

GALE Evaluation Reference Translation Alternation Guidelines for Arabic

Version 2 – 9/18/2009

Linguistic Data Consortium

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

Table of Contents

1	Introduction.....	3
2	Syntactical Alternations	4
2.1	Tense.....	4
2.2	Negative	5
2.3	Interrogative.....	5
2.4	Passive	5
2.5	Causative.....	6
3	Pronouns	6
4	Idioms, colloquial expressions	6
5	Proper names and aliases	7
5.1	Proper names	7
5.2	Aliases.....	7
6	Metonymy	7

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]**, whereas alternatives that should not be included in the translation are represented in *italics*.

2 Syntactical Alternations

2.1 Tense

In Arabic the imperfect present tense has one form, which is equivalent to both the present continuous and the simple present in English. Therefore it is possible that an Arabic source sentence can have more than one English translation that are not semantically equivalent. Annotators doing translation QC then need to provide extra acceptable translations to convey the possible meanings to a native English speaker.

واليوم يشهد الوطن العربي أحداثا مؤلمة تتمثل بالاحتلال والحصار وانتهاك حقوق الإنسان

And today the Arab world **[is witnessing //witnesses]** painful events such as occupation, siege, and human rights violations.

ونحن قلقون من الاسلحة التي يرسلونها الى العراق وما اذا كانوا سيواصلون ارسالها .

We are worried about the weapons that they **[send//are sending]** to Iraq, and whether they will continue to send them.

ينضم الفائزون في النهائيات السبت الى عشرات المتسابقين من نحو ثلاثين دولة في لندن لنهائيات المسابقة العالمية.

Winners of the finals **[join // will join]** dozens of contestants Saturday from around thirty countries in London for the international competition finals.

كشفت مصادر إعلامية إسرائيلية اليوم أن إسرائيل تعتزم بناء ألفي وحدة سكنية جديدة في المستوطنات اليهودية في الضفة الغربية خلال العام الحالي

Israeli media sources **[revealed // have revealed]** today that Israel intends to build two thousand new housing units in the Jewish settlements in the West Bank during the current year.

2.2 Negative

In Arabic, negations in the past tense have alternative English translations:

ما كنت لأرغب في أن أكتب عنك
I would [not want // not have wanted] to write about you.

وما كنت لاجاريك
Nor would I [go along // have gone along] with you.

Whereas a negation in the present has only one translation:

لا أريد ان أكتب عنك
I *do not want* to write about you.

The **conditional form** in Arabic using the particle قد (may/might), combined with a negation can be translated either in the affirmative or the negative form in English. Both are equivalent in meaning and it is not necessary to provide alternative translations.

قد لا تصل الحقيقة في بعض الاحيان الا مفلترة
Sometimes the truth *may only arrive when // may not arrive except* filtered.

2.3 Interrogative

The interrogative in Arabic can have two translations in English. The alternate translations must be provided, as in the example below.

هل يكون التعطيل تنفيذا لارادة شعبية؟
[Is disruption// will disruption be] a fulfillment of the popular will?

هل يحظى المشروع بموافقة المجلس؟
[Does// will] the plan meet with the approval of the cabinet?

2.4 Passive

In the case of the passive form in Arabic, which without vocalization, can be read as the simple past, there is no need to give alternate meanings when the meaning is clear from the context.

قتل أربعة من ركابها .
He killed four of its passengers.
Four passengers were killed.

In case that the past tense can be interpreted either as simple past or passive form in a given context, both translation need to be provided

حذر سيرجي لافروف من أن روسيا ستزد بشكل عملي على توسيع حلف شمال الأطلسي نحو الشرق
Sergey Lavrov [was warned//warned] that Russia would respond pragmatically to the eastward expansion of the North Atlantic Alliance.

2.5 Causative

The causative form of Arabic verbs can have two meanings in different contexts. When the context is not clear, the translation should provide both alternatives.

وبالتالي. إستبعاد المطالبة بسيادة عربية مستقبلا على هذه المدينة
And therefore [excluding// causing the exclusion of] future Arab sovereignty claims on this city.

3 Pronouns

There are often instances when the object or person is being referred to can be inferred from context. However, it is sometimes, difficult or even ambiguous to resolve the referents that are dropped. When translated into English, it is therefore possible that the omitted pronouns are to be translated into different pronouns or referents. Annotators in the QC stage then need to provide those alternations in ambiguous cases.

ومع قلة الدخل بدأ المواطن يتعلم سلوك سوق العقار ويشارك فيه ليصنع دخلاً إضافياً
With [a deficient income//their low incomes], citizens have begun to learn the behavior of the real estate market and to participate in it to create additional income

المخترع الأحساني منذر الحماد نال ميدالية الإبداع والاختراع نظير اختراع علمي فريد في مجال تصنيع المواد الذكية معقدة الأشكال بواسطة الليزر.
The Ahsa'i inventor Mundhir Al-Hammad was awarded the Medallion of Creativity and Innovation for [his// (no pronoun)] unique scientific invention in the field of laser manufacturing of complex formed smart materials.

4 Idioms, colloquial expressions

For Idioms and colloquial expressions, both literal translations and the intended meanings should be included.

رود لارسن: الشرق الاوسط "برميل بارود بفتيل مشتعل"
Roed-Larsen: Middle East ["is a powder keg with a flaming wick"// is extremely dangerous].

قبل أن ينبت الشعر في ذقنه
before [the first growth of his beard//he reaches puberty]

ولا بد لنا أن نقف أمام
And we have to [put an end to // stand up against]

، كما انه في طريقه إلى واشنطن لتسجيل براءة الاختراع في الولايات المتحدة الأمريكية
He is [on his way to apply // in the process of applying] for a patent
Washington, the United States of America.

5 Proper names and aliases

5.1 Proper names

Proper names in Arabic have literal meanings. If the context is not clear whether a word is a proper name or an adjective, both alternates should be given.

فقدت حسني في هذا الحادث.
I lost [Hosny// my beauty] in this accident.

ذهب برضى الى المطار
He [took Rida //went willingly] to the airport

However, it is not necessary to give alternate spelling of Arabic names such as Rafik al Hariri or Rafiq Hariri, King Abdallah or King Abdullah.

5.2 Aliases

Some personalities or important figures have well known aliases, notably the King of Saudi Arabia is known as the Custodian of the Two Holy Mosques. Both translations are equivalent and acceptable.

وصل خادم الحرمين الشريفين الى مطار الرياض
[The Custodian of the Two Holy Mosques// King Abdullah Bin Abdulaziz] arrived at Riyadh airport.

6 Metonymy

This figure of speech consists in using of the name of one object or concept for another object or concept to which it is related, or of which it is a part; both a literal translation and its meaning should be included.

عرق الطلبة وتعبيهم لا يذهب سدى.
Students' [hard work// sweat] and exhaustion do not go in vain.