquement, les morphotonèmes additifs ont pour origine un affixe dont la forme segmentale a disparu et dont seul le ton a subsisté. C’est en particulier le cas pour le morphotonème bas des noms déterminés en malinké de Kita. En effet, ce morphotonème a pour origine un suffixe **-ò** qui est maintenu dans d’autres parlars mandingues (cf. par exemple en mandinka **súu** / **súw-ò** ‘maison’, **kúlúnj** / **kúlúnjò** ‘bateau’, etc.).

7.3. Morphotonèmes substitutifs

Les morphotonèmes substitutifs remplacent le contour tonal de la base sur laquelle ils opèrent par un autre contour tonal, généralement plus simple que le contour tonal de départ. A la différence des morphotonèmes additifs, leur action n’est pas ‘locale’, et peut au contraire s’étendre à la totalité du mot indépendamment de sa longueur.

La morphologie du hausa fait un large usage de morphotonèmes substitutifs, notamment dans la formation du pluriel. L’ex. (20) illustre l’un des types de formation du pluriel en hausa, qui d’une part substitue **ai** à la voyelle finale du nom, et d’autre part substitue au schème tonal de départ un schème tonal BH avec le H associé à la dernière syllabe et le B associé à toutes les autres syllabes ; pour certains des noms qui forment leur pluriel de cette façon il se produit en outre une reduplication, mais cela ne change rien à l’aspect tonal des choses.

(20) Morphotonème substitutif dans les pluriels hausa en **-ai** ¹¹

ɗàlìlì	BHB	pl.	ɗàlìláì	BBH	‘raison’
ɗàlibí	HBH	pl.	ɗàlibáì	BBH	‘élève’
àbòkì	BHH	pl.	àbòkáì	BBH	‘ami’
lìttáfì	HHB	pl.	lìttàttáfáì	BBBH	‘livre’

¹¹ En hausa, le macron note la longueur vocalique.

Le soninké est une autre langue qui utilise largement des opérations morphologiques consistant en une substitution tonale ou faisant intervenir une substitution tonale. On peut distinguer en soninké deux morphotonèmes substitutifs. L'un de ces deux morphotonèmes substitue au schème tonal lexical de la base un contour tonal entièrement B. Il intervient dans la flexion verbale pour donner une forme utilisée en combinaison avec certains marqueurs de négation, lorsque la partie de la phrase qui précède le verbe comporte un groupe nominal focalisé ou une proforme interrogative, et dans les phrases hypothétiques – ex. (21). Il intervient aussi dans la flexion des numéraux de 2 à 10, pour donner la forme prise par ces numéraux lorsqu'ils se combinent à un nom – ex. (22).

(21) Substitution tonale dans la flexion verbale du soninké ¹²

a. **Ŋ́ ɲá ké kónpé sèllà-ná.**
1SG INACP DEM pièce balayer-GER
'Je balaie cette pièce.'

b. **Ŋ́ ntá ké kónpé sèllà-nà.**
1SG INACP.NEG DEM pièce balayer-GER
'Je ne balaie pas cette pièce.'

a. **Kó nà ké kónpé sèllà-nà?**
1SG INACP DEM pièce balayer-GER
'Qui est-ce qui balaie cette pièce ?'

a. **Hàatú yà nà ké kónpé sèllà-nà.**
Fatou FOC INACP DEM pièce balayer-GER
'C'est Fatou qui balaie cette pièce.'

(22) Substitution tonale dans la flexion des numéraux en soninké

səlìŋŋú 'poulets' + **sikkó** 'trois' → **səlìŋŋú sikkì** 'trois poulets'

L'autre morphotonème substitutif du soninké a le même effet que le premier sur les bases monosyllabiques (auxquelles il impose un ton B), mais il substitue au schème tonal des bases de deux syllabes ou plus un schème tonal BH avec H associé à la syllabe finale et B associé à toutes les autres syllabes. Il intervient dans la flexion nominale pour donner une forme utilisée notamment lorsque le nom est la tête d'une construction génitive – ex. (23), ainsi que dans la flexion adjectivale, pour donner la forme des adjectifs qui est utilisée en fonction prédicative – ex. (24).

¹² DEM = démonstratif, FOC = marqueur de focus, GER = gérondif, INACP = inaccompli, NEG = négation, SG = singulier.

(23) Substitution tonale dans la construction génitive du soninké

qálisí ‘argent’ → **Múusá qálisí** ‘argent de Moussa’
kitàabè ‘livre’ → **Múusá kitàabé** ‘livre de Moussa’

(24) Substitution tonale dans l’emploi prédicatif des adjectifs en soninké

a. **yàxàrì háránpàré** ‘belle femme’
femme beau
(l’adjectif **háránpàré** ‘beau’ a sa tonalité inhérente)

b. **Yàxàrén pàrànpàrén nì.** ‘La femme est belle.’
femme.D beau.D COPEQ¹³
(la tonalité de l’adjectif est modifiée par un morphotonème substitutif)

7.4. Propagation tonale, glissement tonal et l’analyse du rôle morphologique des tons

Dans certaines langues (notamment dans beaucoup de langues bantoues de l’est et du sud) l’analyse du rôle morphologique des tons se trouve compliquée par des mécanismes systématiques et généraux de propagation tonale ou de glissement tonal qui ont comme conséquence qu’un ton introduit dans la structure du mot par un morphème particulier peut se manifester de façon tout à fait automatique très loin du morphème en question, parfois même sur un autre mot. Par exemple, il n’est pas très difficile d’établir qu’en swati, la forme verbale **basebentelána** ‘ils travaillent les uns pour les autres (cj.)’ (avec un unique ton haut sur la syllabe pénultième¹⁴) résulte de la concaténation des morphèmes suivants :

bá-	sébent	-el	-an	-a
CL2	travailler	APPL	RECIP	FIN ¹⁵

Autrement dit, deux des morphèmes qui constituent ce mot comportent structurellement une voyelle à ton haut, mais en surface, on a un unique ton haut associé à une voyelle qui appartient à un autre morphème. Mais ici, il ne faut chercher aucun processus morphologique particulier. La relation entre tons sous-jacents et tons hauts réalisés est un problème purement phonologique, qui met en jeu des processus généraux de propagation et/ou glissement tonal opérant indépendamment de la structure du mot en morphèmes.

Dans des situations de ce genre, une fois établies les règles générales de propagation ou glissement des tons, il n’est généralement pas difficile d’établir à quel morphème appartiennent les tons qui se manifestent en surface. Mais si on ne

¹³ D = marque de détermination nominale, et COPEQ = copule équative – le soninké distingue une copule de localisation (**wá**) et une copule équative (**ní**), et c’est la copule équative qui s’emploie dans la prédication adjectivale.

¹⁴ Dans la transcription du swati, l’absence de diacritique indique un ton bas.

¹⁵ APPL = applicatif, CL = classe nominale, FIN = finale, RECIP.

fait pas au départ les bonnes hypothèses, on risque de perdre beaucoup de temps à essayer d'analyser en termes d'alternances tonales morphologiques des variations qui tiennent en fait à des processus purement phonologiques ayant pour effet que des tons structurellement associés à un morphème peuvent se manifester très loin du morphème auquel ils appartiennent.

8. Formes libres et formes liées, affixes et clitiques

8.1. La distinction entre formes libres et formes liées

Il est souvent difficile de distinguer les formes grammaticales liées (clitiques ou affixes) des mots grammaticaux, car ces derniers manifestent universellement une tendance à perdre leur autonomie pour devenir des formes liées, mais ce processus est progressif, et il est souvent délicat de caractériser de façon satisfaisante des éléments grammaticaux qui y sont déjà engagés sans être encore parvenus à l'étape finale. Il convient à ce propos de ne pas se laisser exagérément influencer par les traditions grammaticales et les conventions orthographiques, qui ne traitent pas toujours cette question de façon cohérente. Par exemple, l'orthographe française et la grammaire traditionnelle du français traitent comme mots à part les unités suivantes, qui ont pourtant plus ou moins clairement des propriétés de formes liées :

- les pronoms personnels conjoints (**je**, **tu**, **me**, etc.), qui s'attachent au verbe ;
- le marqueur de négation **ne**, qui s'attache au premier mot du groupe verbal¹⁶ ;
- les prépositions **de**, **à** et **en**, les articles, les déterminants démonstratifs et les déterminants possessifs, qui s'attachent au premier mot du constituant nominal.

Le terme de clitique n'a pas vraiment de définition précise, et il est utilisé de façon assez variable d'une tradition à une autre, mais son utilisation répond au souci d'enregistrer l'existence de formes dont le comportement par rapport aux formes adjacentes n'est, ni tout à fait celui d'un mot, ni tout à fait celui d'un affixe ou d'une séquence d'affixes. Beaucoup de travaux récents se sont efforcés d'élaborer une typologie des formes que les descriptions de langues désignent comme clitiques, mais pour l'instant la seule conclusion évidente est qu'il s'agit d'un problème qui n'a pas de solution simple. Il n'y a notamment aucun consensus sur celles parmi les formes couramment désignées comme clitiques qu'il pourrait être justifié de reclasser comme affixes.

Une forme (simple ou complexe) est libre ou liée selon qu'elle peut ou non constituer à elle seule un mot. Une forme libre est minimale si toute suppression d'un de ses éléments constitutifs lui enlève son caractère de mot.

Un premier type de critère de reconnaissance de la distinction entre formes libres et formes liées a trait à l'indépendance prosodique des formes libres. Une forme qu'il est naturel d'énoncer en isolation, par exemple en répondant à une question, est nécessairement une forme libre. Si une phrase est interrompue par des pauses, les

¹⁶ En français, le premier mot du groupe verbal est presque toujours le verbe, mais la nécessité de préciser que **ne** se rattache au premier mot du groupe verbal plutôt qu'au verbe apparaît à l'infinif - cf. **ne-jamais le-lui-donner**.

pauses tendent à coïncider avec des limites entre formes libres. Il en va de même pour l'insertion d'incises. Les formes libres se prêtent plus facilement que les formes liées à une mise en relief par l'intonation. Dans les langues à accent de mot, seules les formes libres peuvent constituer des unités accentuelles autonomes ; si une forme liée porte un accent, cet accent ne lui appartient pas en propre, il est déterminé dans le cadre de l'unité accentuelle à laquelle la forme liée en question est rattachée.

Un deuxième type de critère repose sur le fait que les possibilités d'insertion d'éléments supplémentaires entre deux éléments d'un même mot sont limitées à un petit nombre de formatifs ayant eux-mêmes des caractéristiques de formes liées. Ainsi, **-ra** dans **L'avion atterri-ra** et **va** dans **L'avion va atterrir** apportent également une information de type temporel sur l'événement représenté par le verbe **atterrir** et une information sur la personne et le nombre du sujet de ce verbe, mais le fait que **-ra** est suffixé à la base verbale **atterri-**, alors qu'il serait incorrect de considérer **va** comme préfixé à l'infinitif, se manifeste notamment par la différence de comportement lors de l'insertion d'adverbes : on peut dire **L'avion va probablement bientôt atterrir**, mais pas ***L'avion atterri- probablement bientôt -ra**.

Un troisième type de critère repose sur l'impossibilité de traiter isolément les formes liées dans des opérations auxquelles se prêtent les mots ou groupes de mots, telles que le clivage ou la coordination. Cette impossibilité est particulièrement remarquable lorsque la forme qui donne lieu à un tel blocage contribue à la signification globale de la phrase de façon équivalente à un mot ou groupe de mots se prêtant aux opérations en question.

On peut notamment évoquer ici le cas des pronoms personnels conjoints du français. Il ne fait de doute pour personne que les pronoms personnels disjoints du français sont des mots autonomes, qui ont une distribution syntaxique de constituants nominaux. Par contre, les pronoms personnels conjoints, bien qu'ils s'écrivent la plupart du temps comme des mots à part, sont des formes liées, comme le prouve l'impossibilité qu'il y a à les coordonner (on peut dire **Marie me connaît**, **Marie te connaît**, mais pas ***Marie me et te connaît**) et à les faire apparaître dans la première partie d'une construction clivée (on peut transformer **Marie cherche Jean** en **C'est Jean que Marie cherche**, alors que **Marie me cherche** ne peut pas se transformer en **C'est me que Marie cherche**)¹⁷.

Un quatrième type de critère repose sur la possibilité d'avoir entre une forme liée et son environnement immédiat des interactions phonologiques spécifiques, autres que celles se produisant régulièrement aux frontières de mots. On peut citer comme exemple l'effacement du **a** de l'article défini féminin singulier du français devant une partie des mots commençant par voyelle ou semi-voyelle.

¹⁷ Le caractère lié des pronoms conjoints du français se manifeste aussi par l'impossibilité de les séparer du verbe par une incise (***Il, à ce que j'ai entendu dire, a l'intention de changer de voiture**) et par des interactions phonologiques qui n'ont pas d'équivalent lorsque le verbe est précédé d'autre chose que d'un pronom conjoint : liaison obligatoire, possibilité d'élision du **u** de **tu** devant voyelle, etc. Les mêmes critères appliqués aux pronoms personnels de l'anglais permettent de les reconnaître uniformément comme formes libres, contrairement au français, où seuls les pronoms personnels disjoints ont ce statut (cf. par exemple **You also have a house in Paris**, **You and I know what should be done**).

Mais l'application de ces critères donne parfois des résultats peu nets, ou difficiles à interpréter. Il faut notamment être très prudent avec les tests d'insertion, qui conduisent souvent à des raisonnements circulaires, car leur interprétation est largement dépendante d'une hypothèse sur ce qu'il convient de considérer comme base de rattachement d'une forme liée.

En cas de doute, l'existence d'une interaction phonologique spécifique doit être considérée comme décisive. Par exemple, on peut hésiter à reconnaître comme forme liée l'article défini du français, car si on considère comme base de rattachement potentielle le nom qui est la tête du groupe nominal, on doit conclure que le test d'insertion donne un résultat négatif (cf. par exemple **la triste histoire**, où un adjectif s'intercale entre l'article et le nom) ; mais l'existence de processus phonologiques spécifiques entre l'article défini et ce qui lui succède immédiatement (notamment l'élision possible de **a** au féminin singulier : **l'étonnante histoire**) conduit à reconsidérer la question, et à reconnaître l'article défini comme attaché au premier mot du constituant nominal.

8.2. Mots grammaticaux, clitiques et affixes

Les formes reconnaissables selon les critères précédents comme liées ne manifestent pas toujours au même degré des caractéristiques d'intégration à un mot qui les engloberait. Notamment, on compliquerait inutilement la description en traitant comme affixes des formes liées qui occupent dans la linéarité de la phrase la même place qu'une forme libre sémantiquement équivalente. Dans de tels cas, il est plus simple de poser qu'on a fondamentalement affaire à des mots dont la position dans la phrase est déterminée par les règles de la syntaxe, mais des mots susceptibles de se *cliticiser*, c'est-à-dire de perdre leur autonomie prosodique pour se rattacher à un mot plein auquel ils sont adjacents. C'est par exemple le cas en anglais pour **'ve** forme réduite de **have**, ou **'s** forme réduite de **is**. On parle souvent dans ce cas de clitiques *simples*, ou *post-syntaxiques*.

Inversement, en ne classant pas comme affixe une forme grammaticale liée qui s'attache invariablement à un mot plein d'un type précis selon des modalités qui ne font pas intervenir l'environnement syntaxique, on a toutes les chances de compliquer inutilement la description.

Toute la question est de savoir comment traiter les formes grammaticales liées qui ne relèvent pas de manière évidente de l'une de ces deux situations extrêmes, et à partir de quel moment il est préférable de considérer une forme grammaticale liée, soit comme la forme cliticisée d'un mot dont la place dans la phrase est déterminée par la syntaxe au même titre que celle de n'importe quel autre mot, soit comme un affixe qui n'est concerné par les règles de la syntaxe que dans la mesure où sa présence peut avoir une incidence sur les propriétés syntaxiques du mot dont il fait partie, soit comme nécessitant un traitement spécifique.

En particulier, on peut envisager d'analyser comme *affixes de constituant* des formes grammaticales liées qui ne sont, ni la variante cliticisée d'un mot grammatical existant par ailleurs sous forme libre, ni un affixe qui n'aurait d'existence que comme élément constitutif d'un type précis de mot, mais qui s'attachent systématiquement au premier ou au dernier mot d'un constituant d'un

type donné sans tenir compte de la nature précise de ce mot.

Par exemple, le marqueur de génitif 's de l'anglais est indiscutablement une forme liée (ne serait-ce que du fait qu'il est impossible de le détacher dans la prononciation), mais ce n'est pas un affixe typique, car il peut s'attacher à n'importe quel type de mot, pourvu que ce mot termine un constituant nominal (cf. par exemple **the man we saw's car**, **the man we saw yesterday's car**). On peut donc *a priori* envisager de l'analyser, soit comme une postposition se cliticisant nécessairement sur le dernier mot du constituant nominal, soit comme un suffixe susceptible de s'attacher à n'importe quel mot qui termine un constituant nominal, et qui détermine la façon dont le constituant nominal au dernier mot duquel il est attaché peut participer à la construction d'une phrase.

Il n'est pas possible d'aller ici très loin dans une question aussi complexe et aussi controversée que la distinction entre différents types possibles de formes grammaticales liées, aussi on se limitera dans ce qui suit à indiquer les critères utilisables pour fonder un classement des formes grammaticales liées d'une langue, selon leur degré d'intégration à leur base de rattachement, comme clitiques (faiblement intégrés) ou affixes (fortement intégrés).

Le fait qu'une forme liée dépende phonologiquement de manière spécifique de la base à laquelle elle est rattachée n'est pas un critère de classement, car il s'agit d'une propriété générale des formes liées. Par contre, le fait qu'une forme liée impose à sa base de rattachement des processus phonologiques autres que ceux se produisant régulièrement à une frontière de mots peut être considéré comme un critère d'affixation. En effet, la base à laquelle se rattache un clitique est en principe une forme libre, c'est-à-dire un mot déjà constitué, qui ne subit en présence du clitique aucune modification spécifique, alors qu'il est courant que les affixes aient pour base une forme qui n'existe pas par elle-même comme mot.

Par exemple en turc, **da** ~ **de** marque du cas locatif et **da** ~ **de** marque discursive (traduisible selon le contexte comme 'aussi', 'à son tour', etc.) sont également soumis à l'harmonie vocalique, ce qui impose de reconnaître ces deux formes comme liées. Mais la première peut imposer à sa base de rattachement une forme qui n'existe pas comme mot isolé, alors que la deuxième s'attache toujours à une forme qui existe par ailleurs comme mot. Par exemple, avec le démonstratif **o**, on aura **on-da** 'dans cela' (**on-** étant un allomorphe du démonstratif conditionné par le rattachement d'un suffixe) mais **o da** 'cela aussi', ce qui peut justifier de les classer respectivement comme suffixe et comme clitique¹⁸.

De manière analogue, il y a en anglais une différence nette entre le marqueur de pluriel **-s**, affixe du nom qui peut imposer une modification de sa base (comme dans **life**, pl. **lives**) et le marqueur génitif **'s**, qui peut s'attacher à n'importe quel mot qui termine un constituant nominal sans jamais entraîner de modification de ce mot.

Dans le même ordre d'idées, dans les langues à accent de mot, il est courant qu'affixes et clitiques diffèrent par leur comportement dans les règles qui déterminent la place de l'accent, mais les détails de la question diffèrent trop d'une langue à l'autre pour qu'on puisse proposer une quelconque généralisation.

¹⁸ Cette analyse est confirmée par le fait que la marque discursive **da** ~ **de** ne sélectionne pas pour hôte des mots d'une catégorie particulière, ce qui est généralement considéré comme une propriété typique des clitiques.

On mentionne aussi parfois comme critère possible que par rapport à leur base de rattachement, les clitiques peuvent connaître des variations de positionnement qu'ignorent en principe les affixes. Ce critère est notamment régulièrement mis en avant par les linguistes francisants qui se refusent à analyser les pronoms personnels conjoints du français comme affixes du verbe et acceptent tout au plus d'en faire des clitiques. Mais dans de très nombreuses langues, y compris parmi les mieux documentées, il y a consensus pour analyser comme affixes des marques grammaticales susceptibles d'apparaître tantôt comme préfixes, tantôt comme suffixes (c'est le cas par exemple des marques de personne du verbe en arabe, basque, géorgien, etc.). On peut soutenir (et c'est la position défendue ici) que le critère décisif pour s'autoriser à reconnaître comme affixes des formes liées qui n'apparaissent pas toujours du même côté de leur base de rattachement est que le choix se fasse au niveau du mot auquel appartient l'afixe, sans intervention du contexte syntaxique, alors que pour un clitique, le choix entre enclise et proclise peut être conditionné par l'environnement syntaxique. De ce point de vue, il n'y a aucun obstacle à analyser comme affixes les pronoms conjoints du français, à la différence par exemple des pronoms conjoints du portugais, dont la position dépend de la présence ou de l'absence de certains mots à gauche du verbe, et peut donc difficilement être traitée par une règle morphologique.

Mais, sans que ceci remette en question l'intérêt de tous ces critères, il faut rester conscient du fait que leur application rigoureuse et systématique à une langue donnée révèle souvent une situation plus complexe qu'une simple dichotomie entre une classe de formes grammaticales liées peu intégrées (clitiques) et une classe de formes grammaticales liées fortement intégrées (affixes). Deux critères différents peuvent donner pour la même forme des résultats qui se contredisent. Il n'est donc pas très étonnant qu'il n'y ait guère de cohérence dans les décisions prises d'une tradition grammaticale à une autre quant au traitement de formes grammaticales qui ne sont, ni des mots tout à fait normaux, ni des affixes tout à fait typiques. Les difficultés d'analyse se reflètent d'ailleurs dans les conventions orthographiques, qui font parfois apparaître comme mots à part des formes liées suffisamment intégrées pour être analysées comme affixes, alors que par ailleurs il leur arrive d'enregistrer le rattachement de formes liées qui sont de manière évidente la variante cliticisée d'un mot existant aussi sous forme libre.

Références

- Creissels, Denis & Pierre Sambou. 2013. *Le mandinka : phonologie, grammaire, textes*. Paris : Karthala.
- Haspelmath, Martin. 2002. *Understanding Morphology*. London : Arnold.
- Haspelmath, Martin. 2011. 'The indeterminacy of word segmentation and the nature of morphology and syntax'. *Folia Linguistica* 45-1. 31-80.