

**Post Editing Guidelines
For
BOLT Machine Translation Evaluation**

Version 1.3
September 22, 2014

National Institute of Standards and Technology
Information Access Division, Multimedia Information Group

and

Linguistic Data Consortium

1	Introduction.....	3
1.1	Post editing process.....	3
2	Post Editor Role.....	4
2.1	Understanding Post Edit Alignment.....	4
3	Editing Rules.....	5
3.1	Capturing the Meaning of the Reference.....	6
3.1.1	Synonyms.....	6
3.1.2	Acronyms.....	6
3.1.3	Symbols.....	6
3.1.4	Numbers.....	7
3.1.5	Abbreviations.....	7
3.1.6	Contractions.....	7
3.1.7	Phrasal Ordering.....	7
3.1.8	Parts of Speech.....	8
3.1.9	Extra Information in MT Output.....	9
3.1.10	Information Missing from MT Output.....	10
3.2	Making the MT Understandable.....	10
3.2.1	Reference Ambiguity.....	11
3.2.2	Idioms and Internet Languages.....	11
3.2.3	Translation Alternatives.....	12
3.3	Minimizing the Number of Edits.....	12
3.3.1	Spelling.....	13
3.3.2	Proper Names.....	13
3.3.3	Systematic Duplicates.....	14
3.3.4	Dates.....	15
3.3.5	Decimal Points.....	15
3.4	Punctuation and Capitalization.....	15
3.4.1	Quotation Marks.....	16
3.4.2	Change/Insert/Delete Punctuation.....	16
3.4.3	Capitalization.....	17
3.5	Odd but Understandable English.....	18
3.6	Special markup.....	18
3.6.1	Unintelligible speech (()).....	18
3.6.2	Partial words ((pw)).....	18
3.6.3	Foreign speech <foreign lang=" " " " > </foreign>.....	19
4	Completely Incoherent MT System Translations.....	19
5	Appendix A.....	21

1 Introduction

Machine Translation (**MT**) is a technology that translates data from a source language into a target language. The source language for the BOLT program is either Arabic or Chinese, and the target language is English. These guidelines provide an overview of requirements for the BOLT program evaluation of machine translation technology. Genres evaluated in the BOLT program are discussion forum, chat and telephone conversations. These guidelines will cover discussion forums first, which will be the genre for Phase 1 of the BOLT program. A discussion forum is an online discussion site where people can hold conversations in the form of posted messages (Wikipedia.org). In a forum, a person posts a message on a particular topic and other people respond. Content in a forum is typically very informal and highly interactive. LDC collects data from various Arabic and Chinese forums and selects certain posts of threads for translation. A thread is one initiating message plus all of its follow-up messages in a forum, and a post is a single message posted by one user in a thread.

The BOLT program will evaluate MT in terms of the quality of the MT system translations. This is accomplished by measuring the **edit distance** between a **system translation** and a **gold standard reference translation**. The term “edit distance” refers to the minimum number of edits that one needs to make to the MT so it is as meaningful and understandable as the gold standard reference translation. The gold standard reference translation is prepared by professional human translators, and then carefully reviewed in three separate rounds of quality control by bilingual linguistic annotators. This is considered the *correct* translation.

This document serves as a reference for those tasked with post editing machine translation output for the BOLT evaluations. Section 2 of these guidelines describes the role of the editor. Section 3 is organized so that each main point addresses a single high level rule, followed with examples and in rare occasions, exceptions. A short summary of each high level rule can be found in Appendix A.

Post editors should read through these guidelines carefully before beginning the editing process, and refer to them as necessary. Appendix A was designed to serve as a one-page crib sheet.

1.1 Post editing process

Post editors work as independent contractors. To perform the post editing task, each editor downloads a “kit” from <https://secure ldc.upenn.edu/mtpe/bolt-p1/index.php>, which contains a specialized editing tool and a pair of MT and gold standard reference documents, to his or her personal computer. Kits are designed to be of comparable size and difficulty. Editors modify the MT output according to the guidelines outlined in the sections below.

Each kit is edited by a first-pass editor, and then is reviewed for completeness by a second-pass editor. Post editors are supplied with a web-based workflow management interface that permits them to check kits out and in. Supervisors are alerted to any new kit submissions and check the kits for quality. Kits that do not comply with the guidelines outlined in this document are returned to the post editors with feedback.

A training session is provided for each post editor, during which any questions about the process or the project goals are answered.

2 Post Editor Role

The job of the MT post editors may be summarized as follows:

“Make the MT output have the *correct meaning, using understandable English, in as few edits as possible.*”

The Post editor should keep in mind that:

- There are equally-accurate alternative ways to edit the machine translation output, all being considered correct, some using fewer edits than others.
- The reference translation is created by humans and may on occasion contain what appears to be an error. In such cases the editor should treat the reference as being absolutely correct (editors are not to “repair” the reference).
- The resulting edited version should have the same meaning as the reference translation; the edited version should not *add* any information that is *not* present in the reference, and the edited version should not *omit* any information that *is* in the reference.
- The editing occurs on a segment-to-segment basis to make it more manageable for the post editors, but it is more correct to view these rules as applying to the translated document as a whole. In particular, translation systems and your edited versions are free to rearrange words and phrases over long distances (translation is not a word-by-word or phrase-by-phrase mapping process).
- One goal is to minimize the number of edits whenever that can be done without compromising the meaning and without making the edited version hard to understand. Each inserted/deleted/modified word or punctuation mark counts as an edit. Moving an unmodified string of any number of words and punctuation marks by any distance (as if by cut-and-paste) counts as a single edit (shifting the location of that string).
- Editors should avoid wholesale copying of the reference.

2.1 Understanding Post Edit Alignment

As stated above:

“...editing occurs on a segment-to-segment basis to make it more manageable for the post editors, but it is more correct to view these rules as applying to the translated document as a whole. ...”

There are many variations of valid translations. It is perfectly acceptable for the machine translation output to place some phrases in different segments than where the reference has them. The system translations have been aligned to the reference translations according to segment boundaries which were created manually. There may be instances where the system translation segment being edited contains information which is in the preceding (or following) reference segment. This is okay.

The editor ***must not*** move the information just to make it align with the reference data. Table 1 shows differing alignments that do NOT require any editing.

<i>Reference Data</i>	<i>System Translation to Edit</i>
The President signed the Transportation bill on Tuesday morning, and flew to Camp David later that evening.	Being overdue for an extended vacation, the President flew to Camp David on Tuesday evening.
He was overdue for an extended vacation.	On Tuesday morning he had signed the Transportation bill.

Misalignment between the reference and the MT output segmentation could only affect a few words or a punctuation mark. Editors must not move the information in the MT output just to mirror the reference segmentation. Table 2 shows how the machine translation should be preserved even when the alignment of information differs from the reference:

<i>Reference Data</i>	<i>System Translation to Edit</i>	<i>Edited output</i>
After another three Chinese engineers were killed in Baluchistan province in Pakistan, a work group formed by the Chinese Foreign Ministry left for Karachi last night.	The progress made in the. The three Chinese engineers in Pakistan's Baluchistan province was killed after the attack, a Chinese Foreign Ministry led by the working group formed last night,	Since another three Chinese engineers in Pakistan's Baluchistan province were killed after the attack, a working group formed by the Chinese Foreign Ministry last night,
It will assist in making arrangements and dealing with resulting problems for the killed engineers.	I have already departed in Karachi. To help deal with the aftermath of the engineers.	departed for Karachi to help deal with the aftermath of the killed engineers.

3 Editing Rules

The goal of the BOLT MTpost-editing process is to take the machine translated output text drawn from discussion forum text, and to turn it into fluent English while preserving all of the meaning present in the original Arabic or Chinese text. Discussion forum data may be difficult to understand, and may contain non-standard punctuation and grammar, incomplete statements, emoticons, slang, and a range of topic content.

- Meaning

The edited machine translation must be faithful to the reference translation in both meaning and style. The translation should mirror the original meaning as much as possible while preserving grammaticality, fluency, and naturalness.

- Style

Try to maintain the same speaking style, tone or register as the reference translation. For example, if the reference translation is polite, the translation should maintain the same level of politeness. If the reference translation is rude, excited or angry, the edited machine translation should convey the same tone.

- Added or Missing Information

The edited MT translation should contain the exact meaning conveyed in the reference translation and should neither add nor delete information.

This section outlines the rules for each of the guidelines identified in Appendix A. The post editor must strictly adhere to these rules.

3.1 Capturing the Meaning of the Reference

Rule #1:
Make the MT output have the **same meaning** as the reference human translation. No more and no less.

Rule #1 represents the primary rule of post editing. Under no circumstances should the post editor sacrifice capturing the meaning of the reference translation in order to better meet the requirements of another rule. The English language allows the same information to be presented in several ways. The post editor should focus on capturing the meaning in the reference translation, not parroting the reference text word-for-word.

It is important to capture the complete meaning. At times the MT output will be missing information that will need to be added while other times the MT output will include extra information that must be removed. Consider the reference *always* to be correct.

3.1.1 Synonyms

If the meaning is correct, synonyms are completely acceptable.

3.1.2 Acronyms

Acronyms **can** be substituted for the expanded written form, and vice versa. For example, the following two terms are interchangeable:

Reference:BBC

MT output: British Broadcasting Corporation

3.1.3 Symbols

Symbols such as the dollar sign ‘\$’, percent symbol ‘%’ and the pound sign ‘#’ are equivalent to the written form and should not be edited.

The following example is fine as is, and need not be edited:

Reference: The company stock rose by 6 percent to \$22.50.

MT output: The company stock rose by six% to twenty-two dollars and fifty cents.

3.1.3.1 Emoticons

Emoticons such as :-) may appear in discussion forums. These should not be inserted or removed during the editing process.

3.1.4 Numbers

As a general rule of thumb, numbers may be either spelled out as in “one hundred and one” or written using numbers as in “101”. Do not change either form to the other. Do not insert or remove hyphens in forms like “twenty-two”.

3.1.4.1 Roman Numerals

Roman numerals are acceptable when they are used unambiguously, for example, to denote years or as markers in lists:

MT output: [...] ten great opportunities and challenges: [...] *Nine* elimination knowledge gap, to allow the use of information networks. **X** promote urban and rural areas, economic, social and natural efficient and coordinated development, building harmonious society.

Since “X” is being used as a number in a list of “ten great opportunities and challenges,” editors would *not* change the numeral during the editing process. On the other hand, something like the following would require clarification:

Reference: I wear a ten extra large.

MT output: I wear size X.

3.1.5 Abbreviations

Abbreviations are completely acceptable. “Mr.” for “Mister” and “lbs” for “pounds” are just two examples.

3.1.6 Contractions

Contractions are perfectly legitimate alternatives to the full spelling; “can’t” is equivalent to “cannot,” and vice versa.

3.1.7 Phrasal Ordering

The order of phrases only needs to be changed to correct the meaning or to follow other guidelines stated here. For instance, the MT in the following example is acceptable as is, and should not be edited:

Reference: Since the light was red, he stopped his car.

MT output: He stopped his car, since the light was red.

But the following example requires a reordering of the phrases in order to capture the meaning of the reference translation:

Reference: The protestors rallied against the government to change the outdated law.
MT output: The protestors will modify the outdated law to rally against the government.
Edited output: *The protestors rallied against the government to modify the outdated law.*

Phrasal reordering may improve **English Understandability** — see section 3.5.

3.1.8 Parts of Speech

It is important that post editors carefully consider differing parts of speech to be sure that the complete and correct meaning of the reference data is captured in the MT output and made readily understandable. Examples of how editors should treat various parts of speech follow in the sections below.

3.1.8.1 Verb Tense

The MT output must have the same verb tense as the reference when this makes a difference in the meaning. In the following example, the MT system produced a verb in the present tense; however, for the MT output to contain the same meaning as the reference, the verb should be changed to the past tense.

Reference: Chinese leaders *attached* a high degree of importance and paid a high degree of attention to this case [...]

MT output: The Chinese leaders. China *attaches* great importance to and concern.

Edited output: *The Chinese leaders attached to this case great importance and concern.*

Verb tense may affect English understandability; see section 3.5.

3.1.8.2 Prepositions

Prepositions may need to be changed, inserted or deleted, in order to correct the meaning.

Reference: The rebels looked at the hills *beyond* the road after the smoke cleared.

MT output: When the rebels looked up the hills *under* the road the smoke cleared.

Edited output: *the rebels looked at the hills beyond the road When the smoke cleared.*

In this example the prepositions “up” and “under” must be changed.

3.1.8.3 Adverbs and Adjectives

At times adverbial and adjectival constructions will need to be changed or added to relate the correct meaning. If the reference used the phrases “*lofty* goals” and “*dangerous* job”, the meanings conveyed by both lofty and dangerous must be included in the MT output.

Reference: Provocative Valentine Messages Cause Embarrassment: Six Recommendations for Texting Romantic Mobile Phone Messages

MT output: Valentine's Day ambiguous messages embarrassing phone transmission of are six 6

Edited output: *Valentine's Day **provocative** messages are embarrassing: 6 Recommendations for **romantic** mobile phone text transmission*

In some cases, the need to make an addition or change will be a judgment call by the post editor.

3.1.8.4 Determiners

Determiners are sometimes required to convey the complete meaning, as in:

Reference: ...representing *the* different historical periods of Egypt.

MT output: ...representing different historical periods of Egypt.

Subtle, but the determiner “*the*” is required to communicate the fact that **all** the historical periods are represented. Determiners in the MT output can also add incorrect meaning, as in:

Reference: The sources added that two other armed Palestinians were wounded in the clash ...

MT output: The sources added that *the* other Palestinians were wounded during the clash ...

Edited output: *The sources added that two other armed Palestinians were wounded during the clash ...*

The reference translation in this example is non-specific: it’s not clear that there were **only** two other armed Palestinians wounded. The editor therefore deleted the determiner “*the*” from the MT output to mirror the generic tone of the reference.

3.1.8.5 Pronouns

Pronouns should be correct and may be used in place of their corresponding noun, as long as the meaning is clear and correct, as in this example:

Reference: Walid Jumblatt Describes Syrian President as "Terrorist" [...] **Jumblatt** said, "The symbol of agency and subjugation to the Syrian regime is based in Baabda",

MT output: Walid Jumblatt, who described the Syrian president "terrorist" [...] **He** said "the symbol of employment and dependence of the Syrian regime abandoned in Baabda, "

Edited output: *Walid Jumblatt, described the Syrian president as "terrorist" [...] He said "the symbol of agency and dependence to the Syrian regime originates in Baabda, "*

3.1.9 Extra Information in MT Output

If it is the case that the MT output contains *information* that is not in the reference translation, delete it. In the following example “evening” should be deleted from the MT output:

Reference: The bomb went off on Tuesday.

MT output: The bomb exploded on Tuesday *evening*.

Edited output: *The bomb exploded on Tuesday.*

COMMENT: This was a simple example. The post editor was able to make a decision based on a single segment. It is the case however, that at times the post editor will have to consider segments *before* and *after* the one being edited in order to correctly determine if the information is indeed extra rather than simply moved from a different sentence/segment. (see alignment guidelines in **Section 2.1**)

The structure of the English language is very different from some languages. In English we might convey information by inflecting verbs or by the form of a noun or pronoun where other languages convey the same meaning by means of more verbose syntax.

MT output may contain extra words without including extra information. It is okay if the MT output has the *same* meaning as the reference but conveys that meaning in a more wordy fashion. For example it would be considered correct if the MT output had “five head of cattle”, where the reference had something like “five cattle”.

However, there are many wordy phrases that do convey some meaning. See the section that follows for guidelines on inserting information into MT output.

3.1.10 Information Missing from MT Output

This is the counterpart of the preceding rule. If some of the *meaning* of the reference is missing from the MT output, it must be added. The MT output is to be a complete translation.

Reference: The interim Iraqi government will hold its first election on January 30. The US has deployed nearly 20,000 additional troops to ensure the security of the *Iraqi* election.

MT output: The Iraqi interim government election is scheduled to be held on January 30. The United States has deployed some 20,000 troops to protect the security of the election.

Edited output: *The Iraqi interim government's first election is scheduled to be held on January 30. The United States has deployed some 20,000 additional troops to protect the security of the election.*

Note:

- (1) The fact that this was the “first” election was added.
- (2) The fact that these are “additional” troops was added
- (3) The last occurrence of “Iraqi” in the reference was redundant and not added.

Since post editors are editing the translation as a whole, not by each individual word or phrase, it is okay if the MT omits duplicated or redundant words and phrases when the meaning is *unambiguously complete and correct* (as was the case with the last occurrence of “Iraqi” in the preceding example)

Editing should never compromise the meaning. So, phrases such as “in the case of”, “for example”, “a number of”, and “on behalf of” should be added or removed to match the meaning of the reference whenever they make any difference in the meaning. ***When required to make an addition, use the shortest phrase (or single word) that conveys the same meaning.***

In the following example, the MT is acceptable ***only if***, because of other context, we know that I fly in an airplane, rather than a helicopter or whatever.

Reference: I fly in an airplane from Shanghai to Beijing.

MT output: I fly from Shanghai to Beijing.

3.2 Making the MT Understandable

Rule #2:
Make the MT output be ***as understandable*** as the reference.

Post editors should make the MT output as understandable as the reference translation, but must not try to make it more understandable.

Note: While every effort is being made to construct reference data that is accurate and fluent English, there is always the possibility of human error (spelling, punctuation, etc.) Always treat the reference translation as correct.

3.2.1 Reference Ambiguity

At times the reference data will be fluent, but not understandable. Consider the following example:

Reference: Journalists and the media were only allowed to tell the people what the data or government dictated that they say.

MT output: Journalists and men information, only to convey to the people the dictates of statements by the government.

There is not a clear distinction in the *reference* between “journalists” and “the media”. “The data” does not have a clear distinction either. Here is what three editors did for this example:

Post-editor1: *Journalists and the media were only allowed to convey to the people the dictates of the data or government.*

Post-editor2: *All that was available to journalists or media was to convey to the people the statements dictated by the data or government.*

Post-editor3: *Journalists and media were only allowed to convey to the people what the data or government dictated they say.*

Note that all three editors were scrupulous about the meaning. Note **carefully** that the editors did not make the MT output *more* understandable than the reference, because doing so might have involved some compromise about the meaning of the reference.

3.2.2 Idioms and Internet Languages

Idioms, colloquial expressions, and ways that language is used on the internet are particularly difficult to translate for human translators, let alone for machine translation engines. In the case of idioms, the reference translation provides both the intended meaning and literal meaning with the intended meaning highlighted and clickable. When you click at the intended meaning, the literal meaning will be displayed. For example:



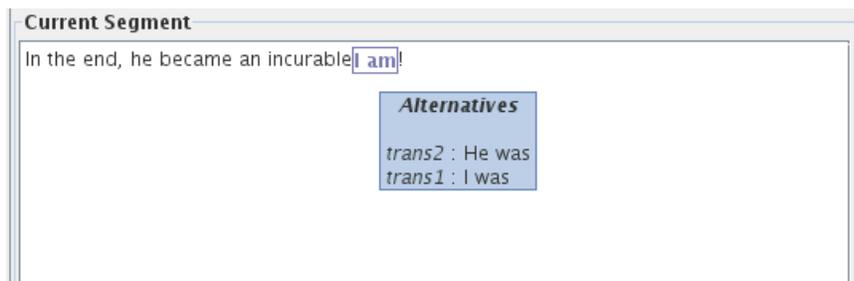
Intended meaning: hypocrite

Literal meaning: person with two faces

No edit is needed if the MT output has the same meaning as either the intended meaning or the literal meaning. If the MT output needs to be edited, modify the MT output **in as few edits as possible** to match the meaning of either the intended meaning or the literal meaning (whichever requires the fewest edits).

3.2.3 Translation Alternatives

Sometimes, the Arabic or Chinese source text is ambiguous which results in multiple good reference translations that are not predictable to native English speakers. Such multiple reference translations are provided as translation alternatives in the following format:



“I am”, “I was” and “He was” are equally good reference translation.

No edit is needed if the MT output has the same meaning as any of the translation alternatives. If the MT output needs to be edited, modify the MT output **in as few edits as possible** to match the meaning of any of the translation alternatives.

3.3 Minimizing the Number of Edits

Rule #3:

Capture the correct meaning in **as few edits as possible** using understandable English.

In case of perceived conflicts among the parts of this rule: correcting the meaning and making the output as understandable as the reference are more important than minimizing the number of edits.

Because minimizing the number of edits is a goal, when content from the reference is *missing* from the MT output, try to *insert* that missing reference content as the shortest replacement (fewest words) that you can think of with exactly the right meaning. Any deletion or insertion can make a significant difference in the edit score. Likewise, changing a verb construction unnecessarily can inflate the edit score. Consider the following:

Reference: He said, "We condemn the crimes of assassinations, incursions, arrests, the demolition of houses and the annexation of the Jordan Valley by imposing facts on the ground by force,"

MT output: He said that we condemn the crimes and unlawful arrests, the demolition of houses and the Jordan Valley to impose facts on the ground, by force.

*Edited MT 1: He said that we condemn the crimes and unlawful arrests, the demolition of houses and **the annexation of the Jordan Valley by imposing** facts on the ground, by force. (HTER = 13.158)¹*

*Edited MT 2: He said that we condemn the assassination crimes and unlawful arrests, the demolition of houses and **the Jordan Valley annexation to impose** facts on the ground, by force. (HTER = 5.714)*

Post editors should not exercise editorial license in altering the existing MT output. Do not delete or change perfectly acceptable words/phrases/punctuation that exist in the MT output to some substitute just because **you** have a preference for something else. Keep in mind that minimizing the number of edits is a goal. One way to think about minimizing edits is that the resulting edited version should preserve (completely unmodified) as much of the original MT output as possible without compromising the other guidelines; note that modifying the form of a word (for example, changing “flyer” to “flyers” or changing “radiates” to “radiating”) is no more desirable than substituting a completely different word.

Perhaps the most interesting challenge about doing your editing task is to be so clever that you find a way to fix the meaning with the minimum number of edits. This is a very important goal.

3.3.1 Spelling

Do not fix spelling variants in spelling such as “the colour grey” that reflect British usage. However, editors should correct misspelled words if they appear in the MT output.

3.3.2 Proper Names

Spelling variants are acceptable.

Sometimes alternate phrases will convey the same meaning as a proper name, for example: “North Korea” and “northern Korea” and “Democratic People’s Republic of Korea” and “People’s Democratic Republic of Korea” represent acceptable alternatives.

3.3.2.1 Proper names in Arabic source data

There are many ways of transliterating names (particularly Arabic names) into English. As a rule of thumb, when the system produces a phonetic transliteration of the proper name in the reference, it is acceptable.

The following examples show acceptable alternatives of Arabic proper names:

Reference:Emile Lahoud

MT output: Emil Lahud

Reference:Khaled Mashaal

MT output: Khalid Meshal

However, editors should take note of proper names that appear to be transliterations but are in fact different names all together:

¹ HTER stands for “Human-targeted Translation Error Rate”, and is the computation of the “edit distance” between the edited MT and the original MT.

Reference: Bassam Abdulmajid
MT output: Bassam Abdullah Al Majid
Edited output: *Bassam Abdul Majid*

Reference: William Brownfield, the American ambassador to Venezuela,
MT output: William Barwanfild, American Ambassador to Venezuela,
Edited output: *William **Brownfield**, American Ambassador to Venezuela,*

3.3.2.2 Proper names in Chinese source data

Editors approach Chinese proper names differently. Names may follow the Taiwanese or the Mainland transliteration system. The gold standard reference translation may contain the Mainland rendition of a name, while the MT output contains the Taiwanese representation of the same name, or vice versa.

Because post editors are not required to be familiar with the acceptable variations for proper names in the Chinese data, these variants will be noted as alternations in the reference text, as in:

Reference: Nanjing / Nan-ching / Nanking
MT output: Nanjing

Personal names may be ordered by family name or first name. Editors do not change the order of the name. For example, both of the following names are acceptable:

Reference: Yao Ming
MT output: Ming Yao

3.3.3 Systematic Duplicates

Phrases or expressions which are systematically repeated throughout a translated document are considered meaningful and should be present in the edited output. When the systematic repetition is missing from the MT output, editors reduce the phrase to the least possible number of edits, for example:

Reference: “rule number one, wipe your feet; rule number two, take off your hat; rule number three, choose a seat...”

In such cases, the editor retains the numerical markers, because they are meaningful, but he or she may reduce the phrase “rule number one” to “1”, and so on.

3.3.3.1 Arabic salutations

A common systematic repetition is an Arabic salutation, “peace be upon him,” which is often stated after the name of a prophet, or in some cases, after the name of an imam. If this phrase is missing in the MT output, editors should insert it with the least number of edits. In this case, an acronym is an acceptable synonym for the phrase, as in this example:

Reference: ... and the prophet, Allah’s blessing and peace be upon him
MT output: ... the Prophet, Allah’s blessing ...
Edited output: ... *the Prophet, Allah’s blessing and PBUH* ...

3.3.4 Dates

Dates must be correct. The date format in the MT output does not have to match the reference, but it must be correct and unambiguous. Dates may be represented in one of many numeral formats. For example, the following are all equivalent:

- 02/14/2006
- 2-14-2006
- 2.14.06

The European equivalent is acceptable when no confusion would arise as in 14/02/06.

Dates may be spelled out in part or in whole (abbreviations are completely acceptable). For example, the following are all equivalent:

- February 14th, 2006
- Feb. 14 '06
- Feb/14/2006
- 14/Feb/2006

In the following example no editing is required:

Reference: The bomb exploded on 7/21/1989.

MT output: The bomb detonated on the 21st of July, 1989.

3.3.5 Decimal Points

It may be the case that a comma is used in place of the decimal point. Post editors should correct this “European” format to the American format.

3.4 Punctuation and Capitalization

The BOLT program will tackle unstructured genres, so non-standard punctuation is expected to prevail. Discussion forums pose a lot of challenges in terms of punctuation as there will be a lot of unconventional punctuation or missing punctuation. For this new genre, editors should not standardize punctuation in the translation when it is missing from the reference translation. Non-standard punctuation (dashes, ellipses, etc.) should also be preserved in the MT output if they are in the reference translation. If the punctuation in the MT output is acceptable, leave it unchanged. Questions should end with a question mark.

It is important that the edited machine translation output have proper capitalization, both for proper nouns, and for proper sentence construction.

Rule #4:

Do not insert, delete, or change punctuation merely to follow traditional rules about what is “proper”. Allow missing sentence ending punctuation if it is a period; but fix question marks. Allow punctuation duplication in the MT output even if the reference translation only has one punctuation mark.

3.4.1 Quotation Marks

Single and double quotation marks must be properly paired and quote the correct material. At times, capturing the meaning may make it important to distinguish direct quotes (quotes with quotation marks), and the editor should use his/her best judgment about ensuring that. Note that sometimes quotation marks span more than the segment currently being edited and may require looking ahead to see where the quote ends.

Headlines, book titles, and so forth do not require quotation marks, but if quotation marks are present they must be properly paired and must quote the correct material.

3.4.2 Change/Insert/Delete Punctuation

Do not add/delete/change sentence-internal punctuation unless it is necessary for understandability or makes a difference in the meaning. The edited version does not have to match all the semicolons, colons, dashes, ellipsis dots (...) and commas in the reference, but it must be understandable.

If the reference has a question mark at the end of a question, the MT must also have a question mark at the end of the corresponding question.

In the following example, a question mark must be inserted in the MT output:

Reference: How can this be?

MT output: How can it be

Edited output: How can it be?

If the reference does **not** have a period at the end of a sentence, do not insert or delete a period in the MT. If the reference does have a period at the end of a sentence, edit the MT to correspond.

In the following example no editing is required for either of the MT outputs:

Reference: This is just purely insane

MT output: This is just purely insane.

Edited output: This is just purely insane

In the following example a period is required:

Reference: This is just purely insane.

MT output: This is just purely insane

Edited output: This is just purely insane.

Allow punctuation duplication in the MT output even if the reference translation has only one punctuation mark. In the following example no editing is required:

Reference: How can this be?

MT output: How can it be????

If the reference translation has duplicate punctuations, while the MT output has one punctuation, no editing is required, as the following example shows:

Reference: How can this be????

MT output: How can it be?

3.4.3 Capitalization

It is important that the edited machine translation output have proper capitalization, both for proper nouns, and for proper sentence construction. The first word of a sentence must be capitalized. When the system translation produces a lower-case word after sentence-final punctuation, editors should capitalize the word. This same rule applies to headlines, personal names, names of organizations and geo-political entities.

MT output: Gore’s lead lawyer **boyce** expressed the view that, [...]

Edited output: Gore's lead lawyer **Boyce** expressed the view that, [...]

At least the first letter of abbreviations and acronyms must be capitalized. The following variations are all acceptable forms of an acronym: CCTV, Cctv, C.C.T.V., C.c.t.v.

NOTE: When ambiguity could arise, for example when the MT output contains the word “Us” as an acronym for “United States”, the acronym must be changed to a commonly accepted capitalized acronym, such as “US” or “U.S.”.

Capitalization of job titles like "National Police Chief" is optional but must be consistent within a sentence. In the case of headlines, book titles, and so forth, there is no standard capitalization—only make sure that the first word of them is capitalized and that proper names are capitalized.

3.4.3.1 Inserting words and punctuation

When inserting new words at the beginning of a sentence or headline, capitalize the inserted word, but do not change the capitalization of the system translation. The following example treats the MT output correctly:

MT output: Successfully developed

Edited output: *Once Successfully developed await order to start launch.*

Along the same lines, when inserting sentence-final punctuation marks, editors do not modify the existing capitalization in the MT output, provided the first word after the punctuation mark was already in the MT, for example:

Reference: We will all be very happy. Since we are a family, lacking anyone is the same as missing a lot.

MT output: we will be very happy, **because** we are one family, were many who missing key,

Edited output: *We will be very happy. **because** we are one family, lacking one is the same as missing many.*

Do not change words that are capitalized in the middle of a sentence. For example, the following is acceptable:

MT output: Stanley Cohen Was born in New York in 1953.

3.4.3.2 Moving words in the MT output

Sometimes editors will need to rearrange the order of the MT output. When a word is *moved* to the beginning of the sentence, editors do NOT capitalize that word. Additionally, editors do not change the capitalization of words that are *moved* to the middle of the sentence, for example:

Reference: Good evening.

MT output: Evening good.

Edited output: good Evening.

3.5 Odd but Understandable English

The translation may be clearly understandable, yet odd. At times, what is odd may be the order of phrases, reflecting the structure of the foreign language in the source material. Here are some examples where the order of phrases reflects the structure of a foreign language that would normally put the topic first, when English would normally put the subject of the verb first. As long as the MT output is readily understandable (as in these examples, if you find them readily understandable), don't rearrange the phrases just to improve the fluency.

MT: This book I have read. **Reference:** I have read this book.

MT: In the yard is parked a car. **Reference:** A car is parked in the yard.

MT: Swimming, I am the best. **Reference:** I am the best at swimming.

MT: A dog have I. **Reference:** I have a dog.

Post editors will, however, usually need to insert missing verb tenses/inflections when the MT is not reasonably fluent English, in order to capture the meaning.

3.6 Special markup

3.6.1 Unintelligible speech (())

“Unclear” or “unintelligible” speech marks are inserted in the reference data during the audio transcription process for Conversational Telephone Speech. When a professional human transcriber cannot confidently transcribe an utterance, whether due to audio conditions such as background noise, distortion, or overlapping speakers, he or she marks the region of uncertainty with double parentheses (()). Editors should know that the MT may be very disfluent or odd at these points in a file. These regions will be dealt with separately during the scoring process (when the translation systems are scored) because they denote places of missing or “impoverished” information. Editors do not need to copy the double parentheses (()) into the edited output. These regions should be edited as normal to the best of the editors' ability.

3.6.2 Partial words %pw

Unscripted, spontaneous speech contains disfluencies such as partial words:

Di- di- did you know about th- that?

Partial words are carefully transcribed from the original audio recording; however, it is difficult to translate them, since the speaker's intent is not always known. Therefore, the human

translators insert %pw in place of each partial word. Editors do not treat these regions differently from any others, but recognize that the MT may contain odd or very disfluent language around partial words.

3.6.3 Foreign speech <foreign lang=" " " " </foreign>

This section is provided merely as explanation for the editor. Nothing in particular will be done by the editor about these occurrences. We include this explanation here in order to avoid distress should editors notice such occurrences.

At times words or phrases in a “foreign language” are spoken in a conversation (often English words or phrases). When an MT system is translating directly from speech audio, the system will have no outside information that the audio includes such foreign speech, and the system may well “misunderstand,” (although the system will naturally be free to notice that some words or phrases in the audio seem to be English or an unexpected language and treat such occurrences appropriately, if it can). The reference translation will, likewise, contain nothing flagging these occurrences, and editors will, therefore, have no opportunity to treat them as anything other than ordinary text, (and scoring will likewise treat these areas as ordinary text, so that scoring will reflect whatever a system was able to do with such occurrences).

When the MT system is translating Conversational Telephone Speech from a human-generated transcription, that source language transcription will contain a special tag surrounding the foreign words. For example:

<foreign lang="French"> Oui merci beaucoup. </foreign>

No such tags will, however, appear in the reference translation, and editors will therefore edit these occurrences as ordinary text (likely with no idea that something “different” has occurred), and scoring will reflect whatever the system was able to do with such occurrences.

4 Completely Incoherent MT System Translations

At times the MT system translation may appear as complete garbage. Avoid the temptation of erasing the MT system translation and copying in the reference translation; rather the post editor should try in earnest to salvage as much of the MT system translation as possible. Consider the following example:

Reference: The interim Iraqi government will hold its first election on January 30. The US has deployed nearly 20,000 additional troops to ensure the security of the *Iraqi* election.

MT output: Middle ground was found on the first day of January despite the US ability to send more aid for travel arrangements.

First DELETE: ~~Middle ground was found on the first day of January despite the US ability to send more aid for travel arrangements.~~

Second INSERT: The interim Iraqi government will hold its first election on January 30. The US has sent 20,000 more troops to ensure security for the Iraqi election.

The following second example shows two approaches to the same MT system translation. Both editors strove to retain as much of the original MT output as they could.

Reference: ...what happened is that when the building was on fire, you know that, of course, poisonous gases were released.

MT output: For the celebrations exempt the CIA at a later date has not yet been resolved. Is the kind of from Abu Dhabi in Sharm el-Sheikh will take some time. The killing. Mean for the demolition of the fire, which lead to the death of Mr. Inspectors.

Editor1, DELETE: ~~For the celebrations exempt the CIA at a later date has not yet been resolved. Is the kind of from Abu Dhabi in Sharm el-Sheikh will take some time. The killing. Mean for the demolition of the fire, which lead to the death of Mr. Inspectors.~~

Editor1, INSERT: When the building was on fire, you know that poisonous gases were released. (HTER - 184.6)

Editor2, DELETE: ~~For the celebrations exempt the CIA at a later date has not yet been resolved. Is the kind of from Abu Dhabi in Sharm el-Sheikh will take some time. The killing. Mean for the demolition of the fire, which lead to the death of Mr. Inspectors.~~

Editor2, INSERT: For you know the building was yet on fire, which led to the release of poisonous gases. (HTER - 161.5)

Both editors made a strong effort to salvage something from the original MT output, while retaining the meaning of the reference. The second editor was able to achieve this goal with fewer changes, so the second edit is preferable to the first, though it is a little more awkward than the first.

5 Appendix A

Appendix A contains the guidelines for editing the machine translation output (MT output). This is the set of rules each post editor should keep near them while they edit until they are committed to memory. *Minimizing the number of edits is a goal, but your primary goal is to make the meaning correct while making the edited version understandable.*

Make the MT output have the *correct meaning*, using *understandable* English, in as *few edits* as possible.

1. Make the MT output have **the same meaning** as the reference human translation. No more and no less.
2. Make the MT output be as **understandable** as the reference.
3. Capture the meaning in **as few edits as possible** using understandable English. If words/phrases/punctuation in the MT output are completely acceptable, use them (unmodified) rather than substituting something new and different.
4. Do not insert, delete, or change punctuation merely to follow traditional rules about what is “proper”. Allow missing sentence ending punctuation if it is a period; but fix question marks. Allow punctuation duplication in the MT output even if the reference translation only has one punctuation mark.