MULTIDIMENSIONAL TRANSLATION IN LINGUISTIC ANNOTATION

KRISHNA PRASAD CHALISE

The core components of linguistic annotation consist of transcription, interlinear glossing and free translation. A linguistic annotation incorporates different types of translations together in itself. From one perspective, the transcription is a type of transliteration and from another perspective; it is very close to translation. Similarly, interlinear glossing is linear morpheme-to-morpheme or word-to-word translation. In the same way the free translation incorporates a variety of translations from literal to idiomatic translation.

Key words: *Annotation, transliteration, interlinear, idiomatic translation*

1. Introduction

It has been a normal practice in the linguistic literature to present linguistic annotations in an interlinear format. It means that different types of annotation are presented in different tires as a result it is a kind of multi-tire annotation. Translation is an important part of linguistic annotation because as a whole linguistic annotation is a multidimensional translation itself. It is multidimensional in the sense that linguistic annotation incorporates diverse forms of translation. The most commonly used model of linguistic annotation in the literature of linguistics is the three tire model which consists of three major components: transcription of the utterance with morpheme break, interlinear glossing and free translation as presented in (1).

(1)   Greedy dog

ek	coTi	yoTaa	kukur	khaanaa-ko	khoji-maa	baahira	nisk-yo.
one	instance	one	dog	food-GEN	search-LOC	outside	come.out-3smL.PST

Once, one dog went out in search of food.   [Genetti (1994:166)]

In the later days, we find a slightly extended form of three tire model of linguistic annotation which divides the transcription level into two levels: utterance level and morphophonemic (text) level. In the utterance level, the exact transcription of the utterance is presented and in the morphophonemic (text) level, the base forms of the morphemes along with the morpheme breaks are presented. In example (2), *xha* is the surface form and in next tire it has been presented in terms of its base form as: 'xhe-a'.

(2) *do   xyungemu   dangsu   byam   xha.*
    do   syung-mu   dangsu   byam   xhe-a
    here   sit-INF   BEN   rug   bring-2SG.IMP
    'Bring the rug here for sitting (= in order to sit down).'[Wills 2007:118]

Nowadays, because of the availability of different kinds of annotation software like Toolbox, ELAN, Transcriber, etc. the job of linguistic annotation has become quite easier and advanced and the multi-tire annotation with multidimensional information has become fashionable as presented in (3).

(3)

\ref	Drought_MIN.044
\ut	bhədəu səmmə gərə cin kirop bələi
\ELAN	00:01:47.500
\ELAN	00 01 50.250
\ELAN	MIN
\txd	भदौ सम्मन् गरा चिन् किरोप् बालै
\tx	bhədəu səmmə gərə cin kirop bələi
\mb	bhədəu səmmə gərə cin ki- rop bəl -əi
\ge	Bhadau up to field EMP PFV- transplant man -ERG
\ftn	भदौसम्म बेत रोप्छन् मान्छेहरुले ।
\ft	The people plant the farm upto Bhadau (August-September).

[LEDBL Corpus 011]

There are two utterance tires, one transcribed in IPA system and another in Devanagari writing system. Similarly, there are two tires for free translation, one English translation and another Nepali translation. Apart from the tires for linguistic annotation, there are tires for other information like the name of the speaker, beginning and end time of the utterance, too.

2. Translation in the linguistic annotation

2.1 Transcription as a form of translation

In linguistic transcription the transcriber listens to the individual sounds in the speech, identifies the phonetic characters of the sounds and represents them in terms of graphic characters. It is the graphic/symbolic representation speech. It is the process of one-to-one mapping of sounds into graphemes. So it is the transformation of phonic substance (speech) into graphic substance (writing) of language. This feature shows transcription to be a kind of transliteration.

But transcription shares some features with translation, too. A speech sound is not a single entity but it is a bundle of its phonetic features as a word is a bundle of its semantic and grammatical features. In transcription, there is source system (phonetic system) and target system (the writing system). The transcriber, like in translation, first should perceive the sounds and represent them into graphemes. The perception, as in translation, is guided by the sound system of his/her native language. Edward (2001:321) notes '… no transcription can be regarded as a direct, unbiased representation of a speech event- it is by necessity filtered and influenced by the annotator's decision, usually according to his/her theoretical goals and definitions.'

2.2 Interlinear glossing

This is a type of linear translation. In a linear translation, the target language (language of analysis) morphemes/words are below the source language morphemes/words in one-to-one correspondence (Newmark1998:45). In grammatical annotation, the interlinear glossing tire is just under the text tire. In the text tire, the speech is presented in terms of sequence of morphemes and in the interlinear glossing tire; the equivalent forms in the target language corresponding to each of the morphemes in the text tire are presented as presented the shaded tires in (4). 'So interlinear glossing demands some degree of prior morphological analysis and some degree of semantic analysis. It is basically needed for the identification of morpheme break where morpheme boundaries are blurred by morphophonemic process (Schultze-Berndt 2006:240)'.

(4) **Bhote People.001**

bhotgəi	lə	nam	lə	kumbagon			nilik
bhot -gəi	lə	nam	lə	kumba -go	-n	ŋi-	lik
Bhot -LOC	EMP	house	EMP	stone -GEN	-EMP	NPST-	become

The houses also are made of stone in Bhot. [LEDBL Corpus 2011]

In the interlinear glossing, we find source language to analysis language translation and source language to meta-language translation. The translation of lexical morpheme is source language to analysis language translation, like the morpheme <*nam*> (Baram) to <house> (English) but <*lə*> (Baram) <EMP 'emphatic'> (meta-language).

2.3 Free translation

In a grammatical analysis, free translation in the language(s) of wider access is obligatory because it makes the other researchers and concerned people able to understand the linguistic annotation made by the researcher. In the sense of translation 'free translation reproduces the matter without the manner or the content without the form of original (Newmark 1998: 46)'. So a free translation is very closer to the target language rather than the source language. But in linguistic annotation, the free translation should be closer to the source language because the readers/researchers should understand the source language and its structure through the target language. Free translation in the real sense can be misleading and problematic in the linguistic annotation so the free translation in linguistic annotation is literal translation from the perspective of translation theory. In literal translation, 'the source language grammatical constructions converted to their nearest target language equivalents but the lexical words are again translated slightly (Newmark 1998:46). Literal translation is very close to the source language as presented in (5).

(5) **BhotePeople . 014**

himaŋ	jale	jedde	alu	kokaŋ	nipi		səppəigəi
hima -gaŋ	ja -le	jedde	alu	kok -iŋ	ŋi-	pi	səppəi -gəi
below -ABL	go -COND	big	potato	steam -CONV	NPST-	give	all -DAT

If (you) go from down, (they) give (you) all big boiled potato.[LEDBL Corpus 2011]

The close grammatical structure and constituent order of translation to its source in literal translation is more helpful into understanding of the structure the source language and less likely to be misleading (Schultze-Berndt 2006:232-238).

Still, the language of free translation in linguistic annotation is tried to make natural and idiomatic as far as possible to make it more readable especially for the people fluent in the target language. The pragmatic effects of the original utterances are tried to be incorporated, too. For the purpose they supply the necessary information in parenthesis to make the translation more natural. In (5) the words 'you', 'they' and 'you' are missing in the utterance because they are understood in the context but in translation they are compulsory to make the sentence natural. This technique is very useful because it helps to maintain the naturalness of the utterance as well as make the translated sentence readable.

Some linguists have tried to manage the situation in another way. They give literal translation in one tire and fully idiomatic free translation in another tire. Some linguists are found to have managed the matter in another way. They provide the literal translation just below the individual utterance and give idiomatic free translation of the whole text at the last.

3. Conclusion

Linguistic annotation is a multitier annotation in which one type of information is presented in one tire. Because of the technological advancement and purpose of the annotation we find various numbers of tires in linguistic annotations but the obligatory components of linguistic annotation are transcription, interlinear glossing and free translation. A linguistic annotation incorporates different types of translations together in itself. In one perspective the transcription is a type of transliteration and in another perspective it is very close to translation. Similarly, interlinear glossing is a linear, morpheme-to-morpheme or word-to-word translation. In the same way, the free translation incorporates a variety of translations from literal to idiomatic translation. So, linguistic annotation is a complex and multidimensional translation.

References

Genetti, Carol (ed.). 1994. *Aspects of Nepali grammar.* Santa Barbara: Department of Linguistics, University of California.

Linguistic and ethnographic documentation of the Baram language project (LEDBL). 2007-2011. An ELDP, SOAS funded MDP (MDP0158). Central Department of Linguistics, Tribhuvan University, Kathmandu.

Newmark, Peter. 1998. *A textbook of translation.* Prentice Hall Europe, UK.

Schultze-Berndt, Eva. (2006). "Linguistic annotation." In J. Gippert, N. P. Himmelmann, & U. Mosel (Eds.) *Essentials of language documentation*, 213-252. Berlin . New York: Mouton de Gruyter.

Wills, Christina M. 2007. *A descriptive grammar of Darma: An endangered Tibeto-Burman language.* A PhD dissertation submitted to the University of Texas at Austin, USA.