

Alice VITTRANT

Comment constituer un corpus d'étude.

Exemple d'une étude linguistique sur le birman vernaculaire

**Comment constituer un corpus d'étude.
Exemple d'une enquête linguistique sur le birman
vernaculaire**

Mots-clés : birman vernaculaire, linguistique de terrain, modalité, corpus.

Alice Vittrant

Comment constituer un corpus d'étude. Exemple d'une enquête linguistique sur le birman vernaculaire

Introduction

Le birman est la langue maternelle de 25 à 30 millions de locuteurs. Cette langue de la famille tibéto-birmane est presque exclusivement parlée en Birmanie. Elle peut être caractérisée comme une langue tonale à tendance isolante et monosyllabique, de type (SO)V.

Une des particularités du birman est la place importante de la modalité dans la grammaire de la langue. Ainsi, pour être complet, le syntagme verbal requiert la présence d'une particule sélectionnée parmi un jeu restreint de morphèmes, qui tous relèvent du domaine de la modalité. En outre, la modalité est présente dans différentes constructions comme la subordination ou la nominalisation (*cf.* exemples 1 & 2).

Les énoncés (a) et (b) de l'exemple (1) montrent que le morphème reliant une proposition relative au syntagme nominal qu'elle détermine est marqué pour la réalisation de l'événement — en (a) — ou l'éventualité de celui-ci — en (b) —, soit la modalité REALIS ou IRREALIS¹.

- (1) a. ... မနက်တံးက တွန့်မဂီ ခေါ်သွားတဲ့နေရာ ၊
 ... mənɛʔ-TɔN-Ka' cema' Ko kho-'θwa Te' neya¹
 ...matin-au moment-S. 1SG(F.P.) OBJ appeler-aller/AUX:DIR. REL:R endroit
 ... l'endroit où [tu] m'as emmenée l'autre matin...
- b. ... မနက်ပုန့်က တွန့်မဂီ ခေါ်သွားတဲ့နေရာ ၊
 ... mənɛʔpyaN cema' Ko kho-'θwa me' neya
 ... demain 1SG(F.P.) OBJ appeler-aller/AUX:DIR. REL:IR endroit
 ... l'endroit où [tu] m'emmèneras demain...

[C]

¹ Liste des abréviations utilisées :

ASS	assertif	IR	irréalis	PV	particule verbale
AUX	auxiliaire	NEG	négation	PVF	particule verbale
CC	concomitant	NOM	nominalisateur		finale
CLF	classificateur	PARF.	parfait	QUOT	quotatif (discours rapporté)
DEM	démonstratif	PDS	particule discursive	R.	réalis
DIR	directionnel	PL	pluriel	REL	relateur
EXCL	exclamatif	POL	politesse	S.	source (« sujet », ablatif)
F.P.	femme parlant	PP	particule de proposition	SG	singulier
H.P.	homme parlant	PROB	probabilité	SPT	spatio-temporel
INACC	inaccompli,	PTC	particule	SUB	subordonnant
				term.	aspect terminatif

Alice Vittrant

Comment constituer un corpus d'étude. Exemple d'une enquête linguistique sur le birman vernaculaire

De même, les deux nominalisateurs /Ta/ et /Ma/ présentés dans les énoncés (2a) et (2c) sont marqués respectivement pour les modalités REALIS et IRREALIS. D'autres en revanche ne véhiculent aucune information modale. C'est le cas du nominalisateur /chêq/ qui apparaît dans l'exemple (2b).

- (2) a. မတီလှပ်တာကို ပြောခိုင်းတာလေ ၊ [A5/14]
 ma'-To' lo? **Ta** Ko `pyo `khaiN Ta le
 grande sœur-PL.COLL faire **NOMR** OBJ dire commander PVF:R. PDS (insist.)
 ... [Elle] t'a demandé de dire ce que vous aviez fait.
 (litt. : [Elle] a commandé de dire ce que grande sœur et les siens ont fait.)
- b. စားပွဲခင်း ထားချက်ကို ထောက်ရှင် ၊ သူပြန်စားလိမ့်မယ် ၊
 Se'Pwe'KhiN `tha **chêq** Ko tho? yiN
 nappe poser **NOM** OBJ s'appuyer SUB:si
 θu pyaN `sa leiN me
 3SG s'en retourner manger PV:PROB. PVF:IR.ASS
 Si l'on s'appuie sur le fait que la nappe est mise sur la table, il rentrera probablement manger.
- c. သူနဲ့ တွေ့ဆုံမိတော့ မှာ မနက်ပြန် ပါ ၊
 θu ne' ceno twe' **Ma** mene?pyaN Pa
 3SG avec 1SG (H.P.) rencontrer **NOMIR** demain PV:POL.
 C'est demain que je le rencontrerai.
 (litt. : Le fait que je le rencontrerai, [c'est] demain.)

Devant l'importance de cette notion en birman, il nous a semblé pertinent d'étudier l'expression de la modalité dans cette langue.

L'enquête linguistique dont il va être question ici est le résultat d'un travail de terrain effectué en Birmanie entre 1998 et 2002. Durant cette période, nous avons séjourné à quatre reprises dans ce pays — durée des séjours comprise entre un et trois mois — pour y collecter des informations sur la langue auprès de locuteurs natifs. Nos deux derniers séjours ont été consacrés plus particulièrement à l'étude de l'expression de la modalité en birman.

Pour cette enquête, nous avons utilisé une méthode de type « empirico-inductive » qui donne la priorité aux faits sur les constructions intellectuelles, par opposition à une méthode de type hypothético-déductive. Cette méthode de travail est caractéristique de la linguistique de terrain² ; le linguiste cherche à développer la compréhension d'un phénomène à partir des faits recueillis plutôt que la validation d'une hypothèse théorique par la confrontation avec des données langagières. Cette méthode n'exclut cependant pas une bonne connaissance théorique du sujet étudié.

² La linguistique de terrain peut être définie comme l'activité de décrire et de documenter des langues peu ou pas décrites par le biais d'enquêtes menées auprès de locuteurs natifs (cf. Craig, 1979 : 4).

Alice Vittrant

Comment constituer un corpus d'étude. Exemple d'une enquête linguistique sur le birman vernaculaire

201

Ce type de recherche sur une langue « exotique » peu (ou pas) décrite implique par ailleurs une approche qualitative plutôt que quantitative dans la mesure où le chercheur n'a pas à sa disposition suffisamment de données linguistiques représentatives des faits de la langue.

La première partie de cet article sera consacrée à définir l'objet de notre étude à savoir le birman vernaculaire parlé en Birmanie Centrale, c'est-à-dire dans une région située entre Yangon (ou Rangoun) et Mandalay (cf. carte). Dans un deuxième temps, nous essaierons de rendre compte des différents éléments qui ont conditionné l'élaboration de notre corpus, pour enfin nous attacher à en décrire le contenu.

Carte de la Birmanie



Alice Vittrant

Comment constituer un corpus d'étude. Exemple d'une enquête linguistique sur le birman vernaculaire

202

1 Définir l'objet d'étude : le birman vernaculaire de Yangon

1.1 Langue littéraire vs langue vernaculaire

1.1.1 Distinction entre les deux idiomes

Le birman est écrit depuis la fin du 12^e siècle et possède un système d'écriture qui lui est propre. Cette langue littéraire, contrairement à la langue vernaculaire, a très peu évolué entre le 15^e et le début du 20^e siècle (*cf.* Allott, 1985 : 135). Et si l'on se réfère aux témoignages des Occidentaux en poste en Birmanie aux 19^e et 20^e siècles, la langue littéraire (ou écrite) et la langue vernaculaire (ou parlée) étaient nettement distinctes et ce, de façon plus marquée qu'elle ne l'est aujourd'hui. En effet, l'arrivée et le développement des médias (presse, radio, télévision...) ont contribué à ce que la langue littéraire se rapproche de plus en plus de la langue vernaculaire.

Malgré tout, la distinction entre les deux langues est aujourd'hui toujours perceptible, et les domaines d'utilisation assez nettement différenciés. Tout document écrit doit l'être en langue dite « littéraire » qu'il s'agisse de discours officiels, d'écrits administratifs, de romans, mais aussi de la presse et des panneaux de signalisation. Il y a cependant quelques exceptions à cette règle : les bandes dessinées, la transcription d'entretiens oraux (interviews), les dialogues des textes littéraires et certains romans³, sont écrits en langue vernaculaire.

Langue littéraire et langue vernaculaire diffèrent essentiellement au niveau du système de particules grammaticales : marques verbales de fin de phrase, marques syntaxiques et connecteurs. (*cf.* exemple (3)). La langue littéraire a, par ailleurs, un système grammatical plus nuancé et donc morphosyntaxiquement plus riche comme cela a été noté par Bernot (*cf.* aussi Allott, 1985 : 135) :

La langue écrite dispose, d'une manière générale, d'articulations grammaticales plus nuancées, donc d'un plus grand nombre de monèmes pour marquer ces articulations ; des monèmes

³ Les années 70 virent la naissance d'un mouvement littéraire prônant une littérature en langue « parlée » (ou vernaculaire), accessible à un plus grand nombre. Originaire de Mandalay (Birmanie centrale), ce mouvement à vocation populaire n'a pas eu l'audience souhaitée au sein du grand public, quoique quelques auteurs écrivent aujourd'hui encore des romans, ou des nouvelles en langue vernaculaire. L'un des textes de notre corpus (B2) est issu de ce mouvement littéraire (texte de Daw A Ma).

Alice Vittrant

Comment constituer un corpus d'étude. Exemple d'une enquête linguistique sur le birman vernaculaire

203

grammaticaux sont communs à la langue écrite et à la langue parlée, certains se correspondent d'un style à l'autre, mais leur forme diffère légèrement ; enfin certains sont propres à l'un ou l'autre style et ces derniers sont plus nombreux en langue écrite. (Bernot, 1980 : 16)

Dans l'exemple (3), les morphèmes grammaticaux qui n'ont pas la même forme dans l'une et l'autre langue apparaissent en gras.

- (3) a. ဤ လူ သည် သူ ၏ ဇိနား နေ သော ရွာ ၌ အလုံး ယှဉ် နေ သည် ၊ [birman littéraire]
 ʔi lu θi θu.ʔi' ʔi'za ne 'θa ɣwa ɲei? ʔa'loN ɣɔ ne θi
- b. ဒီလူ သူ ဇိနား နေတဲ့ ရွာမှာ သိပ် ယှဉ်နေတယ် ၊ [birman vernaculaire]
 di lu Ø θu' ʔi'za ne Te' ɣwa ɲa θei? ɣɔ ne Te
 DEM homme (S.) OSG.GEN amoureux se trouver REL village LOC très é.heureux AUX PV.F.R.ASS
Cet homme est très heureux dans le village où vit sa dulcinée.

1.2 Diglossie ou registre ?

Le statut dominant de la langue littéraire par rapport à la langue vernaculaire pourrait faire penser à une situation de diglossie⁴. Mais, il ne s'agit pas à proprement parler de deux langues distinctes, la différence relevant plutôt d'une variation de registre.

En effet, toute personne, même si elle n'a pas été scolarisée, comprend la langue littéraire, cette dernière étant largement diffusée par les médias, à la télévision, par exemple, où les programmes d'information comme le journal, la météo — tout comme les nombreux discours officiels retransmis — sont « récités » en langue littéraire (Allott, 1985 : 136).

1.2.1 Variation régionale

Le deuxième point nécessitant que nous définissions précisément l'objet de notre étude est l'existence de variations régionales. Le birman, comme toute langue parlée sur un grand territoire, est constitué de plusieurs dialectes. Si certains n'affichent que des différences lexicales avec le parler « standard » de la capitale, d'autres en revanche se distinguent par un système phonologique, voire morphosyntaxique différent. C'est ainsi que l'on note des différences dans l'utilisation des pronoms parmi les dialectes de la vallée centrale. Les locuteurs de Mandalay⁵ ont un emploi plus

⁴ Sur la définition de la diglossie, on peut se reporter à Blanchet (2000 : 130).

⁵ Ancienne capitale royale, Mandalay est aujourd'hui la deuxième ville du pays après Yangon la capitale. Toutes deux sont situées sur le fleuve Irrawaddi, voie navigable de grande importance qui traverse le pays du nord au sud.

Alice Vittrant

Comment constituer un corpus d'étude. Exemple d'une enquête linguistique sur le birman vernaculaire

204

généralisé des pronoms masculins de première et deuxième personnes, leur dialecte ne possédant pas autant de formes que celui de Yangon, comme le montre le tableau (1).

Tableau 1 : Pronoms de première personne et deuxième personne dans les dialectes de Yangon et de Mandalay

Pronom	Yangon	Mandalay
1SG - (familier)	ɑ /ŋɑ/	ɑ /ŋɑ/
1SG - Homme parlant (neutre)	ᵿᵿᵿᵿ /cəŋɔ/	ᵿᵿᵿᵿ /cəŋɔ/
1SG - Femme parlant (neutre)	ᵿᵿᵿ /cəmaʔ/	ᵿᵿᵿᵿ /cəŋɔ/
2SG - (familier)	ᵿᵿ /niN/	ᵿᵿ /niN/
2SG - Homme parlant (neutre)	ᵿᵿᵿᵿᵿ /kə`mya/	ᵿᵿᵿᵿᵿ /kə`mya/
2SG - Femme parlant (neutre)	ᵿᵿᵿ /ʃiN/	ᵿᵿᵿᵿᵿ /kə`mya/

En revanche, dans les dialectes des régions excentrées (Arakan, Tenasserim) ou géographiquement isolées (lac Inlé), les différences sont plutôt d'ordre phonétique ou grammatical. L'arakanais par exemple, qui est parlé à l'ouest du pays à la frontière avec l'Inde et le Bangladesh, a conservé des phonèmes disparus dans le parler « standard », comme la vibrante roulée /r/. Il a aussi conservé l'emploi de morphèmes subordonnants dans des suites de verbes, étape préalable aux grammaticalisations que connaissent les dialectes du centre de la Birmanie.

L'exemple (4) présente en (a) et (b) respectivement, un énoncé de birman « standard » tel que parlé à Yangon et son équivalent en dialecte arakanais⁶. En (a), le verbe « réparer » /pɣuɛ pyiN/ est directement suivi du verbe « poser, mettre » /`tha/ qui s'est grammaticalisé en morphème aspectuel résultatif, comme dans d'autres langues de la région⁷. En revanche, dans l'énoncé arakanais de même sens proposé en (b), le morphème subordonnant /lɔɛ/ est inséré entre le verbe principal et le verbe « poser ».

⁶ L'exemple (4b) est emprunté à Okell (1995 : 23).

⁷ Sur la grammaticalisation du verbe « poser, mettre » T, : /thaʔ/ en birman, on peut se reporter à Bernot (1980 : 311-14), Vittrant (2004 : 209) ; sur la grammaticalisation du verbe apparenté signifiant « poser, mettre » dans d'autres langues de la famille, on peut consulter Matisoff (1991 : 396), Björverud (1998 : 83), Hansson (1985 : 303), DeLancey (1991 : 9).

Alice Vittrant

Comment constituer un corpus d'étude. Exemple d'une enquête linguistique sur le birman vernaculaire

205

- (4) a. $\begin{array}{llll} \text{ပျံ့ပျံ့ဆီလီ} & \text{တီ} & & \\ \text{pyú pyiN} & \text{`tha} & & \text{Te} \\ \text{réparer} & \text{AJX:résult/(poser)} & & \text{PVF:R.ASS} \end{array}$ [birman 'standard']
[lɔs][l'] ont réparé.
- b. $\begin{array}{llll} \text{ပျံ့ပျံ့ဆီလီ} & \text{တီ} & & \\ \text{prú prɔn} & \text{lo' `tha} & & \text{Te} \\ \text{réparer} & \text{SUB AJX:résult/(poser)} & & \text{PVF:R.ASS} \end{array}$ [arakanaïs]
[lɔs][l'] ont réparé.

Après ces quelques précisions sur l'objet de notre étude — à savoir le birman vernaculaire parlé dans la vallée centrale —, nous pouvons maintenant nous intéresser aux données collectées.

2 Choix du type de données et des sources utilisables

2.1 Les contraintes de la recherche

2.1.1 Importance du thème de recherche

La collecte des matériaux linguistiques nécessaires à une étude linguistique dépend bien sûr du but que l'on s'est fixé, lequel déterminera aussi les méthodes à utiliser. Il est évident qu'une étude phonétique ou phonologique ne nécessitera pas le même type de données qu'un travail sur la syntaxe ou la pragmatique. Dans le premier cas, un enregistrement de mots isolés pourra constituer une première étape tout à fait appropriée. Mais un tel corpus ne serait d'aucune utilité pour étudier la structure informationnelle de la langue.

2.1.2 Importance de la structure de la langue

De plus, la structure de la langue étudiée peut, elle aussi, conditionner le type de données à collecter. Si l'on envisage le cas qui nous occupe, à savoir l'étude des expressions modales, une analyse syntaxique des données recueillies s'impose, ce qui implique un travail sur des phrases plutôt que sur des mots isolés. Or en birman, les phrases sont souvent constituées d'un syntagme verbal unique sans aucune référence explicite aux arguments. De ce fait, elles sont difficilement compréhensibles hors contexte. En d'autres termes, et c'est ce qui est illustré par les exemples (5) et (6), dans une phrase verbale, la présence des arguments est optionnelle ;

Alice Vittrant

Comment constituer un corpus d'étude. Exemple d'une enquête linguistique sur le birman vernaculaire

206

plus exactement, elle est liée à l'accessibilité au référent⁸ qu'ils représentent. Si le référent est « actif » (au sens de Lambrecht 1994), c'est-à-dire s'il a déjà été exprimé linguistiquement et est donc identifiable⁹, il ne sera pas nécessaire de le faire apparaître à nouveau.

Des énoncés comme (5), qui ne donnent aucune information lexicale ou grammaticale concernant l'argument à l'origine de l'action, sont très fréquents en birman. Pour entreprendre l'analyse syntaxique (et sémantique) d'un énoncé, il est donc nécessaire de prendre en compte le contexte situationnel, c'est-à-dire les informations données par les énoncés précédents.

(5) မြေအရေခွံ အိတ် ဝယ် တယ် ၊ [A2/B6]
mywe-ʔeɣe'KhwaN ʔeiʔ wɛ Te
serpent-peau (d'animal) sac acheter PVF:R.ASS
[Elle] a acheté un sac 'peau de serpent'. (sac plastique imitation peau de serpent)

L'exemple (6) est composé de deux phrases successives, extraites d'un récit enregistré en 1999 ayant trait aux légendes attachées à une montagne sacrée de Birmanie, le mont Popa. L'informateur raconte la légende d'un forgeron très puissant nommé Tindé, dont la puissance semblait menacer le roi de la région. Assassiné par ce dernier, Tindé devient alors un *nat*, c'est-à-dire un esprit tutélaire, une divinité locale à vénérer. Dans l'énoncé (6a), Tindé est exprimé par le pronom de 3^e personne /ʔu/, lequel apparaît dans la première proposition. En revanche, la deuxième proposition ne contient aucune indication linguistique — ni pronom anaphorique ni marque verbale de personne — faisant référence à Tindé. De même, en (6b), la proposition principale en fin d'énoncé ne contient aucune information explicite sur les protagonistes engagés dans l'action de « tuer ». Mais la présence dans la proposition circonstancielle précédente de l'expression lexicale « le roi » et du pronom de 3^e personne marqué comme objet par le morphème /Kɔ/ permet de reconstruire la situation et de comprendre par déduction que le roi a tué Tindé.

⁸ Une entité (ou une proposition) a le statut de référent dans le discours si le locuteur considère qu'elle est connue de son interlocuteur, qu'elle fait partie du *background* (énoncés antérieurs, situation socioculturelle...) au moment de l'énonciation (cf. Lambrecht, 1994).

⁹ Sur les notions d'identifiabilité et d'accessibilité, se reporter à Lambrecht (1994 : 76-77, 93).

Alice Vittrant

Comment constituer un corpus d'étude. Exemple d'une enquête linguistique sur le birman vernaculaire

- (6) a. သူ အရမ်း အားကောင်းတာပေါ့ ၊ ပန်းဘဲ လှုပ်တယ် ၊ [A1/74]
 မူ ဂေ-ၤယၢN ၤာ-ၤကၢN Ta po' `paN'be lo? Te
 3SG très force-é. bon, bien PVF:R.(excl) Excl. forgeron travailler PVF:R.ASS
Il est très puissant. [Il] travaille [comme] forgeron.
- b. အဲဒါကို ဘုရင်က သူ့ကို ပုန်ကန်မယ် လို့ ထင်တော့ သတ်ပစ်လိုက်တယ် ၊ [A1/75]
 ဂေ-ၤda Ko peyIN Ka' မူ' Ko poNkaN me lo' thiN To'
 DEM.anaph.-cela OBJ roi S.TOP 3sg.DAT OBJ se rebeller PVF:IR.ASS CIT. penser SUB
 မာ? pi? lai? Te
 tuer AUX.exh AUX.term PVF:R.ASS
Et donc, comme [il] pensait qu'il (Tindé) allait se rebeller [contre lui], le roi [l']a tué.

En résumé, un corpus d'énoncés isolés n'est pas exploitable pour une recherche se situant à l'intersection de la syntaxe et de la sémantique, comme celle dont il est question dans cet article.

2.1.3 Nécessité d'avoir recours à des sources extérieures

Une troisième contrainte conditionne le contenu du corpus d'étude. Elle peut paraître évidente, mais il me semble nécessaire de la rappeler ; c'est l'appartenance ou non du chercheur à la communauté linguistique de la langue qu'il étudie. En effet, un chercheur ne travaillant pas sur sa langue maternelle, ne peut se servir de sa compétence linguistique pour valider ses hypothèses. Dès lors, une méthode basée sur l'introspection est exclue dans une telle situation, et le chercheur, obligé de faire appel à des sources extérieures, constitue un corpus¹⁰ de fait.

2.2 Les sources utilisables

Différentes sources sont utilisables par le chercheur qui n'est pas locuteur natif de la langue qu'il étudie, lorsqu'il souhaite collecter des données langagières et constituer son corpus d'étude. D'une manière générale, s'il étudie une langue écrite, il sélectionnera des textes sur les sujets qu'il jugera en adéquation avec son étude. S'il s'intéresse à une langue à tradition orale ou à une langue vernaculaire, il aura recours à des enregistrements sonores.

Pour notre étude des expressions modales en birman vernaculaire, trois sources d'information se sont révélées utilisables : (1) l'enregistrement de locuteurs birmans, (2) des textes utilisant la langue vernaculaire, et (3) bien sûr l'élicitation, c'est-à-dire des enquêtes auprès de locuteurs natifs.

¹⁰ Quelle que soit la définition du mot « corpus » que l'on adopte, elle implique toujours une « collection de données langagières » que celles-ci soient sélectionnées et organisées pour servir d'échantillon de langue (cf. Habert *et al.*, 1997 : 144) ou sélectionnées en fonction de la représentativité des faits à illustrer (Hug, 2002 : 14).

Alice Vittrant

Comment constituer un corpus d'étude. Exemple d'une enquête linguistique sur le birman vernaculaire

208

2.2.1 Enregistrements

Plusieurs heures d'enregistrement de type « observation participante »¹¹ ont été collectées. Ce corpus n'a été que partiellement dépouillé, et une quinzaine de minutes d'enregistrement seulement ont pu servir à notre enquête. Le dépouillement d'un corpus oral est excessivement long et coûteux. Veronis (2004 : 9) évalue à 40 heures de travail la transcription orthographique d'une heure d'oral sur la parole simple et considère que ce temps peut être doublé ou triplé lorsque les conditions d'enregistrement ne sont pas satisfaisantes¹². Notre expérience nous incite à ajouter que ce temps augmente de façon inversement proportionnelle à la connaissance que l'on a de la langue étudiée : le traitement des données langagières d'une langue sans tradition écrite et/ou peu décrite (transcription phonologique, découpage morphologique, analyse syntaxique, catégorisation) prendra plus de temps que celui des données d'une langue connue, écrite et déjà bien décrite.

Notre corpus a été transcrit de la façon suivante (cf. Figure 1). Le birman, étant une langue écrite, la première ligne est une transcription orthographique. Elle est suivie de sa transcription phonologique, du découpage morphologique, de la glose par morphème, et enfin de la traduction de l'énoncé.

Figure 1: Traitement des données collectées, transcription du corpus

\ref	MYS_KKH4.phrase 24 / LOC: Khin Khin
\lx	မိမိကတို့က လူကတယ် ၊
\phono	mə caɪʔ tɛ' lu ka' ʔja tɛ
\mrph	mə caɪʔ tɛ' lu ka' ʔja tɛ
\gl	NEG. aimer REL:R homme S. être rare PVF:R.ASS
\trad	<i>Les gens qui n'aiment pas sont rares.</i>

2.2.2 Textes en langue vernaculaire

Les documents écrits en langue vernaculaire ont constitué une deuxième source exploitable pour notre corpus d'étude. Ils sont, rappelons-

¹¹ Ma participation se résumait le plus souvent à introduire le sujet de discussion, voire à acquiescer lorsque j'étais sollicitée par l'(les) autre(s) participant(s) à la situation d'interlocution.

¹² Chafe *et al.* (1991) considèrent pour leur part que six heures de travail sont nécessaires à la transcription d'une minute de parole en anglais américain, si à la transcription orthographique on ajoute des informations prosodiques tels que les mouvements intonatifs ou l'allongement des syllabes.

Alice Vittrant

Comment constituer un corpus d'étude. Exemple d'une enquête linguistique sur le birman vernaculaire

le, peu nombreux et d'emploi très circonscrit : dialogues dans les romans, interviews dans la presse, nouvelles d'un genre littéraire particulier, et des bandes dessinées (cf. Annexe 1).

2.2.3 Enquêtes linguistiques

La dernière source d'information extérieure utilisée pour constituer ce corpus d'étude est la compétence linguistique de locuteurs natifs. Nous avons ainsi effectué des enquêtes linguistiques semi-directives et directives —avec questions ouvertes et questions à choix limité respectivement— dont le but principal était de vérifier la validité générale des données précédemment recueillies. Ces enquêtes nous ont permis d'obtenir des informations qui n'apparaissaient pas dans les enregistrements ou dans les textes collectés, mais que nous estimions nécessaires à la compréhension du phénomène étudié, comme par exemple le classement des différentes expressions modales exprimant une probabilité (cf. exemple (7)).

Les énoncés de l'exemple (7) contenant des expressions modales épistémiques, ont été classés par nos informateurs selon leur degré de probabilité, *i.e.* du moins sûr au plus sûr.

(7)	a. ရေကန် သွား ရင် သွားလိမ့်မယ် ၊	[C/HINTH]	[-sûr]
	ye-'ku-kaN 'θwa yiN 'θwa leiN' me		
	piscine aller PTC aller PV:PROB. PVF:IR.ASS		
	(Il) ira peut-être à la piscine.		
	b. ရေကန် သွား ရင် သွားမယ် ၊		
	ye-'ku-kaN 'θwa yiN 'θwa me		
	piscine aller PTC aller PVF:IR.ASS		
	(Il) ira probablement à la piscine.		
	c. ဒီစာအုပ် မတ်နေကောင်းတယ် ၊	[C/Y YM]	
	di sa-ʔoʔ phaʔ ne 'kaN-ye'		
	DEM livre lire AUX:INAC.CC AUX-PVF:PROB		
	(Il) est probablement en train de lire ce livre.		
	d. ကျမသုငယ်ချင်း စိတ်ပူလိမ့်မယ် ၊	[A4/113]	[+sûr]
	cəma θəŋe'ChiN seiʔ-pu leiN' me		
	1SG.(F.P.) ami(e) é.inquiet PV:PROB. PVF:IR.		
	Mon amie va sûrement s'inquiéter.		

Ces enquêtes linguistiques participent pour une part non négligeable à la constitution de notre corpus d'étude.

Alice Vittrant

Comment constituer un corpus d'étude. Exemple d'une enquête linguistique sur le birman vernaculaire

210

3 La constitution d'un corpus pour l'étude de la modalité en birman vernaculaire

Après avoir évoqué les conditions préalables à une enquête linguistique de terrain, les contraintes pesant sur l'élaboration d'un corpus et les possibilités offertes au chercheur pour son étude, intéressons-nous maintenant à la constitution du corpus proprement dit.

3.1 Corpus et représentativité

Notre recherche portant sur le birman vernaculaire, notre corpus devait refléter au maximum les habitudes langagières des Birmans en général, et non celles d'un individu en particulier. Il s'agissait non pas de décrire les faits linguistiques étudiés à partir de l'idiolecte d'un individu, mais d'essayer de rendre compte de la langue parlée aujourd'hui à Yangon, et plus généralement dans la vallée centrale ; les sources d'information utilisées pour ce corpus devaient donc être variées.

Il est difficile de sélectionner un public objectivement représentatif de la population étudiée, mais on peut supposer que si, avec des personnes sociologiquement différentes, on obtient des résultats convergents, une extrapolation à l'ensemble de la communauté linguistique étudiée n'est pas déraisonnable. Nous avons donc travaillé avec des informateurs d'âge et de sexe différents, issus de milieux socioprofessionnels assez divers, de façon à ce que l'échantillon des locuteurs interrogés ait un certain degré de représentativité. Mais nous n'irons pas jusqu'à prétendre avoir constitué un corpus de référence au sens défini par Sinclair (1996)¹³.

¹³ Sinclair (1996 : 10) cité par Habert *et al.* (1997 : 144) : « Un corpus de référence est conçu pour fournir une information en profondeur sur une langue. Il vise à être suffisamment étendu pour représenter toutes les variétés pertinentes du langage et son vocabulaire caractéristique, de manière à servir de base à des grammaires, des dictionnaires et d'autres usuels fiables ».

Alice Vittrant

Comment constituer un corpus d'étude. Exemple d'une enquête linguistique sur le birman vernaculaire

211

3.2 Incidence de certains facteurs sur l'enquête

En effet certains facteurs sociaux, psychologiques ou autres, peuvent gêner l'enquête linguistique, voire en fausser les résultats (*cf.* Hug, 2002 : 27 *sq.*)¹⁴.

Peuvent ainsi perturber le déroulement et / ou le résultat de l'enquête :

(1) la position de l'enquêteur dans la communauté linguistique qu'il étudie, ou les relations qu'il entretient avec ses informateurs

(2) le niveau d'éducation des locuteurs interrogés

(3) la situation politique du pays dans lequel est parlée la langue étudiée.

Si d'autres facteurs peuvent eux aussi avoir une incidence sur l'enquête, nous nous contenterons d'illustrer ces trois-là car ils correspondent à des situations auxquelles nous avons été confrontée durant cette étude.

3.2.1 Le contexte social

La société birmane, traditionnelle et bouddhiste à 95 %, n'accorde pas une place égale aux hommes et aux femmes, et les situations de contacts entre ces deux groupes de population sont relativement réglementées. Par ailleurs, le contact de la population locale avec des Occidentaux en dehors des circuits touristiques relève encore du domaine de l'extra-ordinaire. Le pays ne s'est ouvert que récemment aux étrangers, et ces derniers, peu nombreux, jouissent d'un certain prestige auprès de la population locale.

Nous avons ainsi majoritairement établi des contacts avec des informatrices, et notre accès à la population masculine n'a pu se faire qu'au travers d'un cadre très formel, c'est-à-dire à l'occasion de cours individuels de birman, de visites répétées ou de séjours prolongés dans un même établissement, hôtel ou restaurant par exemple.

En outre, il ne nous a pas toujours été facile d'établir un rapport de confiance exempt de toute « autorité » liée au statut d'Occidentale. Or, si l'on souhaite éviter les réponses polies et arrangeantes d'un informateur qui ne veut pas déplaire ou froisser son interlocuteur étranger, il est important d'avoir développé une relation de confiance, voire d'égalité. L'anecdote suivante est révélatrice d'une part de l'importance du statut social d'un Birman au sein de sa famille et plus généralement de la société, et d'autre

¹⁴ Hug (2002 : 27 *sq.*) cite comme « facteurs de perturbation multiples risqu[ant] de fausser les résultats [d'une enquête d'usage] », les relations entre enquêteur et enquêté, le passé social personnel des informateurs, la formulation et l'ordre des questions, la pression de la norme, *etc.*

Alice Vittrant

Comment constituer un corpus d'étude. Exemple d'une enquête linguistique sur le birman vernaculaire

part, de la faculté d'adaptation langagière de certains autochtones face à un enquêteur étranger.

Moe Moe Oo est une couturière birmane d'une trentaine d'années. Fille aînée et célibataire, elle subvient aux besoins de sa mère et de l'une de ses sœurs.

Lors de mon dernier voyage, elle m'a accompagnée dans un périple à travers la Birmanie Centrale. Nous avons passé une semaine ensemble durant laquelle nous communiquons principalement en birman. Au cours de nos conversations, j'ai généralement eu recours aux pronoms personnels les plus neutres du point de vue du statut social, à savoir /cæmaɛ/ (1ère pers.sg.-Femme Parlant (F.P.)) et /òin/ (2e pers.sg.-F.P). Moe Moe Oo faisait de même. Cependant, quelques jours après notre retour, elle me fit la réflexion suivante : « Ma mère n'est pas contente, car depuis mon retour, je dis /cæmaɛ/ quand je lui parle ».

Or, avant notre périple, Moe Moe Oo employait couramment le terme de parenté /çæ`mi/ « fille » comme pronom de première personne (cf. exemple (8) ci-dessous). Ce terme, à la différence de /cæmaɛ/, reflète la volonté du locuteur de se positionner comme inférieur et respectueux vis-à-vis de son interlocuteur (sans nécessairement faire référence à la relation filiale que dénote le terme). L'emploi de /cæmaɛ/ par Moe Moe Oo a donc été considéré par sa mère comme irrespectueux. Cette utilisation outrancière du terme neutre s'était clairement développée à mon contact. En effet, tant ma connaissance imparfaite du système pronominal birman¹⁵ que ma difficulté à me positionner dans la société birmane et à utiliser comme pronoms personnels des termes de parenté ou des titres honorifiques, ont incité Moe Moe Oo à simplifier son discours et à employer de façon systématique les pronoms neutres que je connaissais bien.

- (8) a. သမီးမျှား ဒီလိုရေးနိုင်မယ် သိပ်ကောင်းမှာတဲ့ ။
be`mi-`mya di lo' ye nain yiN θei? `koN ma `be
fille-PL DEM comme écrire AUX:pouvoir SUB:si très ê.bien PVF:IR EXCL.
Ça serait bien si je (ta fille : moi et mes semblables) pouvais écrire comme cela.

En résumé, la place attribuée par la société à l'enquêteur conditionne en partie ses rapports avec les enquêtés, et cela peut modifier la représentativité du corpus.

¹⁵ Ce système que nous avons esquissé dans le tableau (1), est complexe. Outre le sexe du locuteur, il prend en compte son statut social et son âge. Noms propres, termes de parenté et titres honorifiques peuvent aussi servir de pronoms personnels.

Alice Vittrant

Comment constituer un corpus d'étude. Exemple d'une enquête linguistique sur le birman vernaculaire

213

3.2.2 Le contexte éducatif ou la pression de la norme

Un deuxième facteur perturbateur réside dans le niveau de scolarisation ou d'éducation des locuteurs interrogés. Il conditionne souvent la perception que ces derniers ont de leur langue. La pression de la tradition grammaticale est très forte, et les enquêtes linguistiques que j'ai effectuées auprès de Birmans éduqués se sont souvent révélées très difficiles à mener. Dans un pays où l'enseignant a un statut de « Maître » et les élèves celui de « disciple », il est impensable de remettre en cause la parole enseignée et de contester par exemple la validité des catégories grammaticales utilisées par les érudits birmans et héritées de la grammaire *pali*.

De plus, ces informateurs éduqués sont souvent familiers de langues occidentales comme l'anglais voire le français, et de la terminologie grammaticale qu'elles utilisent. Interrogés sur des faits de leur langue, ils ont facilement tendance à leur appliquer une terminologie familière à l'enquêteur. Il n'est donc pas rare d'entendre un Birman érudit parler de marqueur du « prétérit » ou de « passé composé » en birman, langue qui n'exprime pas morphologiquement le temps.

L'exemple (9a) illustre l'absence de morphème spécifique à l'expression de temps présent ou passé. Le syntagme verbal de cet énoncé est composé d'un verbe invariable et d'une particule modale REALIS indiquant un procès avéré (présent ou passé). Le SV de l'exemple suivant (9b) contient en plus la particule verbale (PV) /khêɛ/. Bien que la traduction puisse donner à penser que cette particule marque le passé, il n'en est rien. La PV /khêɛ/ apparaît en effet dans des énoncés faisant référence à une situation hypothétique comme en (9c), dans des énoncés impératifs comme en (9d), et ne peut de ce fait être considérée comme marqueur du passé. Elle peut également avoir un sens spatial — éloignement par rapport au centre déictique — comme en (9d).

En d'autres termes, la PV /khêɛ/, quoique souvent qualifiée par des locuteurs birmans érudits de morphème du passé, ne marque en fait qu'une distance, un déplacement (Bernot, 1980 : 228). Celui-ci peut être spatial, comme dans l'exemple (9d) ou temporel comme dans les exemples (9b) et (9c).

Alice Vittrant

Comment constituer un corpus d'étude. Exemple d'une enquête linguistique sur le birman vernaculaire

214

- (9) a. ကျမ ၎်ဒ်ဒီ၎်ကို အပ်ထားတယ် ၊ [C]
 cəma `Sə`bɛiN (ko) ʔaʔ `tha Tɛ
 1SG.(F.P) vélo (OBJ) confier AUX:résult. PVF:R.ASS
Je (leur) ai confié / leur confie mon vélo.
- b. အဲဒီရွာမှာ အပ်ထားခဲ့တယ် ၊ [A4/42]
 `ʔɛ-di ywa Ma ʔaʔ `tha **Khe'** Tɛ
 DEM.anaph village LOC confier AUX:résult. **PV:spt** PVF:R.ASS
*Je (leur) ai confié/*leur confie (mon vélo).*
(litt. : J'ai confié à ce village-là (dont on vient de parler) [mon vélo].)
- c. အဲဒီရွာမှာ အပ်ထားခဲ့တယ် ၊
 `ʔɛ-di ywa Ma ʔaʔ `tha **Khe'** me
 DEM.anaph village LOC confier AUX:résult. **PV:spt** PVF:R.ASS
Je vais (leur) confier (mon vélo), c'est sûr !
- d. ဒီစာအုပ် ယူလာခဲ့ပါ ၊
 di sa-ʔaʔ yu **khe'** Pa
 DEM livre prendre **PV:dir** PV:POL
Apportez ce livre, s'il vous plaît.

3.2.3 Le contexte politique

Le troisième élément « perturbateur », qui s'est avéré pertinent pour cette étude sur le birman vernaculaire, est la situation (géo-)politique du pays dans lequel l'enquête a été menée. Poser des questions dans un pays où la liberté d'expression est restreinte peut se révéler très délicat. Il faut à la fois rassurer les autorités sur le caractère anodin et purement linguistique des questions posées, et rassurer les enquêtés sur l'utilisation qui sera faite de leurs réponses enregistrées.

3.3 Déroulement de l'enquête : procéder de manière « intégrée »

Psychologues et sociologues font généralement précéder leurs enquêtes d'une pré-enquête. Celle-ci doit leur permettre de définir les problèmes posés par le thème sur lequel ils souhaitent travailler. Dans les enquêtes linguistiques, la problématique à étudier est souvent bien définie ; la pré-enquête contribue alors à affiner la formulation (longueur des questions, clarté de l'expression).

Alice Vittrant

Comment constituer un corpus d'étude. Exemple d'une enquête linguistique sur le birman vernaculaire

215

3.3.1 Première étape : la pré-enquête

Dans le cas d'une enquête de terrain sur une langue peu décrite, il peut s'avérer utile de procéder à une pré-enquête sur le thème étudié, c'est-à-dire de travailler sur le contenu des enquêtes directives à effectuer, en utilisant par exemple un questionnaire préparatoire.

Ainsi, notre travail sur la modalité en birman a débuté par l'élaboration d'une liste de phrases françaises contenant des expressions modales que nous avons ensuite soumise à des informateurs birmans. Le but de ce questionnaire préparatoire était triple :

(1) se faire une idée de la manière dont les Birmans expriment les notions de possibilité, d'obligation, de probabilité, et évaluent leurs dires (modalité axiologique), par le biais des traductions fournies par nos informateurs.

(2) obtenir un relevé (non exhaustif) des expressions modales possibles dans cette langue, ainsi qu'une liste des formes à rechercher dans un corpus d'énoncés attestés.

(3) aider à la formulation d'un questionnaire d'enquête complémentaire des données collectées.

3.3.2 Deuxième étape : constitution du corpus d'étude

Indépendamment du questionnaire préparatoire, nous nous sommes attachée à circonscrire notre corpus d'étude, en essayant de lui donner une certaine représentativité des habitudes langagières des Birmans. Pour ce faire, nous avons constitué trois sous-ensembles de données, chacun avec sa propre spécificité. Le premier de ces corpus spécialisés contenait des enregistrements sonores de différents locuteurs (corpus A). Le second était constitué de documents écrits en langue vernaculaire — deux bandes dessinées et une nouvelle (corpus B). Le troisième résultait d'un travail d'élicitation comprenant des énoncés forgés et des données recueillies au cours d'enquêtes de contre-vérification (corpus C) — l'ensemble de ces trois corpus spécialisés pouvant être considéré comme ayant la représentativité nécessaire à une étude sur le birman vernaculaire.

3.3.3 Troisième étape : analyse des faits et recherche théorique

La troisième étape de cette étude linguistique a consisté en un dépouillement minutieux et une analyse générale des données. Ces deux opérations, effectuées en collaboration avec des informateurs birmans, ont été suivies d'une recherche systématique dans les corpus A et B des formes modales répertoriées à l'aide du questionnaire préparatoire.

Alice Vittrant

Comment constituer un corpus d'étude. Exemple d'une enquête linguistique sur le birman vernaculaire

216

Parallèlement à ce travail, nous avons recherché un modèle théorique permettant de rendre compte des nombreuses formes modales du birman, mais aussi de leur interaction avec les autres morphèmes du système verbal. Notre choix s'est porté sur la grammaire fonctionnelle de Dik (1997) et son modèle hiérarchisé de la phrase. Cependant la confrontation des faits du birman avec ce modèle théorique nous a incitée à modifier ce dernier. Nous avons ainsi modifié le niveau d'intervention de certaines modalités, considérant que l'obligation et la permission par exemple n'impliquaient pas la participation des actants du procès au même titre que la capacité et le désir. Par ailleurs, au niveau de la proposition, nous avons distingué deux types de modalités, *i.e.* les modalités épistémiques et les modalités factuelles. En dernier lieu, nous avons souhaité intégrer les modalités de phrase (ou modalités énonciatives) en créant un cinquième niveau énonciatif (*cf.* Vittrant, à paraître). Ces modifications sont résumées dans le tableau (2) présentant la hiérarchisation des notions modales dans l'énoncé.

**Tableau 2 :
Récapitulatif des différentes modalités et des niveaux auxquels elles interviennent**

Dik	Notre modèle				
1	Niveau (A)	± Potentialité Inhérente Capacité, Désir			
	Niveau (B)	± « déontique » Possibilité, Obligation			
2	Niveau (C)	« épistémico-factuel »			
		Modalités Epistémiques Degrés de probabilité	Factualité (« actuality ») • REALIS • IRREALIS	Polarité • Négation	
3	Niveau (D)	« Subjectif »			
		Modalités Évaluatives		Modalités Évidentielles	
		APPRECIATIVE	VALEUR DE VERITE	(directe) MIRATIF TESTIMONIAL Etc.	(indirecte) QUOTATIF Etc.
(4)	Niveau (E)	« Enonciatif »			
		DECLARATIVE (Assertive)	INTERROGATIVE	IMPERATIVE	EXCLAMATIVE (expressive) POLITESSE

Alice Vittrant

Comment constituer un corpus d'étude. Exemple d'une enquête linguistique sur le birman vernaculaire

pièce et quand je reviens quelque temps plus tard, il n'y a plus rien. Je peux alors faire la remarque ci-dessus. »

L'autre partie de notre corpus d'élicitation (corpus C) regroupe les données collectées par le biais des jugements d'acceptabilité, en utilisant par exemple la « falsification » de données. Elles ont servi à vérifier l'analyse de certaines expressions modales, validant ainsi nos hypothèses.

La séquence [V-/yɪN/-V], composée d'un verbe redoublé dont les deux occurrences sont séparées par un morphème, est l'une des expressions épistémiques complexes révélées par notre pré-enquête. Elle n'apparaît qu'à un seul endroit dans notre corpus¹⁸. Nous avons donc sollicité nos informateurs pour la production d'énoncés supplémentaires (cf. exemple (11)).

- (11) သု ဒီကား ဝယ်ရင်ဝယ်မယ် ၊ [C/HINTH]
 θu di `ka we **yiN** we me
 3SG DEM voiture acheter PTC acheter PVF:IR.ASS
Il va probablement/peut-être acheter cette voiture.

L'examen des différents exemples fournis, a révélé que cette expression était régulièrement employée avec la particule verbale finale (PVF) IRREALIS /mê/. Nous avons donc souhaité vérifier cette caractéristique morphosyntaxique de l'expression, en procédant par falsification. Dans l'exemple (12b) la PVF /mê/ a été remplacée par la PVF /Tê/ à valeur REALIS. L'énoncé a alors été refusé par nos informateurs.

- (12) a. ရေကန်နဲ့ သွား ရင် သွားမယ် ၊
 ye-`ku-kaN `θwa **yiN** `θwa me
 piscine aller PTC aller PVF:IR.ASS
(Il) ira probablement/peut-être à la piscine.
- b. ရေကန်နဲ့ သွား ရင် သွားတယ် ၊
 ye-`ku-kaN `θwa **yiN** `θwa *Tɛ
 piscine aller PTC aller *PVF:R.ASS

Par ailleurs, le sémantisme des différents éléments participant de cette expression comparé au sens produit par l'expression elle-même, nous a incitée à postuler que cette expression à valeur épistémique était une

¹⁸ L'exemple de notre corpus n'est cependant pas représentatif de l'emploi le plus courant de cette expression. Outre le fait qu'elle y apparaît dans une proposition subordonnée, les deux occurrences du verbe redoublé sont toutes deux précédées de la négation.

Alice Vittrant

Comment constituer un corpus d'étude. Exemple d'une enquête linguistique sur le birman vernaculaire

219

construction syntaxique, au sens défini par Schultze-Berndt (2000 : 24)¹⁹ (voir aussi Goldberg, 1995 : 1, 5), c'est-à-dire une association de morphèmes produisant un sens particulier que l'on ne peut déduire des signifiés de chacun des morphèmes présents. Le caractère indécomposable de l'expression a été mis en évidence par l'impossibilité d'insérer certains éléments sauf à en changer le sens (*cf.* exemple (13), ce qui confirme notre hypothèse. Dans l'énoncé (13a) ci-dessous, il s'agit bien d'une construction à valeur épistémique. En (13b), l'ajout de la particule verbale (PV) spatio-temporelle /Khêʔ/ ainsi que la présence d'une pause après la particule /yɪN/, ont pour effet de rendre à chaque morphème son sens originel ; on est alors en présence d'une phrase complexe contenant une subordonnée conditionnelle introduite par /yɪN/.

- (13) a. သု လာရင်လာမယ် ၊
θu la **yɪN** la me
3SG venir **PTC** venir PVF:IR.ASS
Il viendra probablement/peut-être.
- b. သု လာရင် ၊ လာခဲ့မယ် ၊
θu la **yɪN** la **Khêʔ** me
3SG venir **SUB:si** venir **PV:spt** PVF:IR.ASS
**Il viendra probablement/peut-être.
S'il vient (je) viendrai.*

Conclusion

Si nous voulions résumer le déroulement de notre enquête, nous pourrions dire que nous avons procédé de manière « intégrée » : en partant du plus global pour aller vers le plus analytique, c'est-à-dire en partant des données les plus complexes que sont les enregistrements, pour finir par des enquêtes directives validant les différentes hypothèses émises. Il semble que cela soit la meilleure façon de procéder quel que soit le type d'enquête menée (*cf.* Blanchet, 2000 : 41). Cette généralité sur le déroulement d'une enquête ne doit pas faire oublier les spécificités liées à l'enquête linguistique, les contraintes qui pèsent sur l'élaboration d'un corpus et les conditions particulières du travail de recherche sur une langue peu ou pas décrite.

¹⁹ « Constructions can be defined as patterns which are non-compositional, in the sense that the meaning of a complex expression that instantiates this pattern could not be arrived at solely by relying on the meanings of its parts, or the meaning of other constructions. In other words, if the meaning of a complex pattern expression can only be fully stated with reference to the properties of the pattern itself, this pattern has to be recognised as a construction in its own right. »

Alice Vittrant

Comment constituer un corpus d'étude. Exemple d'une enquête linguistique sur le birman vernaculaire

220

Bibliographie

- Allott A., 1985, Language policy and language planning in Burma, in *Papers in South East Asia Linguistics 9*, Canberra : Research School of Pacific and Asian Studies, The Australian National University, 131-154.
- Bernot D., 1999, Transcrivez, transcrivez, il en restera toujours quelque chose, in Rouard A. (éd.) : *Les orientalistes sont des aventuriers. Guirlande offerte à Joseph Tubiana*, Saint-Maur : Editions SEPIA.
- Bernot D., 1980, *Le prédicat en birman parlé*, Paris: SELAF
- Blanchet P., 2000, *La linguistique de terrain. Méthode et théorie. Une approche ethno-sociolinguistique*, Rennes : PUR.
- Björverud S., 1998, *A grammar of Lalo*, Lund : Department of East Asian Languages, Lund University.
- Chafe W., du Bois J.W., Thompson S.A., Towards a new Corpus of Spoken American English, in Aijmer K. & Altenberg B. (eds.) : *English Corpus Linguistics : Studies in Honour of Jan Svartvik*, London : Longman, 64-82.
- Craig Colette G., 1979, Jacalteco : fieldwork in Guatemala, in Shopen T. (ed.) : *Languages and their Speakers*, Winthrop Publishers : Cambridge
- DeLancey S., 1991, The origins of verb serialization in modern Tibetan, *Studies in Language 15/1*, Amsterdam: John Benjamins, 1-23.
- Dik Simon C., [1989] 1997, *The Theory of Functional Grammar. Part 1 : The Structure of the Clause. Part 2 : Complexes and Derived Constructions*, Berlin: Mouton de Gruyter.
- Golberg Adele E., 1995, *Constructions : A Construction Grammar Approach to Argument Structure*, Chicago : Chicago University Press.
- Habert B., Nazarenko A., et Salem A., 1997, *Les linguistiques de corpus*, Paris : Armand Colin.
- Hansson I.-L., 1985, Verb concatenation in Akha, in Thurgood G., Matisoff J. A., & Bradley D. (eds.) : *Linguistics of the Sino-Tibetan Area : Papers presented to Paul K. Benedict for his 71st birthday*, Canberra: Australian National University (Serie C-87), 287- 309.
- Hug M., 2002, *L'enquête d'usage linguistique*, Paris : Honoré Champion.
- Jacobson M., 2004, Corpus oraux en linguistique de terrain, *TAL 45-2*, 63-88.
- Lambrecht K., 1994, *Information structure and sentence form. Topic, focus, and mental representations of discourses referent*, Cambridge : Cambridge University Press.
- Matisoff, James A., 1991, Areal and Universal Dimensions of Grammatization in Lahu, in Traugott Elizabeth C. & Heine B. (eds.) : *Approaches to grammaticalization : Focus on Theoretical and Methodological Issues*, Vol.2, Londres: John Benjamins, 383-453.

Alice Vittrant

Comment constituer un corpus d'étude. Exemple d'une enquête linguistique sur le birman vernaculaire

221

- Okell J., 1995, Three Burmese dialects, in *Papers in Southeast Asian linguistics N°13 : Studies on Burmese languages*, Pacific Linguistics series A-83, Canberra: The Australian National University, 1-138.
- Sinclair J., 1996, *Preliminary recommendations on Corpus Typology*, Rap.Tech., EAGLES (Expert Advisory Group on Language Engineering Standards) May 1996, CEE.
- Schultze-Berndt E., 2000, *Simple and Complex Verbs in Jaminjung : a study of event categorisation in an Australian language*, MPI Series in Psycholinguistics 14, Nijmegen / Wageningen.
- Veronis J., 2004, Le traitement automatique des corpus oraux, *TAL* 45-2, 7-14.
- Vittrant A., 2004, *La modalité et ses corrélats en birman, dans une perspective comparative*, Thèse de Doctorat en Sciences du Langage, Université Paris 8 Vincennes / Saint-Denis.
- Vittrant A., 2005, La modalité dans une approche typologique, in Ibrahim A. (éd.) : *La typologie en France aujourd'hui*, CRL.

Alice Vittrant

Comment constituer un corpus d'étude. Exemple d'une enquête linguistique sur le birman vernaculaire

Annexes :

Extrait d'une bande dessinée birmane



Extrait d'une interview dans un magazine birman

	<p>အမေ လုပ်ချင်တယ်ဆိုပြီး အရင်းအနှီးထုတ်ပေးလို့ လုပ်တာပါ။ သိပ်အဆင်မပြေပါဘူး။ ရှယ်ယာဆိုတော့ သိတဲ့အတိုင်းပဲလေ။ ကိုယ်က မဦးစီးနိုင်တော့ သိပ်အဆင်မပြေဘူး။ သူတို့ကိုပဲ ဆက် လုပ်ခိုင်းလိုက်ပါတယ်။ ကျွန်တော်က ကားရိုက်ပေါ်လာရင် မသွားတော့ဘူးဆိုတော့ရှယ်ယာတစ်ယောက်အနေနဲ့မကောင်းဘူးလေ။</p>	<p>ကျော်ကျော်(မင်းမော်ကွန်း) သားလေးတောင်ကလေးသရုပ်ဆောင်နေရာက ပါနေပြီ။ သတ္တိရဲ့သားလေးကိုဇာတ်ကားထဲထည့်ရိုက်မယ်ဆိုရင် သတ္တိကကော့ခွင့်ပြုမှာလား။</p>	<p>အဆင်ပြေပေါ့မလားဆိုပြီး အဲဒီစိတ်ကြီးနဲ့ အလုပ်လုပ်ရတာ သိပ်အဆင်မပြေဘူး။</p>
<p>တော့မယ်ဆိုတော့ အမျိုးသမီးက ဘာတဲ့လဲ။</p> <p>သူကတော့ ဒီအလုပ်ကို အားပေးပါတယ်။ ကျွန်တော်ကအိမ်ထောင်ကျသွားတော့ ဒီဖက်ကို သိပ်မလှည့်နိုင်ဘူး။ ခုနောက်ပိုင်းတော့အတည်တကျဖြစ်အောင် ဒီဖက်ပဲပြန်လုပ်တော့မှာပါ။</p>	<p>အမျိုးသမီးက ငွေရှာစေချင်တာလား။</p> <p>သူကတော့ အနုပညာအလုပ်ကိုပဲ လုပ်စေချင်တယ်။ သူက ဖြစ်စေချင်တယ်ပေါ့။ သူကတော့ ကျွန်တော့်သဘောပါပဲ။</p> <p>'ဘုန်းမြင့်သတ္တိ'က ဆယ်လ</p>	<p>ရင်သူတို့သားအမိလိုက်လား။</p> <p>မလိုက်ဘူး။ ကျွန်တော်သွားရိုက်တာ သျန်လျင်ဘက်မှာဆိုတော့ ဝေးလည်းဝေးတယ်။ ရာသီဥတုကတည်း အရမ်းမူတော့ကလေးကို ခေါ်မသွားချင်ဘူး။ ဇာတ်ကားရိုက်ရတာသိတဲ့အတိုင်းပဲ။ စိတ်နဲ့လုပ်ရတာဆိုတော့</p>	<p>အမျိုးသမီးက မလိုက်ချင်ဘူးလား။</p> <p>သူကတော့ လိုက်ချင်တာပေါ့။ အရင်က VCD ရိုက်ကွင်းတွေတော့လိုက်ဖူးပါတယ်။</p> <p>သားကအမေကိုပိုခင်တွယ်တယ်လို့ပြောတဲ့ သားရဲ့အဖေ သတ္တိက ဇာတ်ကားတွေသေချာပြန်ရိုက်တော့မယ်လို့ ဆုံးဖြတ်ထားပါတယ်။ သတ္တိ မေမြတ်သူနဲ့ဘုန်းမြင့်သတ္တိတို့ရဲ့ မိသားစုဘဝလေးက အခြားအနုပညာရှင်မိသားစုဘဝလေးတွေလိုပဲ ပျော်ရွှင်ဖွယ် ကောင်းလှပါတယ်။</p> <p>ခင်လိမ္မာ</p>