

Conventions utilisées dans les gloses

Lorsqu'un linguiste rédige un livre ou un article, il peut être amené à donner des exemples dans des langues que le lecteur ne connaît pas. Pour lui permettre de comprendre ces exemples, l'auteur fournit ce qu'on appelle une *glose* ainsi qu'une traduction. Par exemple, la phrase anglaise *Allison planted daffodils* « Allison plantait des jonquilles » peut être représentée de la manière suivante :

- (1) *Allison plant-ed daffodil-s*
 Allison planter-PAS jonquille-PL
 « Allison plantait des jonquilles. »

La langue en cours de discussion est appelée « langue objet ». En (1), la langue objet est l'anglais. Comme vous pouvez le voir, la présentation d'un exemple fait intervenir trois lignes : l'exemple lui-même dans la langue objet, la glose et la traduction, toujours dans cet ordre. Souvent, l'exemple est en italique, contrairement aux deux autres lignes.

Des linguistes travaillant à Leipzig, en Allemagne, ont rédigé un document contenant les règles, ou conventions, qu'il est recommandé de suivre dans ce type de présentation. Ces règles, connues sous le nom de *Leipzig Glossing Rules*, sont accessibles en anglais à [cette adresse](https://www.eva.mpg.de/lingua/resources/glossing-rules.php)¹. La suite de ce document en présente les principales.

1 - Alignement

Les mots de la langue objet sont **alignés à gauche avec leur glose**. Comparez les exemples (1), (2) et (3). Vous pouvez voir qu'en (2), l'alignement est à droite, et qu'en (3), il n'y a pas d'alignement du tout. Ces types de mise en page sont à éviter.

- (2) *Allison plant-ed daffodil-s*
 Allison planter-PAS jonquille-PL
 « Allison plantait des jonquilles. »

- (3) *Allison plant-ed daffodil-s*
 Allison planter-PAS jonquille-PL
 « Allison plantait des jonquilles. »

2 - Tirets

Les préfixes et les suffixes de la langue objet sont séparés du reste du mot par des tirets. **On retrouve le même nombre de tirets dans la ligne de la langue objet et dans la glose**. Dans l'exemple turc en (4), vous voyez que le nombre de tirets entre les deux lignes ne correspond pas. Par conséquent, on ne sait pas quel morphème correspond à quelle glose. La glose correcte, illustrée en (5), fait

¹ <https://www.eva.mpg.de/lingua/resources/glossing-rules.php>

immédiatement apparaître la valeur de chaque morphème : *gel-* « venir », *-me* « négation », *-yeceğ* « futur » et *-im* « 1^{ère} personne du singulier ».

(4) *gel-me-yeceğim*
venir-NEG-FUT-1SG
« je ne viendrai pas »

(5) *gel-me-yeceğ-im*
venir-NEG-FUT-1SG
« je ne viendrai pas »

3 - Majuscules et minuscules, abréviations

Dans la glose, les éléments grammaticaux (« singulier », « pluriel », « masculin », « féminin », « futur », etc.) sont **abrévés et en petites majuscules** ; les éléments lexicaux (« planter », « jonquille », « rouge », « lentement », etc.) sont **en minuscules et non abrévés**. Les *Leipzig Glossing Rules* proposent une liste d'abréviations pour les éléments grammaticaux. D'autres abréviations ont été ajoutées à cette liste initiale par les membres du projet CorpAfroAs (voir [ici](http://corpafras.huma-num.fr/glosses.html)²).

Dans le logiciel Word, pour écrire un mot en petites majuscules, commencez par écrire ce mot en lettres minuscules. Placez ensuite le curseur sur le mot, puis allez dans « Police » et cochez la case « Petites majuscules ».

4 - Point

Dans la glose d'un mot, les éléments ne sont jamais séparés par des espaces. Ainsi, l'exemple (6) est mal glosé.

(6) *potato*
pomme de terre
« pomme de terre »

Pour éviter les espaces, on relie les différents éléments par un point :

(7) *potato*
pomme.de.terre
« pomme de terre »

Cette règle a une exception : l'indication de la personne est directement accolée à l'indication de nombre qui suit (singulier, duel, pluriel, etc.). Ainsi, dans l'exemple (5), le suffixe *-im* est glosé 1SG et non 1.SG. La personne est glosée avec des chiffres arabes (1, 2 et 3).

² <http://corpafras.huma-num.fr/glosses.html>

5 - Crochets carrés

Pour comprendre la règle suivante, prenons un exemple en anglais. Dans cette langue, le pluriel de la grande majorité des noms se forme en rajoutant un suffixe à la forme de singulier : *daffodil* « jonquille », *daffodil-s* « jonquilles ». Il y a donc une marque explicite au pluriel, mais aucune marque au singulier. C'est l'absence de marque qui donne l'information « singulier ». Dans la glose, on utilise les **crochets carrés** pour indiquer **un élément grammatical auquel ne correspond aucun affixe**, ce que vous pouvez voir en comparant (8) et (9).

- | | | | |
|-----|---|-----|---|
| (8) | <i>daffodil</i>
jonquille[sg]
« jonquille » | (9) | <i>daffodil-s</i>
jonquille-PL
« jonquilles » |
|-----|---|-----|---|

Une autre possibilité est de considérer que le singulier s'exprime par un « morphème zéro », représenté par le symbole \emptyset . Dans ce cas, on n'utilise pas les crochets, mais le tiret habituel :

- (10) *daffodil- \emptyset*
jonquille-sg
« jonquille »

6 - Parenthèses

Pour comprendre l'usage des parenthèses, prenons un exemple en russe. Dans cette langue, les noms sont répartis en trois genres : masculin, féminin et neutre. Le genre d'un nom est fixe : on ne peut pas décider de mettre au masculin le nom *kvartira* « appartement », qui est féminin en russe. Autrement dit, le genre d'un nom ne fait pas l'objet d'un choix ; il est « inhérent » au nom. Les parenthèses s'utilisent pour indiquer les informations inhérentes qui ne sont pas exprimées par une marque particulière. La glose du nom *kvartira* précédé d'un adjectif est représentée en (11).

- (11) *ujutn-aja* *kvartir-a*
douillet-sgF appartement(f)-sg
« un appartement douillet »

Vous constatez que l'information « féminin » dans la glose de l'adjectif *ujutnaja* n'est pas entre parenthèses. En effet, les adjectifs, en russe comme en français, n'ont pas de genre inhérent : un adjectif peut apparaître au féminin, au masculin ou au neutre, en fonction du nom avec lequel il se construit.

7 - Chevrons

Ce qu'on appelle « affixe » regroupe plusieurs types d'éléments morphologiques : les *préfixes*, les *suffixes* et les *infixes*. Comme nous l'avons vu, préfixes et suffixes sont séparés par des tirets (règle 2). Pour les infixes, une convention différente est utilisée. Prenons un exemple en hua, langue parlée en Papouasie-Nouvelle-Guinée. Dans cette langue, on peut créer des formes verbales négatives en ajoutant à l'intérieur de la base l'élément *ʔa* (Haiman 1980:195) :

forme positive	forme négative
<i>harúpo</i> « glisser »	<i>haruʔapo</i> « ne pas glisser »
<i>zgavo</i> « êtreindre »	<i>zgaʔavo</i> « ne pas êtreindre »
<i>rvato</i> « être proche »	<i>rvaʔato</i> « ne pas être proche »
<i>rapai</i> « être gros »	<i>raʔapai</i> « ne pas être gros »

Séparer les infixes par des tirets pose deux problèmes. Si nous segmentons par exemple *zgaʔavo* en *zga-ʔa-vo*, nous donnons l'impression que ce mot est composé de trois morphèmes, alors qu'il n'en contient que deux : la base *zgavo* et l'infixe de négation *ʔa*. D'autre part, comment gloser les trois parties de *zga-ʔa-vo* ? êtreindre-NEG-êtreindre ? La solution retenue par les auteurs des *Leipzig Glossing Rules* consiste à **utiliser les chevrons pour séparer l'infixe dans la ligne de la langue objet et dans la glose**, comme vous pouvez le voir en (12).

- (12) *zga<ʔa>vo*
 êtreindre<NEG>
 « ne pas êtreindre »

Remarque

Quand vous rédigez un document, n'oubliez pas d'**indiquer quelque part dans ce document la liste des abréviations que vous utilisez**, par ordre alphabétique. Il n'y a pour l'instant aucune liste d'abréviations communément acceptées, ni en anglais ni en français. Voici la liste des abréviations utilisées dans ce document :

F	féminin
FUT	futur
NEG	négation
PAS	passé
PL	pluriel
SG	singulier

Référence

Haiman, John. 1980. *Hua: A Papuan Language of the Eastern Highlands of New Guinea*. John Benjamins.