

GALE Evaluation Reference Translation Alternation Guidelines for Chinese

Version 1.2 – 9/24/2009

Linguistic Data Consortium

<http://projects ldc.upenn.edu/gale>

Table of Contents

1	Introduction.....	3
2	Discourse Alternations.....	4
3	Syntactical Alternations	5
3.1	Tense.....	5
3.2	Articles and Numbers	6
3.3	Pro-drop.....	6
3.4	Serial verbs.....	6
4	Idioms, colloquial expressions and political expressions	7
4.1	Idioms and colloquial expressions	7
4.2	Political expressions	8
5	Proper names	8
6	Words that can have multiple meanings in English.....	8

1 Introduction

The GALE Translation Evaluation requires a single gold standard reference translation for each document to be evaluated. This reference must be of the highest quality and will undergo multiple rounds of quality control. One challenge is the fact that in some cases there will be more than one acceptable reference translation. In many cases, these translation alternatives are simply different ways of conveying the same meaning in the source text – for instance, one translator may choose the word "also" while another translator chooses "too". Translators have many options, from word choice to sentence structure and word order. Some translation alternatives may convey different levels of formality or stylistic differences, e.g. choosing "kitties" versus "cats". In these cases, the different alternatives may affect the style or tone of the translation – and some choices may be more appropriate than others – but the meaning of the resulting translation is not changed.

The kinds of alternations described above are predictable, even for a monolingual English speaker reading the translated document without reference to the source document. For instance, consider the following sentences:

```
Kitties are animals too.  
Cats are also animals.
```

A native English speaker presented with these two sentences would conclude that they mean exactly the same thing, even though they are somewhat different in terms of word choice, sentence structure and formality.

But the problem is much more difficult when there are multiple translation alternatives that result in different meaning for the resulting translation. This may arise because the words in the source text are ambiguous or can mean more than one thing. The language of the source text might lack some syntactic or morphological distinction (like tense) that is required by the target language. Idioms in the source language may be hard to translate verbatim without losing their meaning entirely. And so on.

To deal with the issue of translation choices that result in a change of meaning in the translated text, the gold standard reference translations for the GALE Translation evaluation must include translation alternatives whenever the alternatives are not predictable. LDC experts (designated as "translation reviewers" in this document) will add these alternatives during the quality control phase of gold standard reference generation. This document provides guidelines for what kinds of alternatives should be included in the gold standard and what should not.

A basic rule of thumb is that alternatives must be added when one or more of the following conditions apply:

- the source text is ambiguous as to its meaning

-the multiple translation variants are not predictable for a native English speaker

The sections that follow describe several common categories where translation alternatives are likely to occur.

In the examples below, translation alternatives that should be included in the gold standard are indicated in **brackets and separated with double slash, i.e. [text // text], highlighted in yellow**, whereas alternatives that should not be included in the translation are represented in *italics, highlighted in blue*.

2 Discourse Alternations

In English, subjunctive or coordinative relationships between sentences are usually conveyed by explicit conjunction words. But in Chinese, discourse relations between sentences are inferred from context. During translation, translators may vary their choice of these "joining" words in ways that alter the meaning conveyed by the English translation. Translation reviewers should therefore add alternatives when they are not predictable, as in the following sentence:

1) 这位发言人说，他们的最终希望是和中国领导层根据谈判达成的基础，解决有关的问题，让西藏人拥有自由，维护他们认为是重要的文化认同。

The spokesperson said that their ultimate hope is to resolve the relevant issues on the basis reached through negotiation with the Chinese leadership [**so that//and**] Tibetans will have the freedom to preserve their cultural identity, which they consider important.

"so that" which implies some kind of causality, "and" is a simple coordinating conjunction. This choice causes a slight change in meaning which would not be predictable to an English speaker, so the alternatives should be represented.

The translation alternatives in Sentence 2 – *since* vs. *and...thus* – do not change the meaning, and therefore should not be included:

2) 他们的基础知识差，出现了一些错别字，无法纠正，结果落选。

They have poor basic knowledge, they wrote some wrong characters, and *since* there was no remedy, they lost in the selection.

They have poor basic knowledge, they wrote some wrong characters; *and* there was no remedy, *thus* they lost in the selection.

他们的意思，没有别的，求你好好看养大男；说大男就是他们，大男好，就等于他们没死。

Well, they want you to take good care of Big Baby, **because** he is them; **if** Big Baby is all right, it is as good as if they were alive themselves.

They just want you to take good care of Big Baby. **As** they said, Big Baby is them. Big Baby's well being is as good as if they were alive themselves.

面对严峻的国际经济环境，我们积极应对，趋利避害，变挑战为机遇，开创了对外开放的新局面。

Despite the grim international economic environment, we have created a new situation in our opening up by responding positively, striving to seek advantages and avoid disadvantages, and turning challenges into opportunities.

In the grim international economic environment, we ...

3 Syntactical Alternations

Chinese differs from English in grammar in that Chinese (a) has no past tense (b) has no definite article equivalent to "the" in English (c) has no plural form (d) does not always express grammatical subjects and copulas and (e) allows two or more verbs concatenated together. Due to these grammatical differences, it is possible that a Chinese source sentence can have more than one English translation which are not semantically equivalent. Annotators doing translation QC then need to provide extra acceptable translations which sound different to a native English speaker.

3.1 Tense

Tense in Chinese is not expressed anywhere in the verb or any auxiliaries, but only as adverbs of time, when needed for comprehension. A lot of times, the time is dropped from the Chinese source text, which makes tense ambiguous. Therefore, translation alternations need to be provided when the source text can be translated in different tenses.

达赖喇嘛的特使发现北京领导人、特别是中共中央统战部的代表比过去更愿意听取他们的看法。

The Dalai Lama's special envoys found that Beijing leaders, especially the representatives from the United Front Work Department of the CPC Central Committee, [**were//are**] more willing than before to hear their opinions.

他说这个想法是非常危险的，因为一旦达赖喇嘛不在，中国就失去了一个解决西藏问题的历史性机会。

He said that this idea is very dangerous because China [**will lose //would lose//will have lost**] a historic opportunity to solve the Tibetan issue once the Dalai Lama passes away.

3.2 Articles and Numbers

Chinese nouns do not have articles, such as "the" and "a" in English, nor do they have a plural form. So the word 大学 can any of the following mean:

a university
the university
the universities
universities

The context usually makes clear which is meant. There are still some cases that are ambiguous and annotators are expected to provide translation alternations for those ambiguous cases.

其开发人有权利选择放弃开发和重新提交评审会。

The [developer has//developers have] the right to choose to abandon the development and submit it again to the assessment committee.

新华社马尼拉 2 月 18 日电菲律宾地方官员 18 日说,菲东部 11 个村庄为防止再次发生类似山体滑坡事件而紧急撤离

Xinhua News Agency, Manila, Feb. 18 Philippine local officials said on the 18th that eleven villages in the eastern Philippines were urgently evacuated to prepare against [another similar landslide//a similar mudslide// other similar mountain landslides].

3.3 Pro-drop

There are often instances when who or what is being referred to can be inferred from context. In Chinese, the pronouns that in other languages would have those referents can be omitted. It is sometimes, however, ambiguous to resolve the referents that are dropped. Hence when translated into English, it is possible that the omitted pronouns can be translated into different pronouns. Annotators doing QC then need to provide those alternations.

关于今天晚上他们遗体到达具体将进行一些什么样的仪式,会有哪些人会出现,呃现在还不是非常的清楚。

Regarding what specific ceremony there will be when the remains arrive with them tonight, and who will show up, er, [it is still not very clear// we are still not very clear] at this time.

嘉日洛珠坚赞在华盛顿说,希望中国政府确实保护西藏独特的文化和传统。

Lodi Gyaltzen Gyari said in Washington that [he//they] hoped the Chinese government would indeed preserve Tibet's unique culture and tradition.

3.4 Serial verbs

Serial verb construction is a basic feature of Chinese grammar, in which two or more verbs are concatenated together.

达赖喇嘛特使重申不求西藏独立

Dalai Lama's special [envoys//envoy] reiterated [they do not//he does not] Seek Tibet's Independence.

4 Idioms, colloquial expressions and political expressions

For Idioms, colloquial expressions- both a literal translation and its meaning should be included. For political expressions, provide acceptable translation alternations.

4.1 Idioms and colloquial expressions

大家都看到跳了龙门的鲤鱼

Everyone sees [those who made it//the carp that jumped the dragon gate].

没有看到那 没 跳过龙门的。

No one sees [the those who didn't make it//those that didn't jump the dragon gate]

那么最近呢一说到伊朗核问题就说到四个字--充满变数。

Recently, whenever one mentions the Iran nuclear issue, one mentions [the phrase//four characters] - [full of variables//extremely unpredictable].

毕竟，打江山容易，守江山难。

After all, [it is easier to win the throne than to keep the throne//it is easy to become dominant, but it is difficult to maintain a dominant status].

中国的法治虽然不完善，但还不至于无法无天。

Although the rule of law in China is imperfect, it is not so bad as to [allow lawlessness//there is no law and no heaven above].

如果还想给他定罪或者还想采取进一步的流氓手段，我们决不会袖手旁观的。

If you even want to convict him or resort to further hooliganism, we will never [look on with folded arms// stand to the side and just watch].

一年在外挣的辛苦钱就这样打了水漂，

Their money earned by their own sweat over a whole year thereby [goes down the drain//is lost//is squandered//skips away so easily].

而这个“喜”不但能让农民从麻将桌上解脱出来，在文化活动中“动”起来，也为建设社会主义新农村添了砖加了瓦。

Furthermore, this "joy" will not only get peasants away from the mahjong table and get them "active" in cultural activities, but

also helps add contributions to [the construction// add "bricks and tiles" to the construction] of new socialist rural areas.

4.2 Political expressions

发展是硬道理

development is [the absolute principle// of overriding importance// fundamental principle]

民主党派

[democratic parties//non-communist parties]

5 Proper names

Provide alternations for people, organizations, events, etc. that can be referred to in English by more than one name, when the names are significantly different and are not predictable. However, it is NOT necessary to alternate minor spelling variations or the different terms used in different English speaking countries.

副总统吕秀莲质疑马英九这番谈话，是不是也要鼓励国民党员加入共产党？

Vice President [Annette Lu//Lu Xiulian//Lu Hsiu-Lien] questioned [Ma Yingjiu// Ma Ying-jeou] on this statement; isn't it going to encourage [Guomindang//Nationalist Party] members to join the Communist Party?

No alternations such as **Mao Tse-tung//Mao Zedong**, **soccer//football**, **Ministry of Health//Health Department** should be provided.

6 Words that can have multiple meanings in English

以张春祥大使为首的所有在场人员向三座棺木三鞠躬。

All people at the scene led by Ambassador Zhang Chunxiang bowed three times [to// in front of] the three coffins.

他说，没有任何国家或组织有权在巴勒斯坦人组建新政府问题上提出任何条件或要求巴勒斯坦人对以色列做出任何让步。

He said that no country or organization has the right to put forward any stipulations on the issue of the Palestinians forming a new government or [ask//require] the Palestinians to make any concessions to Israel.