

Typologie du langage (Comrie 1981) p.39 à 50

2.3 Typologie morphologique

Bien qu'un certain nombre de bases pour des typologies globales aient été proposées sur l'histoire des études typologiques, il en existe deux qui sont particulièrement importantes, du moins d'un point de vue historique. La première d'entre eux, la typologie morphologique, a été prédominante durant le XIXe et au début du XXe siècle, même si elle conserve également une place bien établie dans le texte-livre de linguistique générale, ce qui fait l'objet de la présente section. La seconde, "word order typology" (la typologie de l'ordre des mots), est discutée dans le chapitre 4. Bien que la vue exprimée dans le présent ouvrage est qu'aucune d'elle ne fournit en fait la base d'une typologie globale, chacune d'elles peut servir à fournir des typologisations d'une partie importante de la structure de la langue.

Bien que la typologie morphologique a une longue histoire, remontant au moins au début du XIXe siècle, il y a eu une tendance pour certains des principes de cette typologie à se fossiliser, et dans la présente section, nous voulons non seulement rendre compte de l'habitude traditionnelle concernant la typologie morphologique, mais aussi se pencher sur certaines améliorations qui peuvent et doivent être faites si le meilleur parti possible doit être établi à partir de cette voix des langues typologisants. Mais d'abord, nous examinerons les classifications traditionnelles.

La typologie morphologique reconnaît généralement trois types canoniques de langues : l'isolante, agglutinante et flexionnel, auxquelles on en ajoute parfois un quatrième : polysynthétique (ou incorporante). Une langue isolante est celle qui n'a pas de morphologie, c'est à dire au moins idéalement, une langue où il y a une correspondance 1 pour 1 entre les mots et les morphèmes. Un exemple de langue qui se rapproche de ce type d'isolement est le vietnamien, comme on peut l'illustrer par la phrase suivante:

[Page 40]

Khi toi den nha ban toi, chung toi bat dau lam bai.
Quand je viens maison ami I pluriel I commence faire leçon.
« Quand je me rendais à la maison de mon ami, nous commençons à faire nos leçons. »

Chaque mot de cette phrase est invariable, il n'y a pas de variation morphologique pour la conjugaison (viens/venais, commence/commencer) ou pour le cas (le vietnamien a *toi* pour les deux, « je » et « moi ») ; peut-être même plus frappant, le pluriel est indiqué, dans le cas du pronom, par l'addition d'un mot séparé plutôt que par un signifiant morphologique, ainsi le pluriel de *toi* « je » est *chung toi* « nous ». Par ailleurs, chaque mot consiste en général en un seul morphème, avec l'exception possible de *bat*

dau « commence », qui est de façon discutable un mot reposant sur certains de ces critères, p. ex. unité de signification, bien qu'il puisse être segmenté, au moins étymologiquement, en deux morphèmes : *bat* « saisir » et *dau* « tête » ; nous devrions creuser en profondeur pour établir si oui ou non on a en fait une correspondance un pour un entre les mots et les morphèmes.

Dans certaines discussions sur la typologie morphologique, on emploie le terme de langue monosyllabique, en plus ou à la place du terme de langue isolante, les deux paramètres ont un principe distinct, et le sujet de la typologie morphologique est la structure d'isolation qui en relève. Cependant on peut imaginer une langue où il n'y aurait pas de morphologie mais où chaque mot (= morphème) pourrait consister en un certain nombre de syllabes. Inversement, on pourrait imaginer une langue avec de la morphologie mais où la morphologie serait limitée au changement dans les consonnes et dans les tons, sans affecter la nature monosyllabique du mot. Nous retiendrons par conséquent le terme « isolante » ici.

Dans une langue agglutinante, un mot peut consister en plus d'un morphème, mais les liens entre les morphèmes du mot sont toujours nets ; par ailleurs, un morphème a raisonnablement au moins une forme invariable, ainsi l'identification des morphèmes dans le cadre de leur forme phonétique est aussi simple. Le turc nous servira comme exemple en nous basant sur la déclinaison du nom. En turc, les noms varient en nombre et en cas (et aussi d'autres paramètres non traités ici, p. ex. possession), avec un système de deux nombres (singulier, pluriel) et six cas (nominatif, accusatif, génitif, datif, locatif, ablatif). Cependant, pour la forme d'un nom donné, il est toujours possible de segmenter clairement le radical de l'affixe du nombre (zéro au singulier, *-lar* au pluriel), et l'affixe du cas (zéro au nominatif, *-i* à l'accusatif, *-in* au génitif, *-a* au datif, *-da* au locatif, *-dan* à l'ablatif), comme dans le paradigme suivant du mot *adam* « homme » :

[Page 41]

	Singulier	Pluriel
Nominatif	<i>adam</i>	<i>adam-lar</i>
Accusatif	<i>adam-i</i>	<i>adam-lar-i</i>
Génitif	<i>adam-in</i>	<i>adam-lar-in</i>
Datif	<i>adam-a</i>	<i>adam-lar-a</i>
Locatif	<i>adam-da</i>	<i>adam-lar-da</i>
Ablatif	<i>adam-dan</i>	<i>adam-lar-dan</i>

(Notez que l'affixe du pluriel précède toujours l'affixe du cas). Comme le suggère le terme agglutinant (cf. Latin *gluten* « colle »), c'est comme si les différents affixes étaient juste collés les uns après les autres (ou l'un avant l'autre dans le cas des préfixes).

Dans les langues flexionnelles, cependant, il n'y a pas de liens si précis entre les morphèmes, la caractéristique d'une langue flexionnelle étant que l'expression de différentes catégories avec le même mot sont fusionnées ensemble pour donner un simple morphe non-segmentable. Cela peut être illustré par la déclinaison du russe : le russe a deux nombres (singulier, pluriel), et six cas (nominatif, accusatif, génitif, datif, instrumental, prépositionnel). En russe, par ailleurs, même les affixes unis n'ont pas de forme invariable, puisque dans les différentes classes de déclinaisons, différents affixes

sont utilisés. Cela est illustré ci-dessous avec les formes déclinées du nom *stol* « table » (déclinaison Ia) et *lipa* « citronnier vert » (déclinaison II) :

[Page 41]

	Ia		II	
	Singulier	Pluriel	Singulier	Pluriel
Nominatif	<i>stol</i>	<i>stol-y</i>	<i>lip-a</i>	<i>lip-y</i>
Accusatif	<i>stol</i>	<i>stol-y</i>	<i>lip-u</i>	<i>lip-y</i>
Génitif	<i>stol-a</i>	<i>stol-ov</i>	<i>lip-y</i>	<i>lip</i>
Datif	<i>stol-u</i>	<i>stol-am</i>	<i>lip-e</i>	<i>lip-am</i>
Locatif	<i>stol-om</i>	<i>stol-ami</i>	<i>lip-oj</i>	<i>lip-ami</i>
Ablatif	<i>stol-e</i>	<i>stol-ax</i>	<i>lip-e</i>	<i>lip-ax</i>

Clairement, il n'y a pas de possibilité pour qu'une forme comme le génitif pluriel *stol-ov* « des tables » puisse être segmentée en un affixe du nombre et un affixe du cas, mais plutôt que l'affixe plein *-ov* est un affixe simple combinant l'expression du cas et du nombre (un morphe-portemanteau). Et même en sachant que *-ov* est l'affixe du génitif pluriel dans la déclinaison Ia, nous n'avons pas de moyen de prédire l'affixe du génitif pluriel dans la déclinaison II, qui peut être zéro.

À la place du terme fusionnel, on trouve parfois le terme flexionnel. Cela n'est pas le cas dans le présent travail afin d'éviter une potentielle confusion terminologique : les langues agglutinantes et fusionnelles, puisqu'opposés aux langues isolantes, ont des flexions, et c'est par conséquent équivoque d'utiliser un terme basé sur la flexion pour en désigner un seul des deux. La disponibilité du terme alternatif fusionnel résout ce dilemme terminologique.

Le quatrième type morphologique qui est parfois inclus, bien que n'ayant pas toujours de signification, est polysynthétique ou incorporant. Bien que ces deux termes soient parfois considérés comme interchangeables, il est possible et conseillé de faire une distinction entre eux. L'incorporation réfère à la possibilité de prendre un nombre de morphèmes lexicaux et de les combiner ensemble dans un simple mot. D'une manière limitée, c'est possible en anglais avec différents processus de composition, comme lorsque les morphèmes lexicaux *swim* et *suit* sont composés ensemble pour donner *swimsuit*. Dans certaines langues, cependant, ce procédé est extrêmement productif, donnant naissance à des mots extrêmement longs avec un grand nombre de morphèmes lexicaux, traduisant souvent des phrases entières d'anglais, comme le Chukchi *tə-meyŋə-levtə-pəŋt-ərkən* « J'ai un terrible mal de tête », qui contient trois morphèmes lexicaux : *meyŋ-* « grand », *levt-* « tête », et *pəŋt-* « tête », en plus du morphème grammatical *t-* (première personne du singulier sujet) et *-rkən* (imparfait).

La polysynthèse, cependant, réfère simplement au fait que, dans une langue de ce type, il est possible de combiner un grand nombre de morphèmes, qu'ils soient lexicaux ou grammaticaux, en un simple mot, correspondant souvent à une phrase entière d'anglais, comme en esquimau (yupik sibérien) *angya-ghlla-ng-yug-tuq* « il veut acquérir un gros bateau », littéralement « bateau-AUGMENTATIF-ACQUÉRIR-DÉSIDÉRATIF-3SINGULIER ». En esquimau, contrairement au chukchi, un mot donné contient seulement un morphème lexical, tous les autres étant grammaticaux, c.-à-d. que

l'esquimau est polysynthétique, mais pas incorporant. Nous voyons ainsi que l'incorporation est un cas spécial de polysynthèse, à savoir où les morphèmes lexicaux peuvent être combinés ensemble dans un simple ensemble polysynthétique, et nous devrions par conséquent utiliser polysynthétique comme terme générique pour ce type.

L'une des raisons de l'omission de la polysynthèse de beaucoup de listes de types morphologiques est que son inclusion détruit l'homogénéité de l'ensemble de la typologie morphologique. Bien que nous classifions les exemples du chukchi et de l'esquimau données au-dessus comme des instances de la polysynthèse, ils sont aussi clairement des instances de l'agglutination : dans l'exemple du chukchi, nous pouvons segmenter les morphèmes grammaticaux et lexicaux individuellement, et ceux-ci sont dans l'ensemble largement invariables (l'exception principale étant la largement prévisible occurrence de la voyelle *a* qui fait éclater le regroupement de consonnes, spécialement les liens du morphème) ; également, dans l'exemple de l'esquimau, nous pouvons segmenter facilement les suffixes grammaticaux isolés, et ce en gardant leur forme constante. (Dans une étude plus approfondie du chukchi et de l'esquimau, il deviendrait apparent que les deux langues ont aussi un certain taux de fusion, bien que ça n'interagisse pas avec la polysynthèse). Par conséquent l'agglutination et la polysynthèse représentent des paramètres différents, qui peuvent opérer indépendamment, plutôt que différentes valeurs du même paramètre.

Cependant, exclure la polysynthèse de la typologie morphologique pour cette raison n'est pas nécessairement justifié, en particulier parce-que les langues polysynthétiques fournissent le contrepoint des langues isolantes en terme de nombre de morphèmes par mot : dans les langues isolantes, chaque mot consiste en juste un morphème, tandis que dans les langues polysynthétiques, ou plutôt dans un idéal polysynthétique, chaque phrase consiste en juste un mot, ce mot à son tour consistant en plusieurs morphèmes, autant que nécessaires pour exprimer une signification. Avant tout, cela suggère d'abandonner la typologie morphologique en terme d'un simple paramètre qui serait désigné pour couvrir tous les types morphologiques, et opérer plutôt avec deux paramètres. Un de ces paramètres serait le nombre de morphèmes par mot, et ses deux extrêmes seraient l'isolation et la polysynthèse. L'autre paramètre serait quels morphèmes au sein d'un mot sont facilement segmentables, ces deux extrêmes étant l'agglutination (où la segmentation est simple) et la fusion (où il n'y a pas de segmentabilité). Nous pouvons référer à ces deux paramètres en tant qu'index de la synthèse et index de la fusion. Notez que l'index de fusion est, par définition, non pertinent en ce qui est des langues isolantes. Autrement, celles qui sont appelées traditionnellement langues polysynthétiques deviendraient des langues avec un haut index de synthèse (en plus, elles pourraient ou non avoir un haut index de fusion, bien que pour des raisons discutées plus haut il est inévitable qu'une langue avec un très haut index de synthèse ait aussi un bas index de fusion, même si les deux paramètres sont logiquement indépendants). La classe traditionnelle des langues agglutinantes correspond à celles avec un bas index de fusion (et, selon la quadruple classification, un index intermédiaire de synthèse, c.-à-d. ni isolante ni polysynthétique). Finalement, la classe traditionnelle de langues fusionnelles correspond à celles ayant un haut index de fusion (en gardant à l'esprit que les langues isolantes n'ont ni un haut ni un bas index de fusion : cet index ne s'applique simplement pas à elles).

La précédente discussion a subrepticement introduit un autre aspect de la typologie morphologique. Dès le début de notre discussion nous présumions, très simplement, que la typologie consisterait en trois ou quatre types idéals, parmi lesquels nous pourrions classer les langues du monde. En fait, cependant, bien que nous puissions établir ces types idéals, la majorité (peut-être toutes) des langues du monde ne correspondent pas exactement à l'un ou l'autre de ces types, mais tombent plutôt entre les deux extrêmes de chaque indice de synthèse et de fusion. Donc au lieu de fournir une typologie discrète, la typologie morphologique nous fournit une typologie continue, c.-à-d. que pour une langue donnée nous pouvons lui assigner une place parmi le continuum défini par l'index de synthèse et l'index de fusion. Nous devrions illustrer cela, et certains des autres problèmes qui nous attendent, dans le paragraphe suivant, en commençant avec l'index de synthèse.

Au point de vue dimension synthético-analytique, il y a clairement des langues qui au moins s'approchent la fin du spectre analytique, c.-à-d. où il y a la plupart des correspondances un-pour-un entre les mots et les morphèmes ; le vietnamien, cité plus haut, est un bon exemple. Cependant, il n'existe probablement pas de langue où il est obligatoire de combiner autant de morphèmes que possible en un simple mot, c.-à-d. où il devrait y avoir une parfaite symétrie entre le mot et la phrase. Bien qu'en esquimau, par exemple, il est facile de trouver des phrases consistant en juste un simple mot qui consiste à son tour en un grand nombre de morphèmes, il est juste aussi facile de trouver des phrases consistant en plus qu'un mot, et il y a beaucoup d'exemples où il n'est pas possible de combiner des morphèmes ensemble en un simple mot : comme noté au-dessus, l'esquimau n'a pas de moyen de combiner des morphèmes lexicaux ensemble, notamment si un morphème lexical donné n'a pas sémantiquement pour équivalent un morphème grammatical, alors il n'y a pas de méthode pour que ce morphème lexical puisse être combiné avec d'autres morphèmes lexicaux en un simple mot. Même dans le chukchi, qui a la possibilité de combiner des morphèmes lexicaux ensemble, il y a de sévères contraintes qui limitent les possibilités d'en faire ainsi : par exemple, il n'y a pas de moyen d'incorporer un sujet transitif ou la plupart des sujets intransitifs dans le verbe, de sorte qu'un exemple de trois mots comme *tumγ kupren nantəvatən* « les amis fixent le filet » (littéralement « amis filet fixent ») puisse être réduit à deux en incorporant l'objet direct, donnant *tumγət koprantəvatγ?at*, il n'y a pas de moyen pour que le sujet « amis » puisse à son tour être incorporé pour donner une phrase d'un mot contenant les trois morphèmes lexicaux. L'index de synthèse est donc mieux vu comme un index de degrés de déviation du type analytique idéal dans la direction de la synthèse.

Mais même en essayant d'appliquer l'index de synthèse en termes pratiques, par exemple en divisant le nombre de morphèmes par le nombre de mots, certains problèmes pratiques se posent qui indiquent qu'encore davantage d'attention doit être apportée aux bases théoriques de la typologie morphologique. Peut-être le plus sérieux, et le plus largement débattu dans la littérature, est la question d'établir des liens de mot, et de là le nombre de mots dans une phrase : même en vietnamien, nous notons ce problème avec l'expression (un ou deux mots ?) *bat dau* « commence ». Pendant que la définition canonique du mot comme une « forme libre minimale » donne beaucoup de kilométrage, et est particulièrement utile pour traiter avec les langues riches en polysynthèse, où les morphèmes individuels ne sont fréquemment clairement par des formes libres minimales, les problèmes peuvent se poser dans beaucoup plus de cas banals, p. ex. avec l'article

défini anglais dans *the man*, ou le pronom français clitique dans *je le vois*, où, malgré les conventions orthographiques, il y a peu de raisons pour assurer que *the*, *je* ou *le* sont des formes libres, c.-à-d. prononçables isolément (autrement que par les linguistes). Mais que *je le vois* soit considéré comme un ou trois mot peut faire une différence significative dans l'index de synthèse du français.

Un autre problème pour l'index de synthèse se révèle lorsqu'on essaie de compter les morphèmes, dans des langues contenant des morphes zéro ou des morphes porte-manteau. En anglais, le pluriel *cat-s* est clairement composé de deux morphèmes, mais le nombre de morphèmes dans le singulier *cat* est moins clair : seulement un morphème, ou un morphème lexical *cat* plus un morphème zéro grammatical ? En terme de comparaison entre les langues, décider entre l'un ou l'autre peut être d'une extrême importance pour des statistiques complexes : si l'anglais *cat* consiste en un seul morphème, alors l'anglais devrait être rapproché, en degrés de synthèse, au russe, où le singulier *košk-a* a un affixe à l'instar du pluriel *košk-i*. En analysant un verbe espagnol comme *cantas* 'tu chantes', devrait-on le considérer comme étant deux morphèmes (racine *cant-* ou *canta-* et affixe – *s* ou –*as*), ou plutôt que les morphèmes soient tous fusionnés à la fin (deuxième personne, singulier, temps présent, mode indicatif, premier groupe) donnant, avec le morphème lexical, rien de moins que six morphèmes ? Alors qu'une décision cohérente devrait être prise, au moins arbitrairement, la décision précise modifierait radicalement la comparaison entre une langue comme l'espagnol, avec une généralisation des occurrences de morphes porte-manteau (spécialement dans le système verbal), et une langue agglutinante comme le turc, où il y a peu ou pas de controverse à propos du nombre de morphèmes dans un mot (excepté, peut-être, pour le dénombrement de morphes zéro).

En nous tournant maintenant vers l'index de fusion, nous devrions rappeler les deux composants de l'agglutination qui étaient mentionnés au-dessus : segmentabilité et invariabilité des morphèmes. La segmentabilité est peut-être plus importante en prévision des traitements de l'agglutination, bien que le second ne doive pas non plus être exclu de la prise en compte, spécialement en comparant l'agglutination avec la fusion – qui est, après tout, ce que l'index de fusion fait. Ici, nous pouvons prendre l'agglutination comme la norme : des morphèmes clairement segmentables et invariables, et définir l'index de fusion comme déviation de la norme. L'extrême éloignement de cette norme devrait donc être la supplétion, où il n'y a absolument pas de segmentabilité et d'invariabilité, comme avec l'anglais *went* qui est le temps passé de *go*. Donc une langue qui représenterait le type fusionnel idéal devrait avoir tout dans sa morphologie se définissant en terme de supplétion ; si idéalement elle a aussi un index de synthèse élevé, alors chaque mot devrait simplement être totalement et non-segmentablement distinct de chaque autre mot de la langue. Cette langue donnée consiste en un nombre infini de mots, ce qui est clairement pratiquement impossible. Ce qui signifie en pratique que plus l'index de synthèse augmente, plus le ratio agglutination-fusion doit aussi augmenter ; Nous allons illustrer cela, ainsi que certains autres problèmes, dans les paragraphes qui suivent, en commençant par l'index de synthèse.

Nous pourrions maintenant les problèmes internes à l'index de fusion, en commençant par la segmentabilité et en nous tournant vers l'invariance. Le problème avec la segmentabilité et qu'elle n'est pas une catégorie à part entière. Elle implique plutôt un degrés de segmentabilité. Dans la déclinaison du turc donnée plus haut, la

segmentation était clairement identifiable dans chaque instance. Si nous regardons la déclinaison du hongrois, cependant, la situation n'est pas aussi simple, comme on peut le voir dans les formes suivantes, singulier et pluriel, nominatif et accusatif, de *ház* « maison », *asztal* « table », et *folyó* « rivière ». Il est donc également clair que la forme du pluriel à l'accusatif est constitué de trois morphèmes. Ce qui n'est pas clair, cependant, c'est l'endroit exact où le lien entre morphèmes doit être établi dans ces instances où la consonne de l'accusatif ou du pluriel est précédée par une voyelle qui n'est pas là dans la forme du nominatif ou du singulier, p. ex. est-ce que *házat* doit être segmenté en *ház-ar* ou en *háza-t*, est-ce que *asztalokat* doit être segmenté en *asztal-ok-at* ou en *asztalo-ka-t*, ou de façon plus concevable en *asztalo-k-at* ?

[Page 46]

Nominatif singulier	<i>ház</i>	<i>asztal</i>	<i>folyó</i>
Accusatif singulier	<i>házat</i>	<i>asztalt</i>	<i>folyót</i>
Nominatif pluriel	<i>házak</i>	<i>asztalok</i>	<i>folyók</i>
Accusatif pluriel	<i>házakat</i>	<i>asztalokat</i>	<i>folyókat</i>

Il est clair qu'il y a des morphèmes lexicaux qui consistent au moins en *ház*, *asztal*, et *folyó*, qu'il y a un suffixe accusatif consistant au moins en *-t*, et un suffixe pluriel consistant au moins en *-k*. Il est donc également clair que la forme de l'accusatif pluriel est constituée de trois morphèmes. Ce qui n'est pas clair est l'endroit exact où la limite devrait être établie dans ces exemples où la consonne de l'accusatif ou du pluriel est précédée par une voyelle qui n'est pas présente dans la forme nominative ou singulière correspondante, p. ex. est-ce que *házat* doit être segmenté *ház-at* ou *háza-t*, est-ce que *asztalokat* doit être segmenté *asztal-ok-at* ou *asztalo-ka-t*, voire *asztalo-k-at* ? Des raisons peuvent être avancées pour les deux possibilités logiques, c.-à-d. que les deux incluent la voyelle comme partie de la base ou non ; à ce moment le problème de la segmentation semble irrésoluble. Ce cas diffère donc du cas du turc, où la segmentation était simple. Cependant, il diffère au moins autant du cas du russe illustré ci-dessus à partir du moment où en hongrois il est clair que chacun des morphèmes impliqués a un contenu segmental (p. ex. accusatif *-t*, pluriel *-k*), tandis qu'en russe il n'y a pas de solution pour que des segments de suffixes puissent être identifiés comme indiquant exclusivement le cas ou le nombre. En quelque sorte, nous voulons indiquer un degré intermédiaire entre la segmentabilité qui va de soit et la segmentation impossible.

Si l'invariabilité est incluse comme une caractéristique additionnelle de la morphologie agglutinante, alors le problème des cas intermédiaires est même pire. D'abord, nous devrions illustrer que la segmentabilité et l'invariabilité sont en effet distincts l'un de l'autre. En turc, les morphèmes sont en général facilement segmentables et invariables, mais il y a des exceptions : en particulier, le suffixe de la première personne du pluriel des verbes est facilement segmentable, mais il a deux formes différentes du radical, *-iz* et *-k*, qui se surviennent dans différentes formes de temps, d'aspect et de mode, cf. aoriste *yap-ar-iz* « nous faisons », conditionnel *yap-sa-k* « si nous faisons ». Bien que le manque d'invariabilité soit évident, il n'y a pas de problème de segmentabilité, c.-à-d. que la situation illustrée à l'instant est plus agglutinante que la

déclinaison russe (où il n'y a ni segmentabilité ni invariabilité), mais moins agglutinante que la flexion du non turc (où il y a segmentabilité et invariabilité).

Dans certains exemples, la variabilité de la forme du morphème est complètement prévisible en termes de règles phonologiques générales de la langue en question. En turc, par exemple, les règles de l'harmonie vocalique comptent pour des formes différentes du morphème pluriel dans *adam-lar* « hommes » contre *ev-ler* « maisons » (-*lar* après une voyelle d'arrière, -*ler* après une voyelle d'avant, à partir du moment où l'harmonie vocalique exclut la présence simultanée de voyelles d'avant et d'arrière dans le même mot). Ces exemples, probablement, ne devraient pas être considérés comme des violations de l'invariabilité du moment que la variabilité du morphème est inévitable car découlant d'autres règles de la langue. D'autre part, pourtant, la variabilité dans la forme du morphème représente un continuum atteignant ses extrêmes avec la supplétion, mais traversant une rangée de valeurs intermédiaire en termes de degré de variation et de degré de prévisibilité : donc les alternances de la voyelle accentuée en anglais *divine-divinity* et *strong-strength* sont comparables en termes de distance phonétique entre alternants, mais tandis que la première alternance est essentiellement prévisible en termes morphologiques, la seconde est idiosyncratique.

Une bonne illustration du problème qui se pose en pratique quand on essaie de calculer l'index de fusion peut être vu en comparant la déclinaison du nom en finnois et en estonien, deux langues très proches génétiquement. D'abord, considérons les formes finnoises, de *jalka* « jambe » et *lippu* « drapeau » :

[Page 47]

Nominatif singulier	<i>jalka</i>	<i>lippu</i>
Génitif singulier	<i>jala-n</i>	<i>lipu-n</i>
Partitif singulier	<i>jalka-a</i>	<i>lippu-a</i>
Partitif pluriel	<i>jalko-j-a</i>	<i>lippu-j-a</i>

Avec ces données limitées ci-dessus, la segmentabilité n'est pas un problème : le suffixe du pluriel est *-j*, le suffixe du génitif est *-n*, et le suffixe du partitif est *-a* ; pareillement, il y a peu de variabilité dans la forme du morphème : les alternances consonnantiques *k - Ø* et *pp - p* sont largement (mais pas complètement) justifiables en termes de structure syllabique (le second membre de chaque paire survient dans une syllabe fermée), et l'apparance de *o* dans *jalkoja* est justifiable morphologiquement (mais pas phonétiquement). La situation est très différente en estonien, cependant :

[Page 48]

Nominatif singulier	<i>jalg</i>	<i>lipp</i>
Génitif singulier	<i>jala</i>	<i>lipu</i>
Partitif singulier	<i>jalga</i>	<i>lippu</i>
Partitif pluriel	<i>jalgu</i>	<i>lippe</i>

Bien que ces formes puissent toutes être dérivées diachroniquement de proto-formes proches des formes finnoises actuelles données ci-dessus, il n'y a plus de facilité

de segmentation ni d'invariabilité. Étymologiquement, toutes les formes sont différents allomorphes de la base, c.-à-d. que chacune est non-segmentable et l'alternance parmi les quatre formes de chaque mot est complètement imprévisible en termes phonétiques. Si on adopte l'analyse alternative de segmentation hors des voyelles finales pour le cas où le nombre des suffixes, alors le degré de variation dans la base est réduit, mais la variation est introduite dans les suffixes, p. ex. le « partitif singulier » est *-a* après *jalg*, mais *-u* après *lipp-*. Cet exemple ne sert pas seulement à illustrer le problème associé à l'assignation d'un index de fusion à un système morphologique comme celui de l'estonien, où il y a de faibles traces plutôt que des indications claires de segmentation, mais aussi le point plus général qu'un relatif court laps de temps peut suffire à altérer la typologie morphologique du langage d'une structure agglutinante assez nette à une autre qui est beaucoup plus fortement caractérisée par la fusion.

Nous pouvons résumer cela plutôt que discuter longuement de la typologie morphologique en disant qu'il y a deux indices majeurs, indépendants l'un de l'autre, qui sont nécessaires en typologie morphologique : l'index de synthèse, qui mesure le nombre de morphèmes par mot (bas dans les langues isolantes, haut dans les langues polysynthétiques), et l'index de fusion (mesurant la différence entre agglutination et fusion). Il y a de nombreux problèmes en pratique pour quantifier ces indices ; en particulier, l'index de fusion réfère à son tour à deux paramètres logiques indépendants, segmentabilité et invariabilité des morphèmes. Malgré la longue histoire des études en typologie morphologique, il est clair que beaucoup de problèmes tout à fait basiques de définition n'ont pas encore réellement été affrontés, ce pourquoi d'immenses problèmes pratiques se posent dès qu'on essaie de le faire, plutôt que juste en parler, en typologie morphologique.

Bien que la typologie morphologique serve le but utile de présenter un aperçu du type de structure morphologique de la langue, il demeure peu clair si elle peut être considérée comme un paramètre morphologique significatif (ou ensemble de paramètres) dans le sens d'une corrélation avec d'autres paramètres en dehors de la morphologie. Bien sûr, par définition, il y a peu de paramètres avec lesquels les types morphologiques se lient. Dans le chapitre 8, par exemple, un des types de la construction causative avec laquelle nous serons confrontés est la causation morphologique, par laquelle un causatif est relié à son équivalent morphologiquement non-causatif, chacun étant un simple mot, p. ex. turc *öl-diir* « provoquer la mort, tuer », en relation à *öl* « mourir ». Clairement donc, une construction causative peut seulement exister dans une langue qui n'est pas isolante, mais cela suit logiquement la définition d'un type isolant comme n'ayant pas de morphologie, et ne représente pas une corrélation parmi les paramètres indépendants logiquement. Notre conclusion d'ensemble est donc que la typologie morphologique a une place évidente mais restreinte dans la typologie de la langue, et il faut espérer que les manuels de linguistique générale ne vont pas continuer indéfiniment à donner l'impression que c'est la seule, ou la plus pertinente, solution de classification des langues typologiquement.

NOTES ET RÉFÉRENCES

Pour un survol général des approches sur la typologie des langues, incluant des références sur des études historiques plus détaillées, voir Greenberg (1974). Ma propre pensée sur la relation entre les universaux et la typologie a été influencée et clarifiée par Keenan (1978).

Le principe universel selon lequel le sujet précède habituellement l'objet dans l'ordre de base des mots est numéro 1 chez Greenberg (1966b, 110). Le matériel préliminaire traitant des langues où l'objet est en position finale peut-être trouvé chez Derbyshire & Pullum (1981).

L'information sur les termes désignant la couleur en Hanunoo est de Conklin (1955). La publication initiale sur les universaux des couleurs foci est de Berlin & Kay (1969). Ce travail a été sujet à de considérables affinements et critiques, cependant les résultats semblent se révéler être au moins comme des tendances universelles. L'explication en termes de perception se trouve chez Kay & McDaniel (1978), qui inclut aussi des références plus récentes sur la terminologie des couleurs et sur les théories perceptuelles et linguistiques utilisant des prototypes et des ensembles flous.

Les universaux des règles de tons, avec comme zone explicite privilégiée l'Afrique de l'ouest, sont discutés par Hyman & Schuh (1974). Dixon (1977, 110-12) discute de la pertinence répandue de l'animé en Yidiny.

Le principal travail classique sur la typologie morphologique, réalisé plus tôt par les frères Schlegel, a été fait par Humboldt (1836) ; C'est Humboldt qui a établi la quadruple typologie en incluant le polysynthétisme. L'exemple du Chukchi est de Skorik (1961, 102), et l'exemple de l'esquimo yupik sibérien est de Jacobson (1977, 2-3).

L'approche de la typologie morphologique adoptée ici tient beaucoup à Sapir (1921), chapitre 6 ; en particulier, Sapir introduit les paramètres de synthèse et technique (ce qui deviendra approximativement l'index de fusion). La quantification d'indices de synthèse et de fusion est introduite par Greenberg (1960). Différentes possibilités pour mesurer les indices de synthèse et de fusion sont discutées par Altmann & Lehfeldt (1973, 108-12) ; malgré le titre (« general language-typology »), ce livre traite presque exclusivement de typologie quantitative. Un compte-rendu plus détaillé de l'agglutination face à la fusion en finnois et en estonien peut être trouvé dans Comrie (à venir, d).

Il existe plusieurs travaux qui donnent un aperçu de la structure typologique d'une langue ou plus – en effet, on peut prétendre que ce matériel devrait être présent dans de bonnes grammaires. Meillet & Cohen (1952) fournit un aperçu de la plupart des familles de langues débattues (une troisième édition substantiellement révisée est en préparation). Des esquisses de plusieurs langues (Île de Pâques, japonais et chinois mandarin) sont incluses chez Lehmann (1978b). Bref, plus d'esquisses d'introduction à un large nombre de langues sont incluses chez Shopen (1979a, b). Voir aussi les références au travail de Sandra A. Thompson et John A. Hawkins dans les notes et références du chapitre 3.

Des aperçus typologiques détaillés d'une gamme de sujets syntaxiques sont inclus chez Shopen (à venir) : ces volumes pourraient bien servir comme ensemble de lectures supplémentaires en conjonction avec le présent livre.