

Règles de paraphrasage
Version 1.0, 11/III/2014
Juan Manuel Torres & Gerardo Sierra

Règles de la paraphrase basse (en bleu les phrases d'origine)

1. Garder la même idée en utilisant d'autres mots : utiliser les synonymes, changer le temps verbale, ajouter ou enlever quelques mots (mots de liaison ou autres) :

La fille est tombée dans la piscine
Cette fillette a chuté dans une piscine

2. Utiliser la voix passive, altération de la structure intra-phrased

Une fille est tombée dans la piscine et personne s'en est aperçu
Dans la piscine, sans que personne s'en est aperçu, une pauvre fillette est tombée

3. Description de concepts

La fille est tombée dans la piscine
La fille a tombée dans un réservoir d'eau utilisé pour nager dans les maisons

5. Altération de la structure entre-phrases

La fille est tombée jeudi dans la piscine.

...

La piscine est dangereuse.

Il faut l'interdire.

Cette piscine est très profonde et dangereuse.

Il faudrait carrément la fermer.

...

Hier soir, une fille est tombée dans cette piscine.

6. Ajout ou suppression de maximum 1 phrase

8. Les phrases ajoutées doivent appartenir à la **même thématique : termes en commun avec l'original (termes = noms)**

7. Interdit d'utiliser des phrases d'Internet (ou autres textes) sans les avoir préalablement paraphrasée.

8. Fichiers segmentés manuellement : une phrase par ligne, en utf8.

Note : les textes ne seront pas jugés par leur valeur de vérité ou fausseté

Règles de la paraphrase haute (en bleu les phrases d'origine)

0. Règles de la paraphrase basse

1. Fusion de phrases

Cette piscine est très profonde et dangereuse.
Il faudrait carrément l'interdire.

Puisque cette piscine est profonde et très dangereuse, je pense qu'il vaudrait mieux la fermer au public.

2. Segmentation de phrases :

La fille est tombée dans la piscine (un réservoir d'eau utilisé pour nager).

Une petite fillette est tombée dans un réservoir.
Or, ce réservoir est utilisé dans les maisons, notamment pour nager.
Le réservoir est appelé couramment « la piscine ».

3. Mélange de segments :

Cette piscine est très profonde et dangereuse.
Il faudrait carrément l'interdire.
Hier, une fille est tombée dans la piscine.

La piscine a toujours été profonde.
Il faut la fermer, car elle est devenue très dangereuse.
Une fille est tombée à l'eau.
Cela s'est passé hier, dans la piscine de ma maison.

4. Ajout ou suppression de **maximum 5 phrases ayant la même thématique : termes en commun avec l'original (termes = noms, substantifs)**

Fichier de Références

Il faut créer un fichier de références (original vs paraphrase) où, pour chaque ligne (phrase) du fichier d'origine, il faut mettre en correspondance (contenu sémantique) la ou les phrases de la paraphrase.

Document d'origine :

- 0 Cette piscine est très profonde et dangereuse.
- 1 Il faudrait carrément l'interdire.
- 2 Hier, une fille est tombée dans la piscine.
3. Je n'aime pas nager moi !

Document paraphrase :

- 0 Un réservoir est utilisé dans les maisons, notamment pour nager.
- 1 Le réservoir est appelé couramment « la piscine ».
- 2 Notre piscine a toujours été profonde.
- 3 Il faut la fermer, car elle est devenue très dangereuse.
- 4 Une fille est tombée à l'eau.
- 5 Cela s'est passé hier, dans la piscine de ma maison.

Fichier de références (txt utf8, autant de lignes que le fichier d'origine):

0 : 2 3
1 : 3
2 : 4 5
3 :