

## TP 2

Xin Chen

22105822

1. On peut déterminer le genre du groupe nominal en analysant son contexte, notamment s'il y a une anaphore pronominale (he/she, his/her, him) qui fait référence à cette entité nominale. Par exemple dans la phrase "patron X has finished his job" le pronom "his" fait référence à "patron X" ce qui est un indice que le patron est un homme. Cette association anaphorique entre le nom et le pronom nous permet de prédire le genre du groupe nominal en français.

2.

ii. Dans un groupe nominal il y a notamment deux constituances qui peuvent exprimer le genre: [DET] [NOM]

Cas a: le genre est marqué sur le déterminant et sur le nom;

Cas b: le genre est seulement marqué sur le nom;

Cas c : le genre féminin n'est pas du tout marqué. Le genre du groupe nominal est toujours masculin. Remarques: en français il y a des noms de professionnels dont le genre féminin est très sous-représenté; en revanche dans certaines métiers, le genre masculin est rare (eg. Sage-femme vs Sage-homme)

Cas e : le genre n'est pas explicité à cause de la liaison;

Cas d: le genre est seulement marqué sur le déterminant

Les différentes manières dont le genre est exprimé en français est ici présenté en tableau:

genre	<i>a</i>	<i>b</i>	<i>c</i>	<i>d</i>	<i>e</i>
feminin	la chanteuse	l'auditrice	?	l'artiste	la journaliste
masculin	le chanteur	l'auditeur	le charpentier	l'artiste	le journaliste

i. D'après l'analyse du genre en français la réponse à la première question est plus claire: dans la traduction du fr vers l'anglais on peut déterminer le genre du pronom associés à “work” en analysant le genre de l'antécédent à lequel le pronom fait référence. Et cela n'est pas toujours évident, par exemple dans le cas *d*, où le genre n'est pas du tout explicité.

3. i. Ce corpus est composé de phrases simples dont la structure est : [DET] + [NOM] + [PRED] + [POSS] + [NOM] . Cette structure syntaxique ne change pas lors de la traduction.

Et en anglais, le porteur du marqueur en genre est sur [POSS]; en français c'est sur le 1er groupe nominal et le possessif. çàd la position du marqueur en genre est fixée, ce qui facilite l'analyse.

Cette association est un à un, donc on peut modéliser le biais d'un système de traduction au moyen d'évaluer si cette association est correcte ou fautive, et calculer le taux d'erreur.

Par exemple pour une traduction de l'anglais vers le français:

ANG FR	vers	genre anglais	en	genre français	en	valeur
phrase 1		Masculin		Masculin		True
phrase 2		Feminin		Masculin		Fausse

4. i. Comme l'analyse sur le marqueur du genre en français on a fait ci dessus, un petit corpus avec très peu d'échantillons ne peut pas présenter cette variété en français. On a besoin d'un plus grand échantillon qu'on peut utiliser pour analyser toute sorte de situation.

ii. Dans le corpus, avec cette structure simple, on évite l'ambiguïté dans la structure syntaxique complexes lorsqu'il y a plusieurs antécédents dans le contexte. Par exemple dans la phrase "The patient made an appointment with the psychologist, but he canceled it shortly after because she had an emergency." Les antécédents et les anaphores pronominaux sont ambiguës.

## II

9. Si on découpe un problème en plusieurs sous-problèmes, ça nous permet de traiter ce problème d'une manière plus organisée et plus efficace, parce qu'il nous suffit de résoudre ces sous-problèmes un par un pour arriver au problème au départ. En plus, écrire des fonctions qui traitent un sous-problème à la fois, ça rend le code plus lisible donc lors du test, il est plus facile de détecter des bugs.

13. à partir du résultat de l'évaluation, on observe qu'en général la traduction automatique a un taux d'erreur très élevé pour la prédiction du genre, soit du français vers l'anglais; soit de l'anglais vers le français.

Pour un ensemble de 388 phrases, la traduction automatique de l'anglais vers le français a traduit la plupart de noms féminin en français en masculin(142 sur 194 noms féminin).

La traduction automatique du français vers l'anglais a fait beaucoup d'erreur en traduisant la plupart de noms masculin en féminin(59 sur 191 noms féminin); pourtant, on peut aussi observer que la traduction automatique dans ce sens a plus de précision que l'autre(srcFr-toAng) probablement parce que la machine a du mal à déterminer la relation anaphorique entre le nom et le pronom qui le suit lors de l'analyse de la phrase source. En revanche, quand le marqueur de genre est porté sur le nom ou son déterminant, c'est plus facile à associer la valeur de genre avec le nom(c'est dans la même syntagme.)

Quant aux noms épicènes, le système a toujours distingué entre masculin(16) et féminin(14) dans la traduction du français vers l'anglais parce que dans l'anglais il n'y a pas de mot épicène.

	srcFrench	srcEnglish
target French	388	208
target English	290	388

