

Manual para la elaboración del Corpus de Paráfrasis

El Corpus de Paráfrasis estará constituido por una serie de textos que serán reescritos a partir de un texto base.

En esta ocasión, se realizará un tipo de paráfrasis que denominamos como paráfrasis baja, la cual se explicará más adelante.

Es importante que los textos finales tengan coherencia y sean legibles para todo lector humano.

TEXTO ORIGINAL

El texto principal es un artículo que trata sobre un tema específico; éste documento consta de 25 partes que llamaremos oraciones, cada una está separada por un salto de línea. En este nivel de paráfrasis, se alterará la estructura de las oraciones para lograr el objetivo. Las técnicas que se mencionarán están basadas en los artículos *A German Corpus for Similarity Detection Tasks (2014)* y *Detección automática de plagio: de la copia exacta a la paráfrasis (2010)*.

PARÁFRASIS BAJA

En este nivel, no habrá cambios en la longitud del texto, es decir, el resultado deberá tener las 25 oraciones que se presentan en el texto original. Todos los cambios se harán dentro de la misma oración. A continuación, se muestran algunos recursos que se pueden usar para realizar la paráfrasis; el círculo negro muestra la oración original y el círculo blanco, las variaciones:

- **Sinónimos / antónimos:** El cambio de palabras es factible mientras no se modifique el significado de la oración.

- El **proceso** que **se empleaba** en China era **muy novedoso**.
- La **técnica** que **se usaba** en China era en **extremo nueva**.
- El **proceso** **usado** en China **no** era **muy antiguo**.

Una sola palabra también puede ser reemplazada por más de un elemento.

- Este método de conservación de pescados enteros...
- Esta forma de hacer duraderos los pescados...

- **Cambios sintácticos:** Mover de lugar los elementos de la oración.

- Hoy en día es conocida **en Japón** con el nombre de shiokara...
- Actualmente se le conoce como shiokara **en Japón**...
- **En Japón**, es conocida como shiokara actualmente...

- **Elisión, adición y reordenamiento:** Algunos fragmentos o palabras dentro de la oración se pueden omitir y también se puede agregar nuevos elementos, siempre que la información extraída o agregada no altere, en gran medida, el significado original.

- una especie de salsa de pescado (similar al garum romano).
- un tipo de salsa de pescado.

La elisión es recomendable, principalmente, cuando se trata de una enumeración; en la cual, también sus componentes pueden cambiar de lugar.

- La fermentación proporciona diversos compuestos químicos que **aromatizan**, dan textura y **proporcionan sabor**.
- ... se fermenta y da varios productos químicos que dan **sabor** y **aroma**.

También, cuando un verbo tiene como complemento otro verbo, se puede omitir si el resultado es coherente. De igual forma, se puede hacer lo contrario y añadir un verbo al complemento.

- Las poblaciones de la costa no necesitaban aplicar estos métodos de conservación...
- Las poblaciones de la costa no necesitaban estos métodos de conservación...

- **Sustitución de un elemento por un actante:** Sustituir algún elemento que representa una acción por otro que represente al actor de la misma:

- Uno de los países que acepta japoneses es EEUU para trabajar en la **construcción** de vías ferroviarias.
- Uno de los países que acepta japoneses es EEUU para trabajar como **constructores** de vías ferroviarias.

- **Sustitución palabra-definición:** Sustituir una palabra por su definición o descripción:

- El pescado es un alimento **perecedero**.
- El pescado es un alimento **que se echa a perder rápidamente**.

- **Números:** Su escritura puede ser por medio de cifras (números arábigos o números romanos) o de texto, cuando su uso sea posible.

- casi cerca de cuarenta mil japoneses inmigran a Estados Unidos.
 - cerca de 40 mil japoneses se van a EEUU.
- En la época del s. XVII se inventaba el vinagre de arroz.
 - En el **siglo diecisiete**, a partir de la receta antes dicha, surgió el vinagre de arroz.

Como se puede notar del último ejemplo, las **abreviaturas** también son otro recurso a tomar en cuenta.

Cuando se habla de años o épocas, puede haber variantes.

- en 1910
 - a principios del siglo XX.
 - en la primera década del s. XX

- **Hiperónimos e hipónimos:** Esto se refiere a emplear un término que englobe al significado de la palabra en cuestión, es decir, que sea más general.

- uno de los países que acepta japoneses es EEUU para trabajar en la construcción de **trenes**.
 - en E.U. aceptaron a varias personas de Japón para que construyeran **medios de transporte**.

O, por el contrario, uno más específico.

- Después de ponerlo fuera de los **recipientes**...
 - tras sacarlo de las **ánforas**...

- **Transformación y segmentación:** Se pueden usar cambios estructurales o sintagmáticos dentro de las oraciones, tales como el uso de voz pasiva:

- El moho se dedica a consumir los hidratos de carbono existentes en el arroz...
- Los hidratos de carbono existentes en el arroz son consumidos por el moho.
- Los hidratos de carbono existentes en el arroz son consumidos.

Cuando se usa la voz pasiva, el actor puede ser omitido, como se observa en el ejemplo anterior.

Asimismo, la puntuación puede ser alterada:

- Esta preparación rápida del sushi es la que se conoce mayoritariamente en la actualidad y su preparación rápida la ha convertido en un elemento de fast food.
- Esta preparación rápida del sushi es la que se conoce mayoritariamente en la actualidad. Su preparación rápida la ha convertido en un elemento de fast food.

Hay que tomar en cuenta que el trabajo no está limitado a estos recursos, se puede usar cualquier otro que no se encuentre en este manual, siempre y cuando no haya cambios en la longitud del texto, como ya se mencionó. La longitud de las oraciones es independiente y sí pueden ser más largas o más cortas que la original. Tampoco es necesario usar absolutamente cada uno de estos recursos.

Por último, todas las oraciones deben tener algún tipo de cambio con respecto a la original, esto es, ninguna puede quedar inalterada.

El texto parafraseado es suficiente para la entrega, no hace falta escribir la oración original ni hacer algún tipo de comparación debido a que la estructura del texto debe tener el mismo orden que la original, así como se deben respetar los saltos de línea que indican el final de una oración.

Tampoco es necesario poner título ni que el archivo de texto tenga un nombre en específico. La tarea se puede realizar y entregar en el formato en el que más se acomoden.