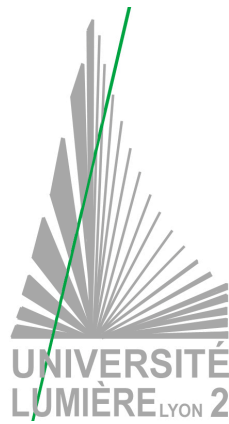


Rapport de stage  
DDL-CNRS / LYON II

Les pratiques de gloses en français, espagnol et anglais.



mai 2010

Jérémy Pasquereau  
M1 SDL Lyon 2

Ce rapport est le fruit de mon stage de recherche au laboratoire CNRS / Lyon 2 Dynamique du Langage effectué pendant le premier semestre 2010. Il poursuit le travail commencé par Jason Huang en 2009 et c'est librement que je réutilise certaines de ses analyses.

Dans le but de rendre compte des différentes pratiques, j'ai utilisé un certain nombre d'exemples provenant d'ouvrages d'une grande variété de linguistes. Je tiens à insister sur le fait qu'aucun linguiste n'est visé quant à ses pratiques de gloses et que les noms apparaissent dans le seul souci de prouver l'authenticité des exemples.

Je remercie aussi les personnes qui m'ont aidé dans ce travail par leurs suggestions et leur participation aux ateliers de réflexion : Colette Grinevald, Denis Creissels, Marine Vuillermet, Natalia Cáceres, Anne-Laure Dotte et particulièrement Lucie Amaro qui a participé à la rédaction des comptes-rendus d'ateliers et a relu ce rapport.

<b>1 INTRODUCTION</b> .....	<b>4</b>
<b>2 ESSAIS DE STANDARDISATION</b> .....	<b>4</b>
2.1 RÉSUMÉ DE « INTERLINEAR MORPHEMIC GLOSSES » DE LEHMANN .....	4
2.2 RÉSUMÉ DE « THE LEIPZIG GLOSSING RULES ».....	6
2.3 MISE EN ÉVIDENCE DES « CONFLITS » ENTRE LEHMANN ET LEIPZIG.....	8
<b>3 EXPOSES DE PRATIQUES DIVERSES</b> .....	<b>9</b>
3.1 ÉVOLUTION DES GLOSES.....	9
3.1.1 <i>Dans les publications en espagnol</i> .....	9
3.1.2 <i>Dans les publications en français</i> .....	11
3.1.3 <i>Dans les publications en anglais</i> .....	14
3.2 RENCONTRE AVEC DES LINGUISTES CHERCHEURS AU LABORATOIRE DDL (LYON) .....	19
3.3 GLOSE ET TRADITION.....	21
<b>4. PROPOS DE NOTRE CONTRIBUTION</b> .....	<b>22</b>
4.1 COMPARAISON DE LISTES D'ABRÉVIATIONS .....	22
4.2 ÉLÉMENTS DE PRÉSENTATION .....	23
4.3 PRINCIPES POUR LA CRÉATION D'ABRÉVIATIONS.....	29
4.4 LISTE D'ABRÉVIATIONS .....	32
<b>5. CONCLUSION</b> .....	<b>36</b>
<b>OUVRAGES CONSULTÉS</b> .....	<b>37</b>
<b>ANNEXE</b> .....	<b>39</b>

# 1 Introduction



En tant que centre de description de langues peu ou pas décrites, le laboratoire Dynamique du Langage suit avec intérêt les divers questionnements sur la question des gloses et n'hésite pas à s'adapter et à abandonner certaines pratiques « traditionnelles » pour d'autres plus standards quand celles-ci servent la compréhension inter-linguistes. Nous précisons *certaines* propositions car elles ne nous paraissent pas toutes égales quand il s'agit d'aider cette intercompréhension.

La glose constitue un niveau intermédiaire de compréhension entre celui de la langue objet et celui de la traduction. Le but de la glose interlinéaire est d'informer sur la signification et les propriétés grammaticales des différents lexèmes et morphèmes (Leipzig : 2008, p. 1), de refléter les diverses manières dont chaque langue exprime le monde à l'entour, ou comme Lehmann le formule « l'esprit de la langue » (2004, p. 2).

Nous réservons le terme *glose* à la morphologie et parlerons de *codage* si nécessaire pour tout ce qui relève d'autres niveaux d'analyse (constituance, syntaxe, ...). Pour qui la glose est-elle utile ? De manière générale, la glose est utile à tout lecteur non familier avec la langue étudiée. Pour les linguistes, elle tend à être une métalangue, avec une visée typologique qui va au-delà des termes idiosyncrasiques des grammaires traditionnelles. Elle permet au linguiste de mettre la structure d'une langue particulière en perspective même s'il n'a pas de connaissances préalables de cette langue. Elle est aujourd'hui incontournable, particulièrement dans les études typologiques où le nombre de langues pouvant être mis en comparaison par un seul linguiste est très important.

Etant donnée la diversité des langues et le flot d'informations qu'un seul énoncé porte en lui à des niveaux différents (le sens d'un mot, la partie du discours, la fonction pragmatique ou discursive), il est matériellement impossible d'en rendre compte exhaustivement dans la glose. Le propos d'une ligne de glose est de rendre compte de l'analyse linguistique effectuée préalablement et de la faciliter tout en restant lisible.

Nous voudrions insister également sur l'étroite corrélation entre le développement de la glose et celui de l'informatique. Une partie de l'intérêt des gloses est de permettre de faire des recherches par glose et d'étudier tous les cas d'un certain affixe dans un corpus par exemple (Grinevald: c.p.). Il existe aujourd'hui des logiciels spécialisés qui permettent de gagner un temps considérable dans l'étude d'un morphème spécifique.

Plus particulièrement, dans le cadre des recherches menées au laboratoire Dynamique Du Langage (DDL), les chercheurs en description des langues publient en français, anglais et espagnol. Le but principal de notre travail est de donner un aperçu des pratiques des linguistes en matière de glose afin de nourrir la réflexion vers une uniformisation plus avancée des gloses. Visant aussi la mise en place d'une liste d'abréviations grammaticales qui soit autant que possible commune pour ces trois langues, nous proposerons une liste de principes généraux non restreints à ces langues.

## 2 Essais de standardisation



### 2.1 Résumé de « Interlinear Morphemic Glosses » de Lehmann

La proposition de standardisation faite par Lehmann remonte à 1982 et a par la

suite été régulièrement mise à jour. La dernière en date est de 2004. Pour Lehmann le but premier d'une glose morphémique interlinéaire (GMI) est de mettre le lecteur dans une position qui le rende capable de comprendre la structure grammaticale de la langue objet. Elle est en outre un instrument de "contrôle" des assertions de l'auteur et permet de passer en revue un corpus en sélectionnant toutes les occurrences d'un phénomène.

Lehmann précise que la GMI est "une sorte" d'analyse linguistique et il résume le problème dont nous nous sommes saisi dans cet article en une phrase "En ce sens, [la GMI] concurrence un pan de la grammaire. Son caractère hybride est la cause d'un certain nombre de problèmes et de différentes manières de glosage". En tant qu'analyse la glose peut donc changer d'un auteur à l'autre pour le même exemple. Quant à ce que la GMI doit contenir, l'auteur est très précis : la glose doit indiquer pour chaque morphème, non la catégorie grammaticale ou syntaxique à laquelle il appartient mais la valeur que ce morphème a dans le contexte. Autrement dit le linguiste ne doit pas chercher à rendre compte des caractéristiques qui unissent différents morphèmes en une même classe ou catégorie, cette dernière tâche va à la grammaire. Cette règle a en même temps des implications non seulement pour l'analyse linguistique à proprement parler mais aussi pour la présentation des gloses. Ainsi la glose pour le terme allemand *laufend* 'courant' n'est-elle pas *courir(part.prs)* mais bien *cour:PART.PRS*. Lehmann ne cite qu'une éventuelle entorse à cette règle pour le cas des verbes sur lesquelles sont indexés plus d'un argument, en effet dans ce cas il peut être utile d'indiquer la fonction syntaxique de chaque morphème.

Lehmann énonce déjà un certain nombre des principes qui font l'objet de notre discussion. Il énonce la possibilité d'avoir plusieurs niveaux de glose suivant le public auquel la glose est adressée: des niveaux "simplifiés" à partir d'un niveau d'analyse optimale, celui destiné aux linguistes. Il pose la contrainte de réserver une ligne de glose par niveau d'analyse et propose sept de ces niveaux, tout exemple devant comporter au moins trois niveaux (une représentation de la langue objet, une GMI et une traduction idiomatique).

Il propose également des solutions standards de traitement de phénomènes linguistiques tels que les morphèmes zéros et l'allomorphie qui pour l'un relève d'un choix personnel de théorie linguistique, pour l'autre soulève un problème qu'il laisse non résolu, celui de la glose du genre des suffixes casuels latins pour les adjectifs qui « contredit » la glose des suffixes casuels nominaux puisqu'il serait contradictoire de gloser un même morphème de deux façons différentes (le genre est inhérent au nom mais pas à l'adjectif).

La distinction entre base lexicale et valeur d'une catégorie grammaticale est rendue par deux typographies : une base lexicale est glosée en minuscules et une abréviation de valeur de catégorie grammaticale en petites majuscules. Les noms propres doivent être rendus de même par leurs équivalents s'ils existent dans la langue cible.

Etant donné que d'une langue à l'autre, il n'y a pas correspondance exacte entre les mots grammaticaux, Lehmann préconise l'emploi d'abréviations référant toujours à une valeur précise que peut prendre une catégorie grammaticale. En effet donner la catégorie grammaticale d'un morphème ne peut par elle-même rendre le « sens » du morphème.

Pour ce qui est de la polysémie, elle doit être traitée en tenant compte de l'ensemble des exemples : si un lexème n'apparaît que dans un sens alors il est inutile de gloser un autre sens, à l'inverse le terme sera rendu par son sens de base (*Gesambedeutung*) dans la glose. L'abréviation d'un morphème servant à marquer plusieurs valeurs grammaticales sera soit une description « englobante », soit une

étiquette décrivant chacun des emplois du morphème soit l'abréviation d'un nouveau terme forgé pour référer au morphème syncrétique. Enfin il aborde les sujets de la dérivation et de son rendu ou non dans la glose et des unités sub-morphémiques telles que les segments épenthétiques qu'il recommande de représenter par un Ø dans la GMI ou de rattacher à un autre morphème.

Après avoir passé en revue les problèmes de rendu de l'analyse linguistique dans la glose auxquels tout linguiste est un jour confronté, Lehmann propose une liste d'abréviations qu'il veut à la croisée de deux principes contradictoires : les abréviations doivent être distinctes et elles doivent être courtes. Ils donnent la priorité au premier principe.

Lehmann termine son exposé par des propositions de codages différentes pour les frontières entre morphèmes. La première règle est qu'il ne peut y avoir de blanc dans la première ligne que si et seulement si il y en a un dans la GMI. Aussi, si deux éléments de la langue objet correspondent à un seul élément dans la langue cible, il propose de lier les deux éléments de la langue objet par un tiret bas  $\_$ . Dans la majorité des cas non problématiques, la frontière entre morphèmes sera rendue par un trait d'union - répété dans la GMI. Il réserve l'emploi des deux-points : aux cas où plus d'un morphème sont présents mais il n'est pas possible de les séparer, c'est le cas des morphèmes portemanteaux. Le signe plus + peut être utilisé pour signaler une frontière entre deux composants d'un mot composé. Le signe égal = peut de même servir à matérialiser une frontière entre un clitique et la base à laquelle il s'attache mais ce signe spécifique n'est pas plus obligatoire que + et peut très bien être rendu par un simple point . . Enfin la valeur d'un morphème zéro sera indiquée entre parenthèses si le linguiste a pris la décision de ne pas faire figurer de marque Ø dans la première ligne. Pour ce qui est des infixes, Lehmann recommande de les isoler par des angles <>. Les mêmes crochets sont utilisés pour les circonfixes. Le tilde ~ au lieu du trait d'union - indique quant à lui la frontière entre une base et un morphème issu de sa reduplication. Enfin les phénomènes d'ordre morpho-phonologiques doivent être traités au cas par cas. L'utilisation de la barre oblique inversée \ est particulièrement utile pour rendre compte des changements de ton et de l'ablaut.

Il est important de noter que Lehmann déconseille l'utilisation du point . comme marque non spécifiée de frontière, cet emploi induisant en erreur.

Pour le cas particulier où une langue indexe sur le verbe et le sujet et l'objet à l'aide d'un seul morphème syncrétique, il recommande l'utilisation du signe & ou du signe >, le dernier ayant l'avantage de remplacer les étiquettes *sujet* et *objet*. Une dernière chose dont Lehmann fait état et que nous reprenons dans notre travail est la possibilité d'utiliser les crochets [ ] pour indiquer les constituants dans un énoncé.

## 2.2 Résumé de « The Leipzig Glossing Rules »

Le document *The Leipzig Glossing Rules* a été élaboré par Bernard Comrie et Martin Haspelmath du département de linguistique du Max Planck Institute ainsi que par Balthazard Bickel du département de linguistique de l'université de Leipzig. Le Max Planck Institute s'intéresse principalement aux universaux linguistiques et à la typologie. Ils sont amenés à étudier des phénomènes en mettant en parallèle un grand nombre de langues. Un aspect essentiel de leur mission est aussi de participer à la description des langues peu décrites, le plus souvent en danger. Balthazar Bickel est principalement intéressé par la typologie des langues et est par ce biais engagé dans plusieurs projets mettant en jeu des langues « minoritaires ». Il a en effet travaillé sur la famille bantu et sur le turc avant de porter son attention sur la région

de l'Himalaya. L'intérêt que représente un système de glose standardisé pour des typologues est évident et nous allons tâcher d'en rendre l'essentiel dans le résumé qui suit.

La glose selon *The Leipzig Glossing Rules* fournit « des informations sur les significations et les propriétés grammaticales de mots individuels et de parties de mots. » Il s'affirme dès le départ non comme un document proposant une standardisation mais plutôt comme rendant explicite des conventions déjà mises en oeuvre fréquemment. Bien qu'ils reconnaissent qu'une GMI puisse être plus ou moins précise selon le public visé, leur proposition énonce des principes applicables à la présentation d'exemples dans un article ou ouvrage scientifique. La glose n'est pas envisagée comme rendant compte d'une analyse complète mais vise seulement à refléter la structure d'un énoncé. De même la glose pour un même exemple peut changer d'un linguiste à l'autre parce qu'elle ne sera pas dans la même langue, parce qu'un autre linguiste ne mettra pas l'emphase sur le même morphème ou bien encore parce que deux linguistes n'auront pas la même analyse.

La première règle que Leipzig énonce est celle de l'alignement de la ligne de la langue objet mot à mot avec celle de la GMI à gauche.

La deuxième règle définit le trait d'union - comme marqueur de frontière entre deux morphèmes et le signe égal = comme marqueur de la frontière entre un clitique et la base. Les segments épenthétiques sont associés soit au morphème précédent soit au suivant. Le trait d'union précédé d'un espace dans la ligne de la langue objet (mais pas dans la GMI) peut, le cas échéant, être utilisé pour marquer la relation morphologique entre deux mots phonologiques.

Dans leur troisième règle, Comrie, Haspelmath et Bickel posent qu'une liste d'abréviations standards (en majuscules) n'est pas une fin en soi. En effet certains phénomènes sont plus récurrents dans certaines langues que dans d'autres et pour des raisons pratiques il est donc logique qu'un linguiste réserve les étiquettes les plus courtes aux phénomènes les plus courants. Qui plus est, selon le but de la GMI, le linguiste a le choix entre une abréviation ou bien l'équivalent du morphème dans la langue cible (par exemple "COM" ou bien "avec")

Le propos de la quatrième règle est de proposer une manière de résoudre les problèmes de présentation qui se posent quand à un terme en correspondent deux ou plus. Il s'agit en fait d'une règle « hiérarchisée ». Pour toutes les situations dans lesquelles à un élément de la langue objet correspond plus d'un élément de la langue cible (d'après l'orthographe), il convient de séparer ces éléments par un point. Bien sûr un linguiste peut avoir besoin d'être plus précis et de différencier les types de relations mots à mot et c'est à cela que la deuxième partie (optionnelle) de la règle s'emploie. Si les éléments sont des lexèmes, leur unité peut être marquée par le trait bas  $\_$ . Les différentes valeurs qu'encode un morphème porte-manteau seront elles séparées par un point-virgule ; . Si le linguiste ne désire pas segmenter un mot (qui pourrait par ailleurs l'être sans problème), les auteurs recommandent l'utilisation des deux-points : pour séparer les valeurs grammaticales. Si une valeur grammaticale est réalisée par un changement morpho-phonologique, donc non segmentable linéairement, le trait oblique inversé sert alors à séparer l'abréviation grammaticale du reste de la glose. Enfin si les éléments segmentables sont des indices verbaux d'agent et d'objet, le crochet anguleux > peut être utilisé, signifiant par là même laquelle abréviation marque l'agent et laquelle le patient.

La cinquième règle établit que les abréviations de personne et de nombre ne doivent pas être séparées par un quelconque caractère typographique mais directement apposées et que les abréviations de nombre et de genre peuvent même être en minuscules dans la GMI.

Pour ce qui est du traitement des morphèmes zéro la sixième règle propose deux choix qui dépendent de l'analyse du linguiste : soit faire figurer le signe  $\emptyset$  dans la première ligne, soit inscrire sa valeur entre crochets dans la deuxième.

Les parenthèses sont réservées au marquage des valeurs inhérentes à un lexème comme le genre.

Dans la huitième règle est abordé le problème des éléments sémantiquement unitaires mais morphologiquement bipartites (ex. circonfixes). La solution proposée est de soit répéter la glose pour chaque partie de l'élément en question, soit gloser un des éléments par une abréviation spéciale telle que 'STEM' par exemple.

La règle neuf préconise l'utilisation des crochets anguleux  $\langle \rangle$  pour isoler un infixe et, de manière parallèle, son abréviation dans la ligne de GMI.

Enfin la dixième règle illustre l'utilisation du tilde  $\sim$  comme marque de frontière entre un élément redoublé et sa base.

Une liste d'abréviations vient compléter et terminer cette liste de règles.

### 2.3 Mise en évidence des « conflits » entre Lehmann et Leipzig.

Les différences dans les recommandations concernent principalement le choix des marques de frontière entre valeurs grammaticales.

Type de frontière	Leipzig	Lehmann
un lexème en L1 pour deux lexèmes en L2	x_y	x.y
deux mots orthographiques en L1 sont une unité sémantique en L2	x -y $\Rightarrow$ z	x_y $\Rightarrow$ z
valeur inhérente d'un lexème	x(y)	x.y
frontière existante non montrée	x:y	x:y
frontière non existante (portemanteau)	x;y	x:y
valeur grammaticale marquée par un changement morpho-phonologique	x\y	x\y
dans un verbe pour séparer les affixes indexant l'agent et le patient	x > y	x > y ou x&y
pour marquer la valeur d'un morphème zéro	x[y]	x(y)
pour isoler un infixe et sa valeur dans la GMI	x < y > x	x < y > x
pour isoler un circonfixe et sa valeur dans la GMI	y-x-y $\Rightarrow$ Y-x-Y ou STEM-x-Y	y > x < y $\Rightarrow$ x < y >
pour marquer une frontière entre un morphème redoublé et sa base	x~x	x~x
pour marquer la frontière entre les composants d'un mot composé	x.y	x + y
pour marquer la frontière entre un clitique et sa base	x = y	x = y
marque non spécifiée séparant différentes valeurs grammaticales	x.y	

Tableau 1 . Comparaison des symboles de frontières entre morphèmes recommandés par Lehmann et Leipzig

Le document proposé par Leipzig prévoit deux niveaux de précision : à un premier niveau d'analyse l'utilisation généralisée du point . entre plusieurs valeurs



grammaticales sans implication aucune quant à la présence d'une frontière morphologique ou, à un niveau plus précis, l'utilisation de toute une série de marques typographiques différentes précisant le type de relation qu'entretient un morphème avec son environnement morphologique direct. Cette utilisation du point comme marqueur non spécifié est fortement découragée par Lehmann qui pour une part pense qu'elle peut induire en erreur et d'autre part qu'un linguiste n'en a pas besoin dans la mesure où s'il ne peut prendre une décision quant au statut d'un morphème (et donc choisir entre des marques plus spécifiques de frontières) sans doute ne devrait-il même pas utiliser l'exemple en question (Lehmann: 2004, p. 27).

Il est vrai que le point est utilisé dans plusieurs situations et notamment lorsqu'une décision d'analyse n'a pas encore été prise, le point est utilisé par défaut.

### 3 Exposés de pratiques diverses

□

#### 3.1 Evolution des gloses

□

##### 3.1.1 Dans les publications en espagnol

Les premières études descriptives n'avaient pas de gloses, au sens strictement morphologique du terme. Dans l'esquisse grammaticale de la langue moxeño datant du dix-septième siècle et réécrite en 1976 par le professeur R. B. Casanovas sous le nom *Diccionario del idioma moxeño a través del tiempo*, le peu d'exemples qui est donné n'est pas glosé. *La Gramática de la lengua Quiché* de De Bourgbourg (1961) ne donne que la traduction littérale sans aucune indication autre.

- 1 El superlativo se hace con este adverbio "hincha" que significa: mucho o, muy; v. g.: nucahuma hincha : estoy muy enfermo; tiurimoi hincha maca : este es hermosísimo. Cuando el superlativo es juntamente comparativo, como: la Santísima Virgen es la mejor de todas las mujeres, se dice así: sacaereichu Virgen Santa María tiuri hincha nayee eseno no; que quiere decir: sólo la Santa Virgen María es muy buena entre todas las mujeres; y también se dice así: sacaereichu esu eseno tiuri hincha; que quiere decir: sólo Ella es excelentísima Mujer.

Anonyme. 1675. *Diccionario del idioma moxeño a través del tiempo. Estudio comparativo sobre su evolución.* Esquisse reprise par Pr. R. B. Casanovas en 1976. p. 107

- 2 Qu'in a col-ta pu chi coh, at ahau! ¡librame de la boca del león, o Señor!

Apachina ri ch' akan ula chiri pa quechelah? ¿quién es este que baja aquí del desierto?

De Bourgbourg, Ch. E. B. 1961. *Gramática de la lengua quiché.* Guatemala. p. 157

En 1978, B. Pothier publie *America latina en sus lenguas indígenas* auquel plusieurs linguistes contribuent par des articles. J. Albó glose de manière assez détaillée ses exemples d'aymara, c'est-à-dire que tous les termes de la langue objet ont une glose. Cependant des informations relevant de niveaux différents sont réunies sur la même ligne (équivalents sémantiques, informations pragmatiques, ...) et les quelques abréviations grammaticales apparaissent dans une troisième ligne.

3	<i>qari</i>	<i>-ta</i>	<i>-raki.</i>	<i>Uk(a)</i>	<i>-jam(a)</i>	<i>uñj(a)</i>	<i>-sin(a)</i>	<i>-sti</i>
	cansarse		también	eso	como	ver		y
		N	I		así		G	I
	cansado		también.		Viéndole		así	
	<i>ma(y)a</i>	<i>jaqi</i>	<i>-wa</i>	<i>jak'(a)</i>	<i>-Ø</i>	<i>-kata</i>	<i>-tayna</i>	
	uno	persona	afirmativo	cerca		lado	había	
			O		V	M	3-3	
							conocim. indirecto	

una persona se le acercó

Albó, J. 1978. "El Aymara". In Pothier, B. (ed) *América latina en sus lenguas indígenas*. Caracas, Monte Avila: Unesco. p. 303

L'article issu du même recueil et écrit par S. M. Dessaint présente des exemples sur six lignes. La première ligne est celle de la langue objet, la deuxième est réservée à l'analyse en constituants, dans la troisième il indique la catégorie grammaticale des lexèmes par des abréviations (LN et LV) mettant en évidence les morphèmes grammaticaux, eux mêmes rendus par des abréviations en minuscules. La quatrième ligne redonne la langue objet mais avec les segmentations, segmentations qui sont reprises dans la cinquième ligne où chaque morphème est rendu par un morphème équivalent de l'espagnol bien que cette technique « défigure » les mots. La dernière ligne donne la traduction idiomatique.

4	<i>Kame ou Paraguaygui ha ojuhu isýpe</i>
	GN1//GVa//GN3//rl//GVa//GN3
	LN//mpa-LV//LN-rl//rl//mpa-LV//mpp-LN-rl
	Kame//o-u//Paraguay-gui//ha//o-juhu//i-sy-pe
	Carmen//"-e"-vien//Asunción-de//y//"-a"-encuentr-//de
	ella-madre-a

Carmen viene de Asunción y encuentra a su madre.

Dessaint, R. B. S. M. 1978. "Esbozo del sistema lingüístico del Guaraní paraguayo". In Pothier, B. (ed) *América latina en sus lenguas indígenas*. Caracas, Monte Avila: Unesco. p. 317

Dans l'exemple suivant, la traduction idiomatique est donnée juste après la ligne de la langue objet et le phénomène que l'exemple illustre est souligné tandis que la ligne de glose est en italique. Du côté des abréviations on remarque qu'elles ne donnent pas la valeur proprement dite mais la catégorie grammaticale de chaque morphème.

5	448)	Kaamiik	x-oq-war	kaan	<u>chi</u>	<u>ruu-chii'</u>	ja	<u>chooy.</u>
		ADV	COM-1Bp-dormir	DIR	PREP	3As-boca	ART	lago.

449) Ja nutata' chi nuuchii' xtz'ub'e' wi'.

*Mi papá se sentó a mi lado.*

Ja	nu-tata'	<u>chi</u>	<u>nuu-chii'</u>	x-Ø-tz'ub'e'	wi'.
ART	1As-papá	PREP	1As-boca	COM-3Bs-sentar	PART

Ixmata, P. G. 1997. *Rukeemiik ja tz'utujiil chii'*. Gramática Tz'utujiil. Editorial Cholsamaj. p. 259

Les langues de Bolivie ne sont pas toutes traitées de la même manière. Comme on le verra dans la partie suivante, des chercheurs usent de gloses très élaborées pour leurs exemples, néanmoins on peut encore trouver des ouvrages qui, bien que de très bonne qualité, présentent des listes exhaustives d'exemples pour chaque phénomène étudié et qui malheureusement ne glosent rien.

- 6 El clasificador depende del complemento directo en los verbos transitivos, y depende del sujeto en los reflexivos, intransitivos, estativos, etc. En los reflexivos por definición sujeto y complemento son el mismo.

*pivemeca* enjabónela (tiene que ser una superficie lisa)  
*vésesuraji tiápujicava* nuestra esponja para refregar *vésesuraji eta curuja tiápujicava* con lo que restregamos los platos es la esponja  
*tiápujicavapa eta nupandearruraya* ya hinchó mi masa para mi pandearroz (del que va a ser mi pan-de-arroz); *eta nuturtane vaiítiápujicaca* mi torta no levantó (no hincha) *vaiítúrihacanu* no me sale bien, no me cae bien

Zubiri, J. O. 2002. *Gramática moja ignaciana*. p. 276

Dans sa grammaire maya du sakapulteka (2007), R.M. Isém choisit d'indiquer les rôles syntaxiques en les soulignant et en inscrivant leurs initiales sur la ligne du dessous. Une troisième ligne reprend la première ligne en la segmentant. « L'adresse » de l'exemple est indiquée à la fin de la traduction littérale entre parenthèses.

- 7 55) Renejeel tyempo kecha' wunaq chan.  
           O                  V          S          OI  
 r-enejeel  tyempo k-e-cha'          wunaq ch-Ø-an  
 E3s-todo  tiempo  INC-A3p-decir  wunaq  PRE-E3s-SR  
*Todo el tiempo le dice la gente.* (Trtex01Sak03:108)

Isém, R. M. 2007. *Rikemiik li Tujaal Tzij. Gramática Sakapulteka*. Guatemala. p. 352

Dans une grammaire maya tektiteka publiée la même année, Vail se sert d'une deuxième ligne pour indiquer les rôles sémantiques.

- 8 d. nkyin aq'unan te Wa'nch  
                                   Benefactivo  
*Trabajo para Juan.*  
 /n-kyin  aq'una-n          t-e          Wa'nch/  
 INC-A1s  trabajar-AP          E3S-SR:para  Juan  
 e. ma tzaj weq'i' te a' te wuk'a'  
   Meta  
*Traje el agua para mi bebida*  
 /ma  0  tzaj  w-eq'i-'  te  a'  t-e  w-uk'a'/  
 C.REC  A3s  DIR  E1s-traer-SC  DET  agua  E3s-SR  E1s-bebida

Vail, J. R. P. 2007. *Xt'olil Yool B'a'aj. Gramática Tektiteka*. Guatemala. p. 213

### 3.1.2 Dans les publications en français

En 1946, Leenhardt (1946) ne donne que la traduction littérale de ses exemples.

- 9 *e to dē pwiapwapwa* = il appelle un magicien.  
*e mage wa pwi u ukai* = le chef est malade.

Leenhardt, M. 1946. *Langues et dialectes de l'Austro-Mélanésie*. Paris: Institut d'ethnologie. p. 81

En 1995, dans l'ouvrage de syntaxe générale de D. Creissels, les gloses deviennent plus précises sans pourtant l'être autant que les gloses des ouvrages

publiés en anglais.

- 10 (74) d̩q̩hòmó tsé m̩òsádí ód̩b̩óñìḡ kwá n̩òk̩éḡ  
| VACHES | LESQUELLES | femme | elle LES a vues | là | rivière-loc. |  
« les vaches que la femme a vues au bord de la rivière »

Creissels, D. 1995. *Eléments de syntaxe générale*. Paris: Presses universitaires de France. p. 321

En 1999, dans les publications bantu, on ne note pas de pratiques de gloses développées.

- 11 t̩ùsyélí bíbám̩b̩ó k̩ùv̩è “nous ne sommes plus des Blancs”  
t̩ùsyélí síḡḡāḡnd̩ù k̩ùv̩è “nous ne sommes plus des crocodiles”

Blanchon, J.A. 1999. “Détermination des classes tonales des nominaux enci-vili, i-sangu et inzébi”. In *Douze études sur les langues du Gabon et du Congo Brazzaville*. Lincoln Europa. p. 36

Il existe un autre type de présentation mis en oeuvre par Denis Creissels pour le tswana. La première ligne rend le tswana tel qu’il s’écrit avec son orthographe standard. Dès la deuxième ligne est donnée la traduction idiomatique. Dans la troisième ligne il donne la transcription phonétique segmentée et la glose dans la quatrième ligne. L’avantage de cette présentation est de ne pas défigurer la langue et de faciliter la lecture pour les lecteurs ayant des connaissances de tswana.

- 12 (13) a. *Lo tlaa gakgamala fa lo utlwa mo ke go itseng*  
« Vous serez surpris en apprenant ce que je sais »  
[l̩ò-t̩l̩à-χáq̩hám̩ál-à fá l̩ó-ùt̩t̩w-á mó  
S2P-FUT-être+surpris-FIN quand S2P-entendre-FIN 15/17JCT  
k̩í-χ̩ò-íts-ì-ḡ]  
S1S-O3:15/17-savoir-FIN-REL

Creissels, D. 2002. "Présentation du tswana", *Lalies*, 23. p. 65

Dans la thèse de G. Polian (2004), on trouve le modèle à trois niveaux avec des abréviations grammaticales en majuscules et le phénomène étudié en gras. On note aussi l’utilisation du symbole  $\exists$  pour marquer le « prédicat existenciel-locatif ». Les crochets ne sont pas référencés dans la liste des abréviations mais nous supposons qu’ils servent à indiquer une valeur sémantique générale.

- 13 (44) In =kati ba ay-at.  
DEM:PROX =[admiratif] où  $\exists$ -B2  
Tu es ici ! (C’est ici que tu es !)

Polian, G. 2004. *Eléments de grammaire du tsetla*. Thèse sous la direction de Loïc Boizou. p. 83

Dans l’ouvrage majeur de D. Creissels, *Syntaxe générale, une introduction typologique*, (2006) on note une volonté d’uniformisation des gloses mais elle a posé problème à l’auteur (de lisibilité notamment) et il a dû se résoudre à un compromis entre les différentes notations.

Étant donné la variété des langues citées et la nature très disparate des données disponibles, il était impossible dans les ouvrages généraux d’avoir pour toutes les langues les mêmes principes de notation des exemples.

(Creissels : 2006, xii).

Ainsi, dans les exemples donnés par Creissels, le mot *finira* peut être glosé de plusieurs façons différentes. Dans l'exemple (14), les morphèmes dans le verbe sont concrètement segmentés et glosés individuellement (3a). Mais le même exemple peut aussi être glosé comme dans l'exemple suivant (3b) sans montrer les frontières entre les morphèmes. De plus, le degré de précision des abréviations grammaticales est aussi flexible : voir par exemple la glose du morphème “-r-” dans l'exemple (3.c) qui indique le futur mais qui est remplacée par une glose très générique TAM quand la valeur de ce morphème n'est pas en question dans la discussion.

□

**14 3a. (Creissels, 2006 ,xiv)**

fini-r-a  
finir-FUT-S3S

**3b. (Creissels, 2006 ,xiv)**

finira  
finir.FUT.S3S

**3c. (Creissels, 2006 ,xv)**

finira  
finir.TAM.S3S

□

Ces exemples montrent que la glose dans un ouvrage général de typologie syntaxique n'est pas toujours uniforme dans tous les chapitres. Elle est flexible et dépend en partie du phénomène discuté. En d'autres termes, le linguiste fait le choix de la manière dont il va présenter ses exemples suivant divers paramètres : va-t-il segmenter les morphèmes ? va-t-il indiquer la valeur de chaque morphème ou rester au niveau moins précis de la catégorie ?

N. Cáceres (2010) présente ses exemples avec deux lignes pour la langue objet, le ye'kwana. Elle a en effet à rendre compte d'un certain nombre de phénomènes morpho-phonologiques, aussi dans la première ligne écrit-elle l'énoncé tel qu'il se prononce et dans la ligne du dessous, elle transcrit et segmente les morphèmes tels qu'ils se présentent en isolation. Dans l'exemple suivant, le même suffixe *jötü* est réduit en discours et est sujet à l'harmonie vocalique et les racines verbales *ajöi* et *utu* sont réduites.

**15** (2) \ref ConvChur.203

<b>ajöijo'tojo</b>	<b>nu'jö'a</b>
Ø-ajöi-jötü-tojo	n-utu-jötü-a
3O-enregistrer-ITER-INSTR	3/3-donner-ITER-NPST
'(Elle) va donner un enregistreur (lit. une chose qui sert à enregistrer)'	

Dans les exemples bantou suivants, le système de glose utilisé par le linguiste n'est pas standard : les termes ne sont pas alignés à gauche, les classes nominales ne sont indiquées que par leur numéro et l'utilisation de « s » est ambiguë.

- 16 *Mbàà bá èsónò bá wúb-ân*  
 1Mba CONN Esono s3 :2 embrasser-REC  
 ‘Mba et Esono se saluent’

*Bí á bérân à wā*  
 S2 :p PRES saluer JONC 2toi  
 ‘Nous te saluons’

*Nfónò àà dzìṅ Minkòò*  
 1Nfono s1 :1.ASP aimer 1Minko  
 ‘Nfono aime Minko’

*Nfónò bá Minkòò bádziṅ-àn*  
 1Nfono CONN 1Minko s2:2 aimer-REC  
 ‘Nfono et Minko s’aiment l’un l’autre’

Mvé, P.M. 2010. *L’expression de la voix moyenne en fang (langue bantu A75)*. Exemplier présenté lors de l’atelier de morphosyntaxe (Lyon-DDL) du 9/04/10.

### 3.1.3 Dans les publications en anglais

En 1972, Dixon dans sa grammaire du dyirbal commence à gloser certains exemples alors que d’autres, à quelques pages d’intervalle, ne sont présentés qu’avec leur traduction littérale en anglais.

- 17 *buni naḍu bagul yugubaraḡu banimbilagu*  
 fire-NOM light-IMP THERE-DAT-I yugubara-DAT come-bila-?  
 make a fire against Yugubara’s coming (i.e. a fire will keep away  
 the unpleasant Yugubara spirit)

Dixon, R.M.W. 1972. *The Dyirbal language of North Queensland*. Cambridge: Cambridge university press. p. 113

(388) *biri naḍa buḍiṅu* I didn’t take any (although I had the opportunity)

(389) *biya ṅinda balay baḍiṅu/naḍabu baḍiṅu* you didn’t fall over there (although you might well have), only I did

Dixon, R.M.W. 1972. *The Dyirbal language of North Queensland*. Cambridge: Cambridge university press. p. 119

De même dans la grammaire du tswana de Cole (1975), les exemples sont directement suivis de leurs traductions idiomatiques. Toutefois il est à noter que l’auteur indique quand il le juge nécessaire une traduction mot à mot en regard de la traduction idiomatique.

- 18 *motho mongwe-fêla* (one person, lit. person one-only)  
*motho ale mongwe* (one person, lit. person he-being one)  
*motho ale mongwe-fêla* (one person, lit. person he-being one-only)  
*nku nngwe-fêla* (one sheep)  
*pitse ele nngwe* (one horse)  
*selêpê sele sengwe* (one axe)

L'ouvrage *The structure of Jacaltec* de Craig (1977) est représentatif des premiers ouvrages à proposer des gloses. A chaque terme de la langue objet correspond une abréviation ou un terme de la langue cible. Qui plus est, elle rend les déictiques par des chiffres et surtout la glose et le terme correspondant de la première ligne sont alignés à gauche. Dans la troisième ligne on trouve la traduction idiomatique. A noter que le terme en question est souligné et qu'elle utilise deux marqueurs différents pour les frontières morphème à morphème. On peut également mentionner l'utilisation d'abréviations renvoyant à des classes grammaticales « asp » et « cl » pour respectivement *aspect* et *classifier* plutôt que directement à leurs valeurs dans ce contexte.

- 19 (134) tita' ch-∅-aw-il-a' naj txoŋbal tinaŋ  
 maybe asp-A3-E2-see-fut cl/him market today  
 'maybe you will see him in the market today'

Grinevald-Craig, C. 1977. *The structure of Jacaltec*. Austin and London: University of Texas Press. p. 87

En 1978, avec sa grammaire maya du mam, N. England arrive à un certain niveau de précision dans ses gloses. Elle fait usage de la distinction *minuscule/majuscule* pour différencier les abréviations grammaticales des lexèmes. On remarque aussi que comme dans la grammaire du jacaltec le phénomène linguistique décrit est souligné.

- 20 at-∅ juun xjaal n-∅-taan t-uj  
 LOC PRED-3sA ONE PERSON prog-3sA-SLEEP 3s-RN/in  
 juun t-tx'aaqan jaa t-jaq' chik'uul t-e  
 ONE 3s-OLD HOUSE 3s-RN/under MOUNTAIN 3s-RN/pos  
 Chna7jal  
 Huehuetenango  
 'There was a person who slept in his old house under  
 the mountain of Huehuetenango.'

England, N. C. 1983. *A Grammar of Mam, a Mayan language*. Austin: University of Texas Press.

La grammaire tzutujil de Dayley (1985) reprend certains principes de glose comme les trois lignes, l'alignement à gauche et l'utilisation des chiffres pour les personnes. On remarque tout de même que ses gloses sont peut-être moins précises en ce que beaucoup de morphèmes grammaticaux sont rendus par des morphèmes équivalents de la langue cible au lieu d'abréviations grammaticales et que rien ne distingue les trois lignes typographiquement parlant.

21 (106) Jar ee k'ulan pa q'atb'al tziij, ja wi ma  
 cleft B3p married in courthouse cleft (?) if not  
 ee k'ulan ta pan iigléesya, ja cheqe  
 B3p married irreal in church cleft just  
 kik'amoon pon kii7, ma ee  
 B3-A3p-have-taken humbly each-other not B3p  
 q'alajinaq ta chwach Dyoos.  
 have-appeared irreal in-front-of God  
 'It's they who are married in the courthouse, (it's)  
 if they are not married in the church, it's they who  
 have just taken each other that have not appeared  
 before God.'

Dayley, J. P. 1985. *Tzutujil Grammar*. Berkeley: University of California Press. p. 387

Les trois volumes édités par T. Shopen (1985) furent pionniers en leur temps car ils essayèrent de coordonner le début des études typologiques et de donner un cadre commun dans le contexte de réaction à la dominance de l'école générativiste chomskyenne. Cette collection, écrite par des auteurs différents ne présente néanmoins aucun effort de coordination des systèmes de glose.

En général les morphèmes et mots, en marge du phénomène étudié, sont glosés par leur traduction littérale dans la langue cible (ici l'anglais) comme dans (1) qui illustre l'encodage de la négation en français. Dans un autre chapitre qui traite de la définitude du sujet, ce dernier et le verbe sont glosés de manière bien détaillée en précisant les fonctions de tous les morphèmes (2). Dans ces exemples sont pertinentes les gloses qui illustrent le phénomène traité dans le chapitre. La manière de gloser dépend totalement du phénomène étudié.

1  
 22 Jean  ne  veut  pas  manger   
 Jean  NEG  wants  not  to eat   
 'Jean doesn't want to eat'

Schachter, P., 1985<sup>1</sup>, "Parts-of-speech systems", In *Language Typology and Syntactic Description: Clause structure*, p.3-61, Cambridge University Press. p. 59

2

Úmwáana  a-ra-lir-a   
 Child  3SG-PRES-cry-ASP   
 a. 'The child is crying'   
 b. \*'A child is crying'

Kimenyi, A. 1980. *A relational grammar of Kinyarwanda*. Berkeley, Los Angeles, London : University of California Press. p. 78

Dans le mémoire de master de R. Zavala (1993) sous la direction de C. Craig de l'université d'Oregon (USA), il est intéressant d'observer qu'il utilise plusieurs marqueurs de frontières, notamment le trait d'union et les deux-points. Et il se sert de deux lignes pour gloser : une ligne pour les « étiquettes » lexicales et grammaticales et une autre ligne pour les rôles sémantiques qui sont de première

---

<sup>1</sup> Deuxième édition 2007



importance ici pour illustrer l'emploi des directionnels.

23	(9)	'ey-ø-aa-toj	no' tz'ikin
	AKA	EXIST-B3-DIR:up-DIR:thither	NCL bird
		<u>Head Path</u>	<u>Figure</u>
		s-q'ab' te' tu'	
		A3-arm NCL DIST	
		<u>Ground</u>	
		'The birds are in the branch of that tree'	

Zavala, R. 1993. *Clause integration with verbs of motion in Mayan languages*. Mémoire de master dirigé par C. G. Craig. University of Oregon. p. 11

Dans la grammaire de l'evenki de I. Nedjalkov, les gloses sont très précises, tous les morphèmes grammaticaux sont rendus par des abréviations mais ils ne sont pas différenciés des lexèmes par la typographie. D'ailleurs les trois lignes des exemples ne sont pas différenciées non plus.

24	(726)	Ollomimni bira-va dag-re-n.
		fisherman river-accd cross-nfut-3sg
		'The fisherman crossed the river.'
	(727)	Asi hokto-vo bargisin-a-n.
		woman road-accd cross-nfut-3sg
		'The woman crossed the road.'
	(728)	Nginakin votar-va mikcha:n-e-n.
		dog fence-accd jump-nfut-3sg
		'The dog jumped over the fence.'

Nedjalkov, Igor. 1997. *Evenki*. London, New-York: Routledge. p. 180

Dans les publications austronésiennes, on trouve cette année-là la grammaire du toratán (ratahan) qui présente des exemples dont les gloses sont assez élaborées. La première ligne est réservée à la langue objet telle qu'elle s'écrit et c'est dans une deuxième ligne que sont segmentés les morphèmes. Les phénomènes linguistiques que l'exemple illustre sont mis en valeur par la typographie.

25	(166)	araq manuwang kapiale	<b>makatupók.</b>
		araq maN-tuwang kapiale	maka-tu-pok
		if AV -fell be_careful_not_to	AV.POT-INVOL-cut
		When felling (trees), be careful not to <b>accidentally cut</b> (yourself). (elicited)	

Himmelman, N. P., Wolff, J. V. 1999. *Toratán (Ratahan)*. Muenchen: Lincom Europa. p. 53

Les gloses des publications australiennes sont aussi assez développées. M. Laughren utilise un modèle à trois niveaux et elle distingue le premier niveau en le mettant en italique. Ses gloses sont alignées à gauche avec le terme correspondant de la première ligne et le morphème étudié est en gras dans les première et deuxième lignes.

26	(6)	<i>Nyiya-rla-lu</i>	<i>yarlpurru-pardu-ju</i>	<i>ya-nu</i>	<i>ngurra-kari-kirra-ju?</i>
		what-LOC-PL.SUBJ	coinitiate-DEF-TOP	go-PAST	camp-other-ALL-TOP
		'Why has (my) age-mate-brother moved to another sleeping area?'			

Laughren, M. 2001. "Forty year on Ken Hale and Australian languages". In J. Simpson, D. Nash, M. Laughren et al. (eds). *Pacific Linguistics*. p. 210

Dans cet exemple de lardil, N. Richards se sert d'un modèle similaire mais il n'indique pas dans l'exemple même quel terme est illustré, en l'occurrence l'emploi du pronom de troisième personne singulier *niya* où l'on attendrait plutôt la forme plurielle.

- 27 (18) *Nyingki yukarr, karan ngakurrwen mangarda jika?*  
 you husband where our.DU.INCL.HARM child many  
*Niya denkawakun wajbelkan laka.*  
 (they) dance-ACT white.person.GEN way  
 'Hey, husband, where are all our children? They're doing disco.'

Laughren, M. 2001. "Forty year on Ken Hale and Australian languages". In J. Simpson, D. Nash, M. Laughren et al. (eds). *Pacific Linguistics*. p. 444

Du côté de l'Himalaya, les linguistes spécialistes des langues sino-tibétaines adaptent leur façon de gloser au système en vigueur. Dans cet exemple d'un dialecte yue, la transcription phonétique de chaque morphème est donnée en regard des sinogrammes (sur la même ligne). Toutes les gloses sans distinction sont en minuscules.

- 28 (109) a. SM 自從 *tsʰʰts'ʊŋ* 認識 *zənʰʰsl* 你 *niʰʰ* 以來 *jiʰʰlaj* 李四 *liʰʰsrʰʰ*  
 from know 2sg since Li Si  
 沒 *mejʰʰ* 喝 *xrʰʰ* 過 *kwɔ* 酒 *tɕjowʰʰ*  
 neg-per-asp drink exp-asp liquor  
 b. SM 李四 *liʰʰsrʰʰ* 自從 *tsʰʰts'ʊŋ* 認識 *zənʰʰsl* 你 *niʰʰ* 以來 *jiʰʰlaj*  
 Li Si from know 2sg since  
 沒 *mejʰʰ* 喝 *xrʰʰ* 過 *kwɔ* 酒 *tɕjowʰʰ*  
 neg-per-asp drink exp-asp liquor  
 'Since knowing you, Li Si has not drunk any liquor.'

Thurgood, G., LaPolla, R. J. 2003. *The Sino-Tibetan Languages*. London: Routledge. p. 121

En ce qui concerne la famille mon-khmer, la grammaire du semelai (2004) de N. Kruspe offre un autre exemple où le contexte est résumé en quelques phrases mises en italique.

- 29 (35) *The protagonists, who live alone in the forest, return home to find that someone has mysteriously prepared a meal for them. This had happened once before.*  
 lɔc ʔyot, dom, pn-ca daʔ haʔ dɔl  
 already return AFF NMZ-eat EXIST AT house  
 (When) they returned, indeed (again) there was foodstuffs in the house.

Kruspe, N. 2004. *The grammar of Semalai*. Cambridge: Cambridge university press. p. 276

La grammaire du kulung de G. J. Tolsma (2006) de la famille sino-tibétaine présente des gloses mêlant abréviations, diverses typographies et symboles graphiques. Il utilise l'écriture en exposant <sup>pi</sup> pour signifier que le *nous* est pluriel inclusif et la flèche indique la direction d'une relation transitive.

- 30 8.1 *ridum kei-a mas-am-no hənna*  
 rite we<sup>pi</sup>-ERG lose-1p→3.NPT-NEG.NPT now  
*kei-a lam-am-ke*  
 we<sup>pi</sup>-ERG search-1p→3.NPT-ASS  
 'We<sup>pi</sup> don't want to lose our<sup>pi</sup> rituals. Now we<sup>pi</sup> are searching for them.'

La grammaire du cavineña de A. Guillaume présente des exemples glosés de manière très précise en utilisant notamment différents symboles de frontière et la typographie pour faire ressortir le morphème que l'exemple sert à illustrer. On note aussi que ce phénomène est indiqué au-dessus de l'exemple, ici le jussif négatif.

31 (6.39) Negative jussive

*Tumebae=dya* [mikwana-ja bakwa=kwana]<sub>s</sub>  
 also=FOC 2PL-GEN child=PL

*jakacha pa-ju=ama escuela=ju!*  
 be.absent JUSS-be=NEG school=LOC

‘Also, your children should not miss (lit. be absent at) school.’  
 di1648

Guillaume, A. 2008. *A grammar of Cavineña*. Berlin, New-York: Mouton de Gruyter. p189.

A. Soubrier du laboratoire DDL prépare une thèse sur la description de la morphosyntaxe du dialecte uwi de l'ikposo, langue niger-congo parlée au Togo. Elle utilise la première ligne pour écrire l'énoncé tel qu'il se prononce. Dans une deuxième ligne, elle reprend le même énoncé et transcrit chaque morphème phonologiquement. Elle-même déclare qu'elle n'utilise que le point et le trait d'union pour marquer la frontière entre deux morphèmes.

32 (3) n̄nyī máyō kpātjā là títjê yì náyúdūnwě

n̄ n̄nyī má-yō kpātjā là títjê yì nú áyú-údúnū-ě  
 COMP her.husband FUT-take machete FOC chase O3S LOC POSS3S-house-DEF  
 ‘(he said) that her husband will take a machete to chase her from the house’

Soubrier, A. 2010. Base de données de texte, Ikposo-Uwi.

3.2 Rencontre avec des linguistes chercheurs au laboratoire DDL (Lyon)

Jason Huang s'est entretenu en 2009 avec trois chercheurs du laboratoire CNRS Dynamique du Langage de Lyon. Même s'ils écrivent dans des langues différentes, français, anglais, ou espagnol et que les langues qu'ils étudient n'appartiennent pas aux mêmes familles, il leur arrive de se concerter sur des points de glose. Le but était surtout de recueillir les préoccupations de ces chercheurs très investis dans la description des langues ainsi que leurs façons de faire. Je rapporte ici leurs entretiens.

**Denis Creissels** a expliqué que dans ses travaux sur des langues spécifiques, il essaye de respecter les règles de Leipzig et en cas d'absence de terme, il regarde dans l'article de Lehmann. S'il n'arrive pas à trouver une bonne solution, il cherche alors dans les conventions d'étude de la langue. Le principe est de pouvoir communiquer facilement avec les autres, et que les travaux soient lisibles pour les chercheurs dans la communauté. D. Creissels cite l'exemple d'un morphème verbal qui marque l'absence du locuteur dans un événement passé. La glose de ce morphème n'est prévue ni dans les règles de Leipzig ni dans celles de Lehmann. Il a donc décidé de

créer une glose pour cela mais il a trouvé plus tard, dans les ouvrages d'autres chercheurs, la glose UW « unwitnessed » qui est, en même temps, assez répandu dans certains travaux. Comme ce terme est convenable pour son analyse, il l'a intégré dans son travail.

Il précise qu'il n'existe pas de normes pour les gloses en français. Sans avoir un manuel de glose en main, les chercheurs français doivent faire des recherches pour trouver une abréviation adéquate ou bien créer leur propre abréviation. Les termes anglais étant souvent transparent en français, ils peuvent généralement se traduire sans difficulté.

**Gérard Philippson** confirme qu'il ne connaît pas de tradition de glosage pour les langues bantou et suit le modèle utilisé dans les publications de Denis Creissels.

**Antoine Guillaume** écrit en anglais et, dans son ouvrage *A grammar of Cavineña* (2008), il dit qu'il n'a pas suivi un modèle précis. Il a regardé dans des grammaires, dans des ouvrages de typologie ou des articles de revues et a réutilisé certaines façons de gloser selon ses propres besoins.

Son souci majeur est d'avoir des gloses qui soient les plus transparentes possibles car certains morphèmes du cavineña ont des valeurs typologiquement marquées. La solution pour lui est de bien choisir les consonnes et voyelles, et que les abréviations restent assez « longues » pour qu'elles soient plus facilement reconnaissables et n'entraînent pas trop de travail pour le lecteur qui va devoir aller les chercher dans la liste d'abréviations.

Il peut arriver qu'une langue exprime une notion spécifique qui ne se retrouve pas dans d'autres langues. Dans ce cas-là, le chercheur doit forger un terme et une abréviation. Le choix d'un terme dépendra du cadre théorique du chercheur et aussi des principes qu'il suit dans sa façon de gloser. En l'occurrence, le choix d'un terme peut très bien évoluer au fur et à mesure que l'analyse devient plus fine.

Par exemple le morphème verbal qui marque le sémantisme de « always » est une notion qu'on ne trouve que dans la grammaire du cavineña. A. Guillaume a donc créé une abréviation « ALWS » pour ce morphème. Il aurait très bien pu choisir « AL » ou « AW » qui n'entraîneraient pas d'ambiguïtés mais en accord avec son critère de « transparence », il a choisi d'utiliser « ALWS » qui est clair et facile à reconnaître. De même pour la notion « contrary to evidence », il utilise comme abréviation CONT.EVID. Cette glose reste transparente et le terme EVID est présent dans les autres listes d'abréviations. Il a donc créé une glose et une abréviation par la combinaison de deux éléments.

La façon de gloser de **Françoise Rose** dépend du besoin du texte. Dans son travail personnel elle utilise aussi le logiciel *toolbox* pour créer une base de données. En glosant avec *toolbox*, elle utilise cinq lignes au total, c'est d'ailleurs la forme du corpus fourni dans l'annexe de sa thèse *Morphosyntaxe de l'émerillon -Langue tupi-guarani de Guyane française* (2003).

33	<i>texte original</i>	<i>owulaõwãꞥe watekotj.</i>
	<i>découpage de morpheme</i>	<i>o- wul -a -õwã -eꞥe watekotj</i>
	<i>Glose (sémantique)</i>	<i>3- monter -a -un.peu -intens en.haut</i>
	<i>Glose (partie du discours)</i>	<i>IP1 -V -a -CL -CL- ?</i>

Rose, F. 2003. *Morphosyntaxe de l'émerillon -Langue tupi-guarani de Guyane française.* □□□□□□□□

□

Ce système de gloses très complet apporte toutes les informations nécessaires et c'est un outil très efficace pour faire l'analyse syntaxique, morphologique ou phonologique. Toutefois l'utiliser dans des publications serait trop lourd. Elle adapte donc ses exemples à un modèle plus standard de présentation : elle enlève la ligne du texte originel et fusionne les deux lignes de glose et partie du discours.

### 34 *aʒe kōʒem-i-l-ehe...*

*DEM demain-i-RELN-POSTP*

*Le lendemain...*

Rose, F. 2003. *Morphosyntaxe de l'émerillon -Langue tupi-guarani de Guyane française.* □□□□□□

□

La terminologie peut être modifiée selon le public visé. Dans une tradition spécifique il peut y avoir des termes qui ne s'alignent pas sur la terminologie des grammaires typologiques générales. En fonction du public pour lequel un linguiste écrit, pour les locuteurs ou pour les chercheurs qui utilisent ces termes « traditionnels », on peut suivre cette tradition et pourquoi pas commenter l'usage de la terminologie dans les notes. D'un autre côté si l'interlocuteur est un typologue qui ne connaît pas forcément bien cette famille de langues, il vaut mieux mettre à part la tradition et reprendre les terminologies plus générales.

Par exemple dans sa thèse (2003) F. Rose suit la tradition des descriptions de langues tupi-guarani et fait par exemple référence à un morphème par la dénomination « causatif comitatif » qu'elle glose CAUS.COM. Dans les articles de typologie qu'elle a écrit sur ce sujet, elle l'appelle par contre « causatif associatif » (*sociative causative* en anglais) et le glose CAUS.SOC.

Elle essaye aussi de simplifier les abréviations quand cela ne rend pas la lecture plus difficile, ainsi pour les morphèmes de masculin ou féminin qui sont souvent à côté des marques de personne dans la langue étudiée, elle pense qu'on peut remplacer FEM par F et MAS par M et il semble que les lecteurs s'y adapte vite. Ce n'est pas par contre ce qu'a choisi A. Guillaume à cause de son souci de transparence des gloses. F. Rose remarque que quand la glose apparaît beaucoup dans un article, une abréviation minimaliste ne crée pas d'ambiguïté avec d'autres termes.

### 3.3 Glose et tradition

Le paramètre sans doute le plus contraignant pour une standardisation des gloses est la prise en compte des habitudes bien ancrées des linguistes travaillant sur des langues appartenant à des familles différentes, ce que nous avons appelé *tradition*. Il ne suffit pas en effet de « coller » une étiquette sur un phénomène linguistique plus ou moins arbitrairement, il faut aussi prendre en compte les traditions qui existent dans différentes familles. Bien que le but soit de faciliter la comparaison et l'intercompréhension entre les linguistes qu'ils appartiennent aux mêmes cercles ou qu'ils étudient des langues de familles différentes, la tradition est une force qu'il ne faut pas ignorer. Par exemple la pratique dans la famille maya est d'appeler l'ergatif SET A et l'absolutif SET B. Dans les publications russes, le terme résultant de la nominalisation d'un verbe est *masdar*.

De l'aveu même de linguistes bantouistes, il n'existe pas de tradition de glose pour les langues bantoues et même plus généralement chez les africanistes. À l'inverse la famille maya a une tradition de glose bien ancrée et cela est sans doute dû en partie à l'existence de l'association guatémaltèque OKMA dont le but est la description des langues mayas par des linguistes locaux. Par le biais de ses nombreuses publications, OKMA diffuse sa liste d'abréviations et ses pratiques de glose. Ainsi dans toutes les grammaires que nous avons pu consulter, la présentation et les abréviations restent les mêmes. Une si large utilisation d'une pratique de la glose tend à l'établir comme standard or OKMA ne publie qu'en espagnol ce qui, dans notre entreprise d'uniformisation des gloses en trois langues, conduit à de nombreux « conflits ».

Pour ce qui est des publications en russe, depuis que les exemples sont glosés, les lexèmes sont rendus par leur équivalent russe et les morphèmes par des abréviations grammaticales en anglais.

35 (2.49) *sufra.ji-k t'ant'a-jar k-e-a.*  
 скатерть-Сопт пятно-Pl Сопт-находиться-Prs

*На скатерти пятна (есть).*

Ganenkov, D.S. 2005. *Kontaknye lokaluzatsii v nakhsko-dagestanskikh jazykakh i ikh tipologicheskie paralleli*. Moscou. p. 93

## 4. Propos de notre contribution

### 4.1 Comparaison de listes d'abréviations

Dans *Syntaxe Générale une Introduction Typologique* (2006), D. Creissels résume les difficultés auxquelles il a été confronté en tant que typologue.

*Ensuite, il y a la question du choix des abréviations à faire figurer dans les gloses. Même si cela est peu satisfaisant au niveau des principes, dans la pratique il ne serait pas raisonnable de prétendre se tenir toujours de façon stricte à un usage terminologique en cohérence totale avec une analyse faite selon les principes proposés, et il faut se résoudre à des compromis avec les termes plus ou moins traditionnels dans la description des langues citées, à la fois pour des raisons de lisibilité et aussi tout simplement parce que reprendre systématiquement l'analyse de toutes les données présentées serait une entreprise démesurée.*

A partir des listes proposées dans « Interlinear morphemic glosses » (Lehmann 2004), « The Leipzig Glosing Roules » (Leipzig 2008), *A Grammar of Cavineña* (Guillaume 2008), *Syntaxe générale, une introduction typologique* (Creissels 2006) et *Morphosyntaxe de L'Emerillon : Langue tupi-guarani de Guyane française* (Françoise Rose 2003), J. Huang a généré un tableau. La liste de Lehmann étant très détaillée, (plus de trois cents abréviations), il a décidé de ne pas prendre en compte les termes qui n'étaient pas présents dans les autres.

Le choix des abréviations est parfois très différent d'une liste à l'autre. Ceci étant, la majorité des abréviations reste transparente.

	Lehmann	Leipzig	Guillaume	(français)	Creissels	espagnol	OKMA
imperfective	IPFV	IPFV	IMPFV				
verbalizer	VR, VBZ		VLZ			verbalizador	VL
antipassive	APASS	ANTIP	ANTIPASS	antipassif	ANTIPASS	antipasivo	AP

Tableau 2. Exemple comparatif de gloses

Nous constatons qu'effectivement les abréviations dans la liste de A. Guillaume sont souvent plus longues et plus transparentes que les abréviations dans les autres listes.

Pour les gloses des arguments nucléaires du verbe, par exemple pour le cas du sujet d'une construction intransitive, on trouve différentes terminologies. Creissels (2006) utilise la glose « U » (unique) pour désigner l'unique argument d'une telle construction. Donc l'abréviation « U/P » (indice se référant à un actant unique ou à un patient, dans un système de type ergatif) est présente dans la liste et la marque « U » est aussi utilisée dans son ouvrage (2006). Mais dans l'article « The Leipzig Glossing Rules » (2008) c'est l'abréviation « S » pour « single argument of canonical intransitive verb » qui est utilisée. Lehmann (2004) utilise aussi « S » pour « intransitive subject ». Les deux « S » renvoient à la même notion sémantique dans les deux propositions.

(ANG)	Lehmann	LPZ	A.G.	(FRA)	D.C	F.R
actor topic	A					
agent-like argument of canonical transitive verb		A				
transitive subject				indice se référant à un agentif	A	
patient-like argument of canonical transitive verb		P		indice se référant à un patient	P	
single argument of canonical intransitive verb		S			U	
subject	SBJ	SBJ		Sujet/ indice de sujet	S	S
intransitive subject	S					
transitive patient	P					
transitive subject	A		A			
				indice se référant à un actant unique ou à un patient, dans un système de type ergatif	U/P	

Tableau 3 . exemples comparatifs de gloses des arguments phrastiques

## 4.2 Eléments de présentation

### 4.2.1 Présentation générale des exemples

#### 4.2.1.1 Résumé du contexte

Plutôt que de citer toute une phrase pour mettre en évidence le fonctionnement d'un morphème, il est possible de ne prendre que la partie de la phrase où se trouve

le morphème et de résumer le contexte.

Dans ces exemples, A. Guillaume résume le contexte afin que le lecteur puisse comprendre l'emploi d'une certaine forme verbale.

- 36** In (6.41), a grandfather (A) communicates to his grandchild (B) his wish to go back home (after a fruitless fishing night). He uses dual hortative morphology because this wish includes the grandchild too.

(6.41) A: *Ne-diru-nuka*                      *Utsekwa!*  
           HORT.DL-go-REITR                grandchild  
           ‘Let’s go back, Grandchild!’ ps014

B: *Ne-diru =taa*                      *Baba-chi!*  
           HORT-go =EMPH                grandfather-AFFTN  
           ‘Alright, let’s go then, Grandpa!’ ps016

Guillaume, A., 2008, *A grammar of Cavineña*. Berlin, New-York: Mouton de Gruyter. p. 189

Bien sûr tout dépend de ce que l'exemple illustre, par exemple le switch-reference ne peut être illustré que dans des énoncés longs. Pour les cas où un résumé du contexte est possible, il est envisageable de remplacer les pronoms anaphoriques par les participants auxquels ils se réfèrent.

#### 4.2.1.2 Eléments culturels

Des éléments culturels peuvent poser problèmes dans la glose, nous pensons qu'il convient alors d'ajouter une note de bas de page pour les questions d'ordre culturel ou anthropologique.

- 37** ~dàhâw / ~îd<sup>n</sup> / ~dúhú

*aguacatillo*<sup>14</sup> / 1PL / passer au tamis

"Nous étions en train de passer au tamis l'*aguacatillo*." [N6.]

14 Le *aguacatillo* est un fruit très apprécié pour le haut contenu d'amidon que contient la noix. Pour le préparer, il faut d'abord rompre la dure écalure qui enveloppe la noix ; ce procédé se réalise avec la partie non aiguisée des couteaux. Ensuite, la noix est râpée avec l'écorce épineuse d'un palmier. La pâte obtenue est tamisée et lavée avec de l'eau pour obtenir l'amidon.

Ospina, A.M. 2002. *Les structures élémentaires du yuhup makú, langue de l'amazonie colombienne: morphologie et syntaxe*. sous la direction de J. Landaburu. p. 142

M. Vuillermet fait état d'un dilemme auquel elle se trouve confrontée lorsqu'elle glose des exemples d'ese ejja en espagnol. Les locuteurs de cette langue ont en effet deux prénoms: un prénom ese ejja et un prénom espagnol. Dans ses gloses, elle répète le nom ese ejja et elle donne le nom espagnol dans la traduction idiomatique et c'est en note de bas de page qu'elle donne une explication avec notamment la traduction littérale du prénom.

- 38** (2) *Macario=∅*    *Enawipa=∅*    *Shomako=∅*    *tres*    *esho'i=∅*    *jaa'oke-ki-naje*  
           M.=ABS    E.=ABS    S.=ABS    tres    niño=ABS    bajar-IR-PAS  
           *besa-a.*  
           bañarse-FIN.MOC



‘Macario, Pedro y Romuald,<sup>2</sup> los tres niños bajaron a bañarse.’ (KyBiñ.004)

Et voici son explication en note de bas de page:

*1 : Ocorre que un Ese Ejja pueda tener dos nombres, el uno en ese ejja, y el otro en castellano. Por eso no corresponden todos los nombres ese ejja en la traducción al castellano.*

Elle précise cependant qu’elle gardera les prénoms *ese ejja* dans la traduction idiomatique pour sa thèse et ajoutera plutôt un paragraphe explicatif dans une partie de sociolinguistique afin d’expliquer que les prénoms *ese ejja* peuvent être traduits soit littéralement soit par leur “équivalent” espagnol. Voyez d’ailleurs l’exemple suivant où elle n’a d’autre choix que de donner la traduction littérale dans la glose.

- 39 b. *Ino-wo'o [o-ja chii-a ino wo'o xeshe-majamaja-ka-(a)ni-jojo].*  
Fil-rouge 3-GEN père-ERG fil rouge acheter-danser/ITR-3.ERG-PRS-parce.que  
(Elle s’appelle) ‘Fil-rouge’ parce que son père n’arrête pas d’acheter du fil rouge. *Lit : son père achète-danse du fil rouge (soo\_bax.147)*

#### 4.2.1.3 - Usage de la typographie

La typographie nous paraît être une ressource encore trop peu utilisée. Le caractère compliqué des manipulations informatiques est souvent invoqué pour expliquer cela et pourtant les « raccourcis » permettent aujourd’hui un usage aisé et rapide. En particulier ces ressources nous paraîtraient avoir du rendement dans les situations suivantes :

i - Cas des paires « bipolaires »

accompli/inaccompli, animé/inanimé, certain/incertain (N. Cáceres),  
transitif/intransitif, passif/antipassif, défini/indéfini ...

Il est clair que la relation d’opposition doit être gardée dans la glose et pourquoi pas systématisée. On pourrait ainsi envisager des abréviations du type AN/inAN, ... même si c’est aller à l’encontre des usages traditionnelles.

ii -Taille de la police

Le phone [ø] est en minuscule et la marque zéro est en majuscule –Ø pour les distinguer.

Avoir la police de la ligne de glose dans une taille moindre nous paraît une façon de faire qu’il faut perpétuer. En plus de distinguer cette ligne des autres, avoir une police moindre a un avantage pratique puisque les abréviations grammaticales prennent souvent plus de place que les mots qu’elles décrivent.

Les petites majuscules sont utiles pour distinguer les abréviations abstraites des lexèmes.

iii -Usage des chiffres arabes.

Lorsque l’on est amené à utiliser les chiffres arabes pour autre chose que les personnes, on peut recommander de les différencier par la typographie. En ye'kwana, il existe trois types de démonstratifs (proche, médian et distant),

chacun d'eux possédant aussi deux formes (animée et inanimée). Il est possible de les gloser DEMA<sub>1</sub>/DEMI<sub>1</sub> (pour démonstratif animé proche/ démonstratif inanimé proche) mais le chiffre porte à confusion avec la glose de personne. Une solution est de mettre le chiffre en indice DEMA<sub>1</sub>/DEMI<sub>1</sub> et de marquer la distinction *animé/inanimé* par des minuscules (évitant ainsi toute confusion) : DEMa<sub>1</sub>/DEMi<sub>1</sub>. Comme on le verra dans le paragraphe "Paramètres variables", ce qu'on met en indice et ce qu'on met en minuscule reflète l'organisation de la langue. Pour les démonstratifs du ye'kwana, on se figure donc que les démonstratifs animés ont une racine commune, cette dernière variant en fonction du paramètre *distance*.

N. Cáceres utilise aussi les chiffres pour gloser inclusif et "exclusif. "1+2" pour inclusif et "1+3" pour exclusif.

#### 4.2.1.4 Glose et codage

Il est envisageable de prévoir plusieurs lignes, plusieurs niveaux de description pour un exemple. *A* par exemple, renvoie à « absoluif » dans certains cas et à la désignation du constituant sujet d'une construction transitive prototypique dans d'autres (par opposition à *P* et *U*). Ou encore *CONSEC* renvoie à la notion de subordonnée consécutive dans la grammaire traditionnelle mais aussi à un enchaînement d'évènements. Pour « signifier » ce dernier phénomène il convient donc d'utiliser plutôt *SEQ*. De manière générale nous pensons que ces ambiguïtés n'ont pas lieu d'être car les notions se situent à des niveaux différents. Réserver une ligne par niveau permettrait ainsi d'éviter toute ambiguïté et cela même avec des abréviations identiques. On pourrait dresser plusieurs listes de gloses, une pour chaque niveau et définir à chaque fois à quel niveau se situe la glose.

D'un autre côté *D. Creissels* considère qu'étiqueter les éléments de la syntaxe bloque le dialogue entre les linguistes plus qu'il ne le favorise. Il donne d'ailleurs l'exemple du *LACITO* dans les années 1950 et des grammaires tagmémiques qui sont très difficiles à lire. En outre, si l'on a besoin d'indiquer les parties du discours, il propose de le faire sur la ligne de glose en utilisant les crochets [ ].

### **4.2.2 Rendu dans la GMI de certains phénomènes problématiques**

#### 4.2.2.1 Syntaxe de la GMI

Nous faisons figurer cette partie ici car dans certains systèmes de glose, la même abréviation *S* est utilisée avec deux significations différentes: singulier et sujet. L'ambiguïté est levée par l'ordre dans lequel est utilisée l'abréviation:

- S précède un chiffre : S=sujet
- S suit un chiffre : S = singulier

Dans cet exemple "S3S" signifie "sujet de troisième personne singulier".

40            [Etxe    berri-ak]    sei    atal    ditu  
                  maison    neuf-ERG.SG    six    pièce    avoir.PRES.S3S.O3P  
                  'La maison neuve a six pièces'

Creissels, D. 2006. *Syntaxe générale, une introduction typologique*. Paris : Hermès Sciences : Lavoisier. p. 54

#### 4.2.2.2 Grammaticalisation

Les phénomènes de grammaticalisation ouvrent deux possibilités de glosage :

- gloser le résultat de la grammaticalisation (notamment quand le sémantisme est

complètement perdu)

- rendre le processus (surtout lorsque le sémantisme est préservé) en glosant et la signification d'origine et le « résultat » grammaticalisé.

Un cas représentatif est l'encodage de l'aspect en *ese ejja* par des verbes de posture qui se suffixent aux verbes. Dans l'exemple (41), M. Vuillermet a choisi la première solution en ne considérant de *ani* que la valeur grammaticale (1). Dans l'exemple (2), les morphèmes *-ani* et *-ki* conservent leur sémantisme d'origine. La valeur sémantique « assis » (2a) et « debout » (2b) peut ainsi être indiquée, en même temps que la valeur grammaticale acquise : les deux sont séparées par la barre oblique.

- 41 (1) a. *Daki wana-bame-ka-(a)ni.*  
habit poser-bien-3ERG-PRS  
'Elle range des habits.'
- b. *Iya-bame-ka-(a)ni.*  
plier-bien-3.ERG-PRS  
'Elle plie (des choses).'
- (2) a. *Esho'i-ø taaa-(a)ni.*  
enfant-ABS crier-assis/PRS  
'Un enfant crie (assis).'
- b. *Esho'i taaa-ki.*  
enfant-ABS crier-debout/PRS  
'Un enfant crie (debout).'

Dans l'exemple suivant il est superflu de montrer l'origine. Il présente une marque de présent qu'elle aurait pu marquer "assis/PRS" mais cela aurait alourdi la glose et surtout l'exemple n'illustre pas la notion d'aspect mais celle de reduplication du verbe pour l'intransitiviser.

- 42 *Oxaña-a ekwana ba-ka-ani meshijaji-jo ixya-ixya-jya-(a)ni.*  
tous-ERG 1EXCL.ABS regarder-3-PRS sable-LOC manger-RED-sans.soin-PRS  
'Tout le monde nous regarde **manger** n'importe où sur la plage.' (KaEkí.018)

#### 4.2.2.3 Paramètres variables

Pour une abréviation donnée de valeur grammaticale, nous recommandons la mise en indice du paramètre variable.

Prenons le cas de la focalisation en wolof. On trouve dans cette langue deux morphèmes de focalisation, l'un indique que le sujet est focalisé tandis que l'autre indique qu'un autre terme (que le sujet) est focalisé. Ces deux morphèmes ont en commun d'être des focalisateurs, ce qui les différencie est le terme auquel chacun d'eux s'applique. C'est donc ce dernier paramètre que nous proposons de gloser en indice.

focalisateur	FOC	focalizador	focaliser
focalisateur du S	FOC <sub>S</sub>	focalizador del S	S. focaliser
focalisateur de O	FOC <sub>O</sub>	focalizador del O	O. focaliser

focalisateur de V	FOC <sub>v</sub>	focalizador del V	V. focaliser
focalisation d'un autre terme que le sujet	FOC <sub>x</sub>	focalizador de otro término que el S	focaliser of another term than S.

A-L. Dotte rapporte que dans les langues océaniques les classificateurs sont des morphèmes servant à construire la possession ou la numération, etc. Si la langue ne contient qu'un seul type de classificateur, CL sera utilisé sans plus de précision, s'il y a en plusieurs comme en ponapéen on précise alors *num*, *poss*, etc.

- 43 lepi -n                      tuhke sil    -lep  
 oblong.piece-GEN    tree    tree    three-CL.OBLONG.PIECE  
 'three oblong pieces of wood'

Carlson, Robert & Payne, Doris. 1989. *Genitive Classifiers, Proceedings of the 4th Annual Pacific Linguistics Conference*. University of Oregon, Eugene. p. 96

Ici la valeur précise du classificateur pourrait très bien être indiquée en indice : CL<sub>num</sub>, CL<sub>pos</sub>, ...

#### 4.2.2.4 Abréviation sous-spécifiée

Bien que, comme Lehmann et l'équipe de Leipzig, nous convenions qu'il ne faut pas gloser par la catégorie grammaticale, nous souhaiterions cependant tempérer ce parti. Plusieurs cas peuvent demander des abréviations moins précises: le public à qui est destiné un exemple glosé peut être plus ou moins spécialiste ou un morphème peut-être si polysémique que lui donner une abréviation de fonction induirait en erreur. Dans les deux cas, il est envisageable d'avoir deux niveaux d'abréviations, de définir des « hypergloses » et des « hypogloses ». Une abréviation générale non-spécifique et plusieurs abréviations, chacune « étiquetant » une valeur particulière. Le linguiste pourrait ainsi gloser plus ou moins précisément selon un effet de zoom/dézoom.

De même, dans des exemples qui illustrent un point particulier, des gloses plus détaillées, donc plus longues, peuvent être utilisées. Dans le reste des cas, une glose sous-spécifiée sera utilisée.

Prenons l'exemple karata suivant:

- 44    **gordi    č'ana-č'o    kec'e**  
 dress    thorn-OR<sub>i</sub>[LOC]    snag:PF  
 The dress was snagged on a thorn  
 La robe s'est accrochée à une épine

Pasquereau, J. 2010. *Emplois des formes spatiales en karata*. Mémoire de master 1. Université Lyon II. p. 16

Le marqueur de configuration -č'o fait partie d'une série de huit marqueurs. Compte-tenu du sémantisme varié de chacun de ces marqueurs, il était préférable de les gloser par une valeur grammaticale très générale *configuration marker* et de les numéroter afin de ne pas induire les lecteurs en erreur. Notamment -č'o a une quantité de significations différentes et gloser une valeur spécifique serait faux pour un beaucoup d'exemples. Cette situation est différente des cas où on peut gloser par

des abréviations bien connues (illatif, superessif, ...) dont les linguistes connaissent bien les sphères d'emplois.

Dans la même veine, Denis Creissels donne en exemple: le cas où un déterminant peut être à un niveau d'évolution où il n'a plus de valeur précise et où gloser «D» pour «déterminant» est la meilleure solution.

De la même manière, en ce qui concerne la glose des particules de discours, faut-il absolument trouver un équivalent dans la langue cible ou bien toutes les gloser de la même manière, "DISC" par exemple, et renvoyer à une section spéciale pour une explication plus détaillée?

#### 4.2.2.5 La polysémie

*Les lexèmes polysémiques:*

Faut-il rendre le caractère polysémique d'un terme dans la glose? Il nous semble que ne pas faire varier la glose d'un terme polysémique risque de rendre la lecture difficile et d'exotiser la langue. Gloser par la traduction appropriée au contexte tout en évitant de faire varier inutilement la glose (pour des questions de style) nous paraît être un compromis raisonnable.

*Les morphèmes polysémiques :*

Nous souhaitons aborder ici le cas des morphèmes encodant plus d'une fonction, c'est par exemple le cas du suffixe «-l» en karata qui encode l'ergatif et l'instrumental. Ce point est abordé dans la proposition de Lehmann (2004, p. 12) à la section « Ambiguity » où il propose trois solutions :

-l'utilisation d'une abréviation « englobante », ici ERG/INSTR

-l'utilisation d'une abréviation différente suivant le contexte, soit ERG soit INSTR

-l'utilisation de l'abréviation d'un nouveau terme

D. Creissels fait observer que dans la pratique, une solution consistant à privilégier dans l'étiquetage l'une des fonctions possibles est souvent implicitement appliquée. Notamment, dans l'étiquetage des cas, il est courant qu'une marque casuelle soit étiquetée en tenant compte uniquement de sa fonction dans le marquage des fonction nucléaires, sans tenir compte de possibles fonctions périphériques.

#### 4.2.2.6 Valeur par défaut

Est-il utile de gloser "singulier" (en dehors du cas des morphèmes portemanteaux) s'il n'existe pas de morphème singulatif particulier? On a observé plusieurs usages: rien au singulier et un + en exposant pour le pluriel (D. Creissels), S et P (Lehmann) ou encore rien par défaut au singulier et P au pluriel (C. Grinevald).

### **4.3 Principes pour la création d'abréviations**

#### **4.3.1 Les principes**

##### 4.3.1.1 Principe d'économie

Il nous paraît raisonnable d'arriver à une abréviation dont le nombre de caractères ne dépasse pas six lettres.

⇒ Sous principe de fréquence

Les phénomènes les plus récurrents sont les plus faciles à retenir donc leurs étiquettes de glose peuvent être moins explicites/plus courtes.  
SIM/SIMULT: Lehmann glose SIM mais étant donné la faible fréquence de ce phénomène dans les langues, il vaut mieux opter pour le deuxième choix plus transparent.

⇒ Sous-principe de préférence des consonnes  
Il consiste à supprimer les voyelles plutôt que les consonnes.

#### 4.3.1.2 Principe de transparence

Les abréviations doivent rester faciles à retenir et “suggestives”.

#### 4.3.1.3 Principe de respect de la tradition

Nous essayons dans la mesure du possible de garder les abréviations les plus ancrées.

#### 4.3.1.4 Principe de représentativité

Pour établir des abréviations communes aux français, à l'espagnol, et à l'anglais, l'abréviation doit logiquement contenir des lettres présentes dans les termes des trois langues.

-ADJR (adjectiviseur) est à préférer à AZR par exemple qui ne fonctionnerait pas en français.

-ADVR (adverbialiseur)

-NMR (nominaliseur)

-VBR (verbaliseur)

-PSF (passif) ou PSV (justifiable en français par *voix passive*)

-aPSF (antipassif) ou aPSV

-PAS (passé)

- ...../HAB pour expliciter le sémantisme d'un verbe de posture dans le cas où il marque l'aspect d'un verbe

#### 4.3.1.5 Principe par défaut

S'il arrive que les termes renvoyant à un même phénomène linguistique ne soient pas cognats dans les trois langues dont nous nous occupons ici, nous pensons qu'il faut alors soit baser l'abréviation sur le terme anglais soit sur une racine latine.

Ainsi le terme pour “moyen” est “middle” en anglais et “media” en espagnol. Nous choisissons le recours à l'anglais “MID”. De même pour “séquentiel”, “sequential” en anglais et “secuencial” en espagnol, nous choisissons “SEQ”. A noter que, compte-tenu des conventions orthographiques de l'espagnol, on pourrait aussi admettre que “q” se lise “cu” en espagnol.

### **4.3.2 Concurrence de principes**

#### 4.3.2.1 Principe de représentativité Vs tradition

Parmi les obstacles majeurs à la création d'abréviations communes aux trois langues figure le problème des doubles consonnes en français et en anglais par rapport à l'espagnol. Comme la glose utilisée pour *allatif*, il semble qu'un certain nombre de gloses soient devenues conventionnelles du fait même de leur usage. Le

problème est que ces gloses traditionnelles transgressent notre principe d'avoir des gloses dont les caractères sont présents dans les termes des trois langues. D'un autre côté, créer de nouvelles gloses, telles PAS ou ALA (pour passé et allatif respectivement), c'est aller à l'encontre d'une tradition de plusieurs dizaines d'années.

Problèmes relevés :

ALL Vs alativo

ADI Vs additive

AFI Vs affirmative (ou bien AFM)

Dans le cas précis des doubles consonnes, garder une abréviation à une consonne nous semble être un compromis peu "gênant" dans la mesure où elle en devient même plus transparente.

#### 4.3.2.2 Principe d'économie Vs tradition

L'abréviation pour *causatif*: CAUS va dans le sens de la tradition mais CAU va dans le sens de l'économie

#### 4.3.2.3 Tradition Vs transparence

Nous proposons tout de même quelques changements à la tradition quand cela nous semble améliorer la transparence ou bien réduire une ambiguïté.

PERF pour "perfectif" (temps ou aspect?) Pour résoudre l'ambiguïté, sont proposées les gloses suivantes :

IPFV/PFV renvoient aux temps verbaux

IPF/PF renvoient à l'aspect grammatical (inaccompli/accompli)

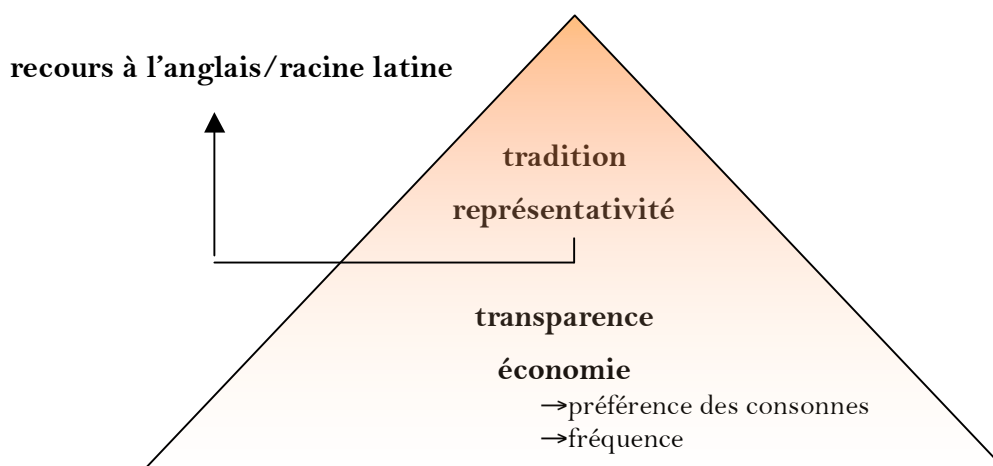
PST (past) = ambiguïté ("past" en anglais "présent" en français MAIS "pasado" en espagnol). Nous proposons PAS.

#### 4.3.2.4 Préférence des consonnes Vs transparence

Combinant la convention de représentation des gloses en majuscules et notre principe de préférence des consonnes, nous arrivons à des gloses très difficiles à déchiffrer, par exemple l'utilisation de CNFRM pour confirmatif/confirmativo/confirmative. Afin de garder une certaine transparence dans les gloses, il est donc nécessaire de "hiérarchiser" les critères suivant l'importance qu'on leur accorde.

### 4.3.3 Hiérarchie des « principes »

Ci-dessous; la liste des critères hiérarchisés du plus important au sommet au moins important à la base.



### 4.4 Liste d'abréviations

Français	ABRV.	Espagnol	Anglais
agent	A	agente	agent
abilitatif	ABIL	abilitativo	abilitative
ablatif	ABL	ablativo	ablative
absolue (forme-)	ABSL	absoluta (forma -)	absolute (form -)
absolutif	ABS	absolutivo	absolute
accompli/perfectif	PF	perfectivo	perfective
accusative	ACC	acusativo	accusative
adessif	ADES	adesivo	adessive
adjectif	ADJ	adjetivo	adjective
adjectiviseur	ADJR	adjectivizador	adjectiviser
adverbe	ADV	adverbio	adverb
adverbialiseur	ADVR	adverbializador	adverbialiser
adversatif	ADVERS	adversativo	adversative
affection	AFE	afectivo	affection
agentif	AGT	agetivo	agentive
aliénable	AL	alienable	alienable
allatif	ALA	alativo	allative
anaphorique	ANA	anafórico	anaphoric
animé	AN	animado	animate
antipassif	aPSV	antipasivo	antipassive



antiaccusatif	aACC	antiacusativo	antiaccusative
antiergatif	aERG	antiergativo	antiergative
antilogophorique	aLOG	antilogofórico	antilogophoric
appréhensif	APRH	aprehensivo	apprehensive
approximatif	APRX	aproximativo	approximative
applicatif	APL	aplicativo	applicative
associatif	ASOC	asociativo	associative
aspect	ASP	aspecto	aspect
augmentatif	AUMT	aumentativo	augmentative
auxiliaire	AUX	auxiliar	auxiliary
bénéfactif	BEN	benefactivo	benefactive
causatif	CAUS	causativo	causative
centrifuge	CTF	centrífugo	centrifugal
centripète	CTP	centrípeto	centripetal
classificateur	CLS	clasificador	classifier
classe (marque de -)	CL	clase (marca de -)	class (marker)
comitatif	COMIT	comitativo	comitative
comparatif	COMP	comparativo	comparative
complémenteur	COMPLR	complementizador	complementiser
complétive	COMPL	completivo	completive
conditionnel	COND	condicional	conditional
concessif	CONC	concesivo	concessive
constatatif	CONST	constatativo	constative
construite (forme -)	CSTR	construída (forma -)	construct
contrastif	CONTR	contrastivo	contrastive
converbe	CVB	converbo	converb
coverbe	COV	coverbo	coverb
continu	CONT	continuativo	continuous
copule	COP	copula	copula
coréférentiel	COREF	coreferencial	coreferential
datif	DAT	dativo	dative
déclaratif	DECL	declarativo	declarative
défini	DEF	definido	definite
démonstratif	DEM	demonstrativo	demonstrative
dérivatif	DER	derivativo	derivational m.
désidératif	DES	desiderativo	desiderative
déterminant	DET	determinante	determiner
diminutif	DIM	diminutivo	diminutive

direct	DR	directo	direct
directionnel	DIR	direccional	directional
discursive (particule -)	DISC	discursivo	discursive
distal	DIST	distal	distal
distributive	DISTR	distributivo	distributive
duel	DU	dual	duel
dubitatif	DUB	dubitativo	dubitative
dummy	DUM	dummy	dummy
duratif	DUR	durativo	durative
emphatiser	EMPH	enfatzador	emphatiser
épenthétique	∅	epentético	epenthetic
ergatif	ERG	ergativo	ergative
evidentiel	EVID	evidencial	evidential
exclusif	EXC	exclusivo	exclusive
expectatif	EXPCT	expectativo	expectative
explétif	EXPL	expletivo	expletive
féminin	F	femenino	feminin
focalisateur	FOC	focalizador	focaliser
futur	FUT	futuro	future
génitif	GEN	genitivo	genitive
habituel	HAB	habitual	habitual
hortatif	HORT	hortativo	hortative
hypothétique	HYP	hipotético	hypothetical
illatif	ILA	ilativo	illative
immédiat	IME	inmediato	immediate
impératif	IMP	imperativo	imperative
imparfait	IMPF	imperfecto	imperfect
imperfectif	IPFV	imperfectivo	imperfective
inaccompli/imperfectif	IPF	imperfectivo	imperfective
indéterminé	inDET	indeterminado	indeterminate
inaliénable	inAL	inalienable	inalienable
inclusif	INC	inclusivo	inclusive
indéfini	inDEF	indefinido	indefinite
indicatif	IND	indicativo	indicative
inessif	INES	inesivo	inessive
infinitif	INF	infinitivo	infinitive
instrumental	INSTR	instrumental	instrumental
intensif	INTS	intensivo	intensive

interrogatif	INT	interrogativo	interrogative
intransitif	inTR	intransitivo	intransitive
irréel	IRR	irreal	irrealis
itératif	ITER	iterativo	iterative
jussif	JUS	yusivo	jussive
locatif	LOC	locativo	locative
logophorique	LOG	logofórico	logophoric
masculin	M	masculino	masculine
moyen	MID	media	middle
négation	NEG	negación	negation
neutre	N	neutro	neuter
nominaliseur	NMR	nominalizador	nominaliser
nominatif	NOM	nominativo	nominative
non présent	nPRS	no presente	non-present
patient	P	paciente	patient
participe	PTCP	participio	participle
parfait	PF	perfecto	perfect
partitif	PART	partitivo	partitive
particule	PCL	partícula	particle
passé	PAS	pasado	past
passif	PSV	pasivo	passive
patient	P	paciente	patient
perfectif	PFV	perfectivo	perfective
perlatif	PERL	perlativo	perlative
pluriel	PL	plural	plural
possessif	POS	posesivo	possessive
post-position	POSP	posposición	postposition
potentiel	POT	potencial	potential
prédicatif	PRED	predicativo	predicative
préposition	PREP	preposición	preposition
présent	PRS	presente	present
préverbe	PREV	preverbo	preverb
progressif	PROG	progresivo	progressive
prohibitif	PROH	prohibitivo	prohibitive
proximal	PROX	proximal	proximal
question fermée (marq.)	Q	pregunta cerrada	'yes/no' question
question ouverte (marq.)	WH	pregunta abierta	'wh' question
réceptif	RECIP	recipiente	recipient

reciproque	REC	recíproco	reciprocal
réfléchi	REFL	reflexivo	reflexive
relative (prop.)	REL	relativo	relative
séquentiel	SEQ	secuencial	sequential
singulier	SG	singular	singular
subjonctif	SUBJ	subjuntivo	subjunctive
subordinateur	SBR	subordinador	subordinateur
sujet (marq. nominale)	SJT	sujeto (marca nominal)	subject (nominal m.)
sujet (marq. verbale)	S <sub>1</sub>	sujeto (marca verbal)	subject (verbal m.)
superessif	SUPES	superesivo	superessive
topique	TOP	tópico	topic
temps-aspect-mode	TAM	tiempo-aspecto-modo	tense-aspect-mood
transitif	TR	transitivo	transitive
transitiviseur	TRSV	transitivizador	transitiviser
verbe	V	verbo	verb
verbaliseur	VBR	verbalizador	verbaliser
vocatif	VOC	vocativo	vocative

□

## 5. Conclusion

Ce travail est en grande partie le reflet de quatre ateliers de réflexion auxquels ont participé des chercheurs, professeurs et doctorants travaillant sur des langues différentes et publiant dans une ou plusieurs des trois langues mentionnées. Nous sommes conscients du caractère parfois hybride que revêt cette étude mais il est volontaire. Nous n'avons pas tenté de recommander l'usage de règles ou plus généralement de « façons de faire » mais plutôt d'exposer celles qui nous paraissaient intéressantes parce qu'innovantes ou au contraire porteuses de confusion. La partie 4 est toute entière dédiée aux propositions des linguistes ayant participé aux ateliers de travail.

Nous espérons que ce travail serve à faire avancer la réflexion vers un modèle standard et évolutif de glose, un modèle proposant des règles encadrant le rendu des décisions qu'aura prises le linguiste et en mesure d'en intégrer de nouvelles. □

## Ouvrages consultés

- Anonyme. 1675. *Diccionario del idioma moxeño a través del tiempo. Estudio comparativo sobre su evolución*. Esquisse reprise par Pr. R. B. Casanovas en 1976.
- Andrews, A., 1985<sup>3</sup>, "The major functions of the noun phrase", In Shopen, T. *Language Typology and Syntactic Description: Clause structure*, p.62-154, Cambridge: Cambridge University Press
- Blanchon, J.A. 1999. "Détermination des classes tonales des nominaux enci-vili, i-sangu et inzébi". In *Douze études sur les langues du Gabon et du Congo Brazzaville*. Lincoln Europa.
- Carlson, Robert & Payne, Doris. 1989. *Genitive Classifiers, Proceedings of the 4th Annual Pacific Linguistics Conference*. University of Oregon, Eugene.
- Creissels, D. 1995. *Eléments de syntaxe générale*. Paris: Presses universitaires de France.
- Creissels, D. 2002. "Présentation du tswana", *Lalies*, 23
- Creissels, D. 2006. *Syntaxe générale, une introduction typologique*. Paris : Hermès Sciences : Lavoisier.
- Dayley, J. P. 1985. *Tzutujil Grammar*. Berkeley: University of California Press.
- De Bourgbourg, Ch. E. B. 1961. *Gramática de la lengua quiché*. Guatemala.
- Desmond, T. Cole. 1975. *An introduction to Tswana grammar*. Cape Town: Longman
- Dixon, R.M.W. 1972. *The Dyirbal language of North Queensland*. Cambridge: Cambridge university press
- England, N. C. 1983. *A Grammar of Mam, a Mayan language*. Austin: University of Texas Press.
- Ganenkoy, D.S. 2005. *Kontaknye lokaluzatsii v nakhsko-dagestanskikh jazykakh i ikh tipologicheskie paralleli*. Moscou.
- Guillaume, A., 2008, *A grammar of Cavineña*. Berlin, New-York: Mouton de Gruyter.
- Grinevald-Craig, C. 1977. *The structure of Jacalteco*. Austin and London: University of Texas Press
- Himmelman, N. P., Wolff, J. V. 1999. *Toratán (Ratahan)*. Muenchen: Lincom Europa.
- Isém, R. M. 2007. *Rikemiik li Tujaal Tzuij. Gramática Sakapulteka*. Guatemala.
- Ixmatá, P. G. 1997. *Rukeemiik ja tz'utujil chii'. Gramática Tz'utujil*. Editorial Cholsamaj.
- Kimenyi, A. 1980. *A relational grammar of Kinyarwanda*. Berkeley, Los Angeles, London : University of California Press.
- Kruspe, N. 2004. *The grammar of Semalai*. Cambridge: Cambridge university press
- Laughren, M. 2001. "Forty year on Ken Hale and Australian languages". In J. Simpson, D. Nash, M. Laughren et al. (eds). *Pacific Linguistics*.
- Leenhardt, M. 1946. *Langues et dialectes de l'Austro-Mélanésie*. Paris: Institut d'ethnologie.
- Lehmann, C. 2004. "Interlinear morphemic glosses". University of Erfurt Leipzig. 2008. "The Leipzig Glossing Rules: Conventions for interlinear morpheme-by-morpheme glosses: revised version" Department of Linguistics of the Max Plank Institute of Evolutionary Anthropology.
- Liao H., 2002, "The Interpretation of tu and Kavalan Ergativity". *Oceania linguistics*, 41, 1.

---

<sup>3</sup> Deuxième édition 2007

- Mvé, P.M. 2010. *L'expression de la voix moyenne en fang (langue bantu A75)*. Exemplier présenté lors de l'atelier de morphosyntaxe (Lyon-DDL) du 9/04/10.
- Nedjalkov, Igor. 1997. *Evenki*. London, New-York: Routledge.
- Ospina, A.M. 2002. *Les structures élémentaires du yuhup makú, langue de l'amazonie colombienne: morphologie et syntaxe*. sous la direction de J. Landaburu.
- Pasquereau, J. 2010. *Emplois des formes spatiales en karata*. Mémoire de master. Dir: C. Grinevald. Université Lyon II.
- Pixabaj, T.A.C. 2007. *Jkemiik yoloj li Uspanteko. Gramática Uspanteka*. Guatemala.
- Polian, G. 2004. *Éléments de grammaire du tseltal*. Thèse dirigée par M. G. Rebuschi. Université Paris III-Sorbonne Nouvelle.
- Pothier, B. (ed) 1978. *America latina en sus lenguas indígenas*. Caracas, Monte Avila: Unesco.
- Rose, F., 2003, *Morphosyntaxe de l'émerillon -Langue tupi-guarani de Guyane française*, Thèse sous la direction de C. Grinevald, Université Lumière Lyon 2.
- Schachter, P., 1985<sup>4</sup>, "Parts-of-speech systems", In Shopen T. *Language Typology and Syntactic Description: Clause structure*, p.3-61, London, Cambridge: Cambridge University Press.
- Shopen, T.(ed.), 1985<sup>5</sup>, "*Language Typology and Syntactic Description: Complex Construction*", Cambridge, London: Cambridge University Press
- Thurgood, G., LaPolla, R. J. 2003. *The Sino-Tibetan Languages*. London: Routledge.
- Tolsma, G. J. 2006. *A grammar of Kulung*. Leiden, Boston: Brill.
- Vail, J. R. P. 2007. *Xtxolil Yool B'a'aj. Gramática Tektiteka*. Guatemala.
- Vuillermet, M. 2009. "Complexité aspectuelle: le cas de l'ese ejja". *Séminaire des Linguistes*, STL, Université Lille 3, (3 avril).
- Zavala, R. 1993. *Clause integration with verbs of motion in Mayan languages*. Mémoire de master dirigé par C. G. Craig. University of Oregon.
- Zubiri, J. O. 2002. *Gramática moja ignaciana*.

---

<sup>4</sup> Deuxième édition 2007

<sup>5</sup> Deuxième édition 2007

## ANNEXE

### Questionnaire

**Mother tongue:** .....

**What is your favourite abbreviation ?**

passif/passive/pasivo	PSF	PSV	indifferent	other: .....
passé/past/pasado	PAST	PAS	indifferent	other: .....

**Which abbreviation catches your eye most ?**

affectif /affective/afectivo	AFE	AFF	indifferent	other: ....
allatif/allative/alativo	ALA	ALL	indifferent	other: ....
accusatif/accusative/acusativo	AC	ACC	indifferent	other: ....
affection/affective/afectivo	AFE	AFF	indifferent	other: ....
approximatif/approximative/aproximativo	APRX	APPRX	indifferent	other: ....
associatif/associative/asociativo	ASOC	ASSOC	indifferent	other: ....
illatif/illative/ilativo	ILA	ILL	indifferent	other: ....
immédiat/immediate/inmediato	IME	IMM	indifferent	other: ....

**« Bipolar » pairs : what do you think about rendering the opposition typographically ?**

passive / antipassive	PSV/aPSV	PSV/APSV	indifferent	other: ....
ergative / antiergative	ERG/aERG	ERG/AERG	indifferent	other: ....
logophoric/antilogophoric	LOG/aLOG	LOG/ALOG	indifferent	other: ....
accusative / antiaccusative	AC/aAC	AC/AAC	indifferent	other: ....
animate / inanimate	AN/inAN	AN/INAN	indifferent	other: ....