

Guía de traducción para el corpus de elicitación
Language Technologies Institute
Marzo 2006

Introducción

¡Hola equipo de traducción! Gracias por participar en esta tarea para aprender más sobre lenguas poco enseñadas (LPE). Podremos obtener información sobre la gramática de su idioma a través de sus traducciones. A este corpus se le conoce como corpus de elicitación; los lingüistas de campo han usado la elicitación durante años para aprender sobre lenguas nuevas o poco enseñadas. La elicitación tradicional es una colaboración entre una persona hablante de la LPE y una/a lingüista. El/la lingüista puede no hablar la LPE, así que se necesita a una persona bilingüe como usted, consultor/a de lengua, para comunicarse con el/la lingüista. El/la lingüista entrega unas oraciones para ser traducidas a la lengua del/la consultante y aprende los detalles sobre dicha lengua al analizar oraciones similares de manera continua.

Nuestro corpus se basa en esa metodología. En este caso traducirá una serie de oraciones en inglés, pero un ordenador y no un/a lingüista se las entregará. Así mismo, muchas oraciones contendrán vocabulario repetido y cada oración podría ser sólo progresivamente diferente de otra. Eso está pensado para estandarizar las oraciones elicítadas y para facilitar la comparación de oraciones a los/as lingüistas o a los programas informáticos. Como el ordenador no conoce su idioma, algunas de las oraciones parecerán redundantes o irrelevantes. Sin embargo, podrían ser relevantes para otro idioma.

Como nota final sobre esta guía, usaremos varias lenguas como el japonés o el español como ejemplos para mostrar la manera adecuada o errónea de traducir usando el corpus de elicitación. Es casi seguro que serán diferentes de la lengua que usted está traduciendo y ninguna de ellas es una LPE.

Traduciendo oraciones

A través del proceso de elicitación se les entregará una serie de oraciones y campos contextuales. Cada oración deberá ser considerada un enunciado independiente, lo cual quiere decir que podemos entregarles las oraciones elicítadas en cualquier orden y la traducción debería ser igual. En la Figura 1 tenemos dos oraciones no relacionadas. Si escucháramos esas oraciones consecutivamente en una conversación normal, interpretaríamos que la persona a la que Simon perdonó fue Frank. Sin embargo, en el corpus de elicitación esas dos oraciones deberían ser consideradas independientemente. Su traducción de 1a y de 1b debería ser igual sin importar el orden en el que aparezcan las oraciones y la información de una oración no se aplica a ninguna otra. Sería bueno pensar en Simon en 1a como una persona diferente y sin relación alguna al Simon de 1b, en una situación diferente y sin ninguna relación.

1 a. Oración: Frank golpeó a Simon.
Contexto: no corresponde

1 b. Oración: Frank golpeó a Simon.
Contexto: no corresponde

Figura 1: Cada oración es un enunciado independiente. Las oraciones **no** son una secuencia relacionada de sucesos. Estas dos oraciones deberían traducirse como sucesos separados, no relacionados. Es decir, la información en 1a no debería influir la traducción de 1b y viceversa.

Además, buscamos las traducciones más naturales y consistentes posibles. Queremos estudiar su lengua cómo se usa con espontaneidad. A pesar de ello, puede que haya algunas palabras o enunciados que no tengan una traducción exacta en su idioma. Por favor, use el equivalente más natural posible y úselo consistentemente. Las traducciones consistentes nos ayudarán a sacar conclusiones exactas sobre la gramática de la LPE en cuestión. Al mirar los ejemplos en la figura 2 podemos ver dos errores. El primero, mientras que la palabra japonesa ‘Jidouhanbaiki’ significa literalmente “Máquina de ventas automática” (tal y como se ve en el ejemplo 2a), la frase “máquina expendedora” usada en el 2b es mucho más natural. Para nuestros propósitos es mejor usar la traducción natural que la literal cuando en la lengua a la que se traduce existe una frase natural equivalente. “Máquina expendedora” contiene la misma información para un hispanohablante que “máquina de ventas automática”, pero es más clara.

En segundo lugar, ahora que ya sabemos que la mejor traducción es “Máquina expendedora”, se debería usar esta frase cada vez que aparece ‘Jidouhanbaiki’. Tener una traducción uniforme es importante a la hora de comparar oraciones y aprender sobre la gramática de su idioma.

2 a. Oración: Jidouhanbaiki ga kowareta.
Contexto: no corresponde
Traducción: La máquina de ventas automática se rompió.

2 b. Oración: Jidouhanbaiki go kowareru.
Contexto: no corresponde
Traducción: La máquina expendedora se romperá

Figura 2: Las palabras y frases deberían traducirse por la frase más natural disponible y debería ser traducida así consistentemente. El ejemplo de más arriba con las traducciones del japonés al inglés usa una traducción inconsistente en 2a y 2b y escoge un giro extraño en la frase en el 2 a.

Puede que en su idioma haya múltiples traducciones posibles que suenen naturales para un enunciado determinado. Sólo puede escoger una traducción para cada oración, así que debe escoger cuál es la que más probablemente usaría.

3 a

Oración: She is a teacher.
Contexto: no corresponde
Traducción: Ella es una Maestra.

3b

Oración: She is a teacher.
Contexto: no corresponde
Traducción: Es una Maestra.

Figura 3: Ambas traducciones pueden ser igualmente correctas para una determinada oración. La anterior oración en inglés se puede traducir de varias maneras en español, que es una lengua que tiende a omitir elementos, es decir: la oración se puede traducir con y sin sujeto. La decisión de qué oración se va a usar depende del/la traductor/a. Sin embargo, si deciden omitir el sujeto, deberían hacer lo mismo con todas las oraciones similares si pueden traducirse de tal manera. Lo mismo aplica cuando el/la traductor/a decide incluir el sujeto. La consistencia es importante.

Así mismo, como ya se ha mencionado anteriormente, nuestras frases en inglés usarán vocabulario o estructuras repetidas. A lo largo del corpus, algunas de las oraciones pueden parecer redundantes o irrelevantes, pero de todos modos podrían ser relevantes para otro idioma. Es totalmente aceptable traducir algunas oraciones de manera idéntica para su lengua. De nuevo, debemos promover la consistencia. En la figura 4 observamos que las mismas frases en inglés se traducen de manera diferente en español (4 a y 4b), pero igual en tagalo (4c y 4d). Si está traduciendo a una lengua que, como el tagalo, no distingue género, traduciría esas oraciones de la misma manera. Es posible declarar de manera explícita que el maestro en 4c es masculino, pero para nuestros propósitos eso no es lo ideal. Hacer eso cambiaría la estructura gramatical de la traducción en tagalo. De nuevo, hay que ser consistente con la manera en que se traduce cada enunciado.

4 a

Oración: She is a teacher.
Contexto: no corresponde
Traducción: Ella una Maestra.

4b

Oración: He is a teacher.
Contexto: no corresponde
Traducción: Él es un Maestro.

4c

Oración: She is a teacher.
Contexto: no corresponde
Traducción: Guro siya.

4d

Oración: He is a teacher.
Contexto: no corresponde
Traducción: Guro siya

Figura 4: Comparación entre dos grupos de oraciones traducidas al tagalo y al español. En las traducciones en español 4 a y 4b la diferencia de género es evidente en dos o tres sitios: el sujeto, el predicado nominal y en el dependiente del predicado nominal. En las traducciones en tagalo 4c y 4d no hay ninguna diferencia.

Como usar el campo contextual

El campo contextual está destinado únicamente a proveer información que no está codificada en la oración en inglés. Sin embargo, usted estará traduciendo únicamente la oración en español. Como este corpus no se diseñó con una lengua en mente, el campo contextual puede influir o no influir cómo traduce usted un enunciado específico. Por ejemplo, en la figura 5, dependiendo de cómo se marque el número en una LPE particular, las dos oraciones traducidas pueden ser idénticas. La traducción en árabe (5 a y 5b) incorpora la información del campo contextual, mientras que la traducción en español (5c y 5d) no lo necesita.

5 a.

Oración: You wrote.
Contexto: You= 5 hombres
Traducción: antum katabtum.

5b

Oración: You wrote.
Contexto: You= dos hombres
Traducción: antumaa katabtumaa.

5c

Oración: You wrote.
Contexto: You= cinco hombres
Traducción: escribieron.

5d

Oración: You wrote.
Contexto: You= dos hombres
Traducción: escribieron.

Figura 5: Las dos oraciones traducidas al árabe moderno estándar (5 a y 5b) se traducen de manera diferente dependiendo del número de personas que “You” representa. Sin embargo, las traducciones en español quedan igual en 5c y 5d.

Por lo que respecta a los campos contextuales, debería considerar cómo influye la información adicional a su traducción. No se sienta obligado/a a incluir en la traducción la información proporcionada en el campo contextual. Por ejemplo, en la figura 6 la fuente de información de

una oración puede influir en la traducción en algunos idiomas. Hay lenguas que traducen un enunciado de manera diferente dependiendo en si el/la hablante presencié el suceso o si lo supo por otra persona. Es importante considerar cómo se manifiesta esta diferencia en su idioma. ¿Traduciría de manera distinta la oración “Bill se comió el pastel” si el/la hablante presencié el suceso versus si alguien se lo contó? Si la respuesta es que sí, entonces debería incorporar la información en el campo contextual de su traducción. Si es que no, entonces ambas oraciones deberían traducirse de igual manera.

6 a.

Oración: Bill se comió el pastel.
Contexto: el/la hablante vio a Hill comerse el pastel.

6b.

Oración: Bill se comió el pastel.
Contexto: el/la hablante oyó que Hill se comió el pastel.

Figura 6: Dos oraciones diferentes por la fuente de información

Hablante vs. Oyente

Habrán referencias a “Hablante” y “Oyente” en este documento. Hablante es la persona que origina un enunciado y se puede considerar “primera persona”. Si no se indica género, el hablante se puede interpretar como un hombre. Sin embargo, tenga presente que el/la hablante cambiará de género a lo largo del corpus; esos cambios se indicarán en el campo contextual.

De la misma manera, el/la oyente es la entidad a la cual se dirige el enunciado. Es el equivalente de la segunda persona. El oyente puede interpretarse como un hombre a menos que se especifique de otra manera.

Hablante y oyente deberían considerarse como de igual rango y edad a menos que se especifique de otra manera. Para las lenguas en las cuales la solidaridad es importante, se debería considerar que ambos tienen la familiaridad de colegas de trabajo.