

Production de paraphrases pour les systèmes vocaux humain-machine

Jonathan Chevelu

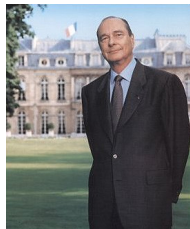
Thèse réalisée dans l'équipe Island du Greyc et à Orange Labs





Point de départ

- Quel est le lien entre ce qui est dit et celui qui le dit ?
- Comment choisir la formulation la plus adapté à une voix ?



Saperlipopette, mon repas s'est renversé sur le président de États-Unis !

Ho punaise, j'ai fait valdinguer ma bouffe sur le boss des Américains !

Comment produire une paraphrase pour améliorer un système vocale humain-machine ?

Définir

Systèmes vocaux

Paraphrases

Étudier

Évaluation classique

Production classique

Limites

Proposer

Une autre évaluation

Une autre production

- **Définir : Une tâche et un contexte**
Systèmes vocaux humain-machine
Paraphrases
- **Étudier : le cadre classique de la paraphrase**
Évaluation des paraphrases
Production de paraphrases
Les limites de la production et de l'évaluation classique
- **Proposer : un autre modèle plus adapté**
Une autre façon d'évaluer les paraphrases
Une autre façon de produire des paraphrases

Définir

Systèmes vocaux

Paraphrases

Étudier

Évaluation classique

Production classique

Limites

Proposer

Une autre évaluation

Une autre production

- **Définir : Une tâche et un contexte**
Systèmes vocaux humain-machine
Paraphrases
- **Étudier : le cadre classique de la paraphrase**
Évaluation des paraphrases
Production de paraphrases
Les limites de la production et de l'évaluation classique
- **Proposer : un autre modèle plus adapté**
Une autre façon d'évaluer les paraphrases
Une autre façon de produire des paraphrases

Définir

Systèmes vocaux

Paraphrases

Étudier

Évaluation classique

Production classique

Limites

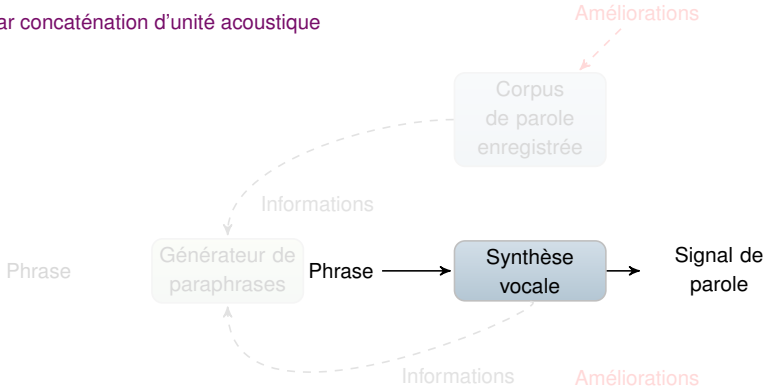
Proposer

Une autre évaluation

Une autre production

La synthèse vocale

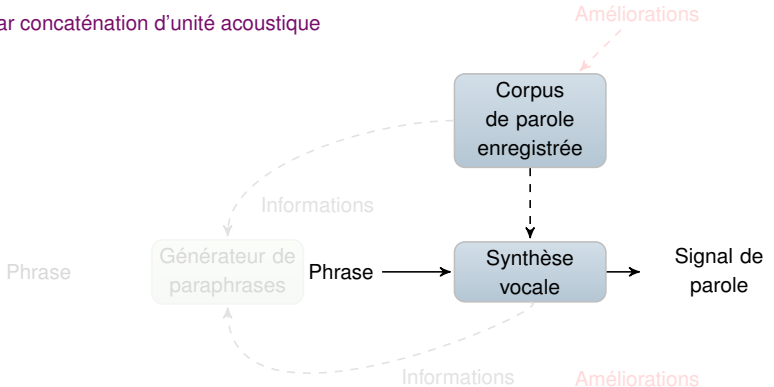
Par concaténation d'unité acoustique



- La qualité d'une voie dépend de l'adéquation entre le corpus de synthèse et les phrases à synthétiser.
- Le coût de développement d'une voie est dépendant de la taille du corpus de synthèse.

La synthèse vocale

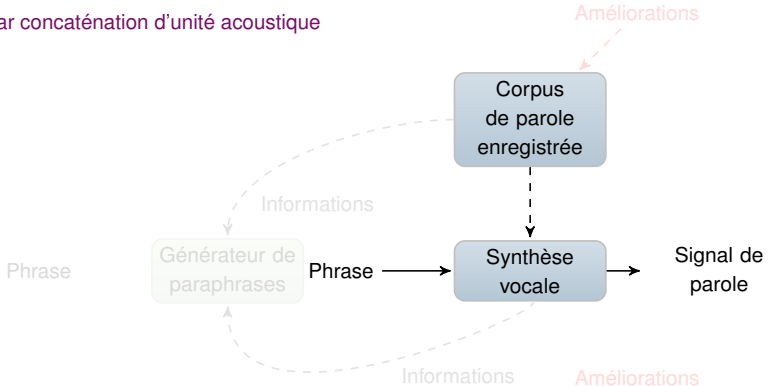
Par concaténation d'unité acoustique



- La qualité d'une voie dépend de l'adéquation entre le corpus de synthèse et les phrases à synthétiser.
- Le coût de développement d'une voie est dépendant de la taille du corpus de synthèse.

La synthèse vocale

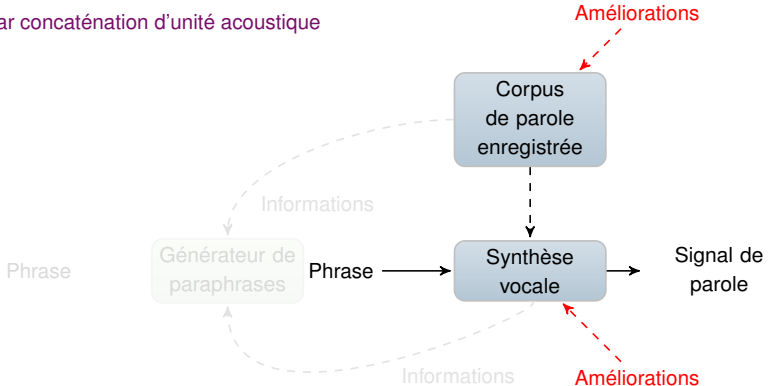
Par concaténation d'unité acoustique



- La qualité d'une voie dépend de l'adéquation entre le corpus de synthèse et les phrases à synthétiser.
- Le coût de développement d'une voie est dépendant de la taille du corpus de synthèse.

La synthèse vocale

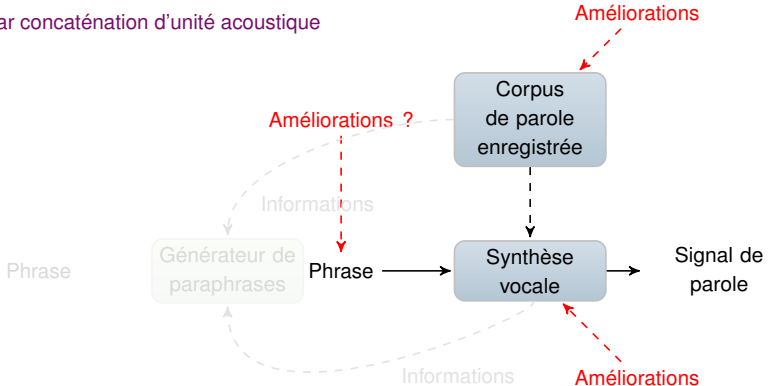
Par concaténation d'unité acoustique



- La qualité d'une voie dépend de l'adéquation entre le corpus de synthèse et les phrases à synthétiser.
- Le coût de développement d'une voie est dépendant de la taille du corpus de synthèse.

La synthèse vocale

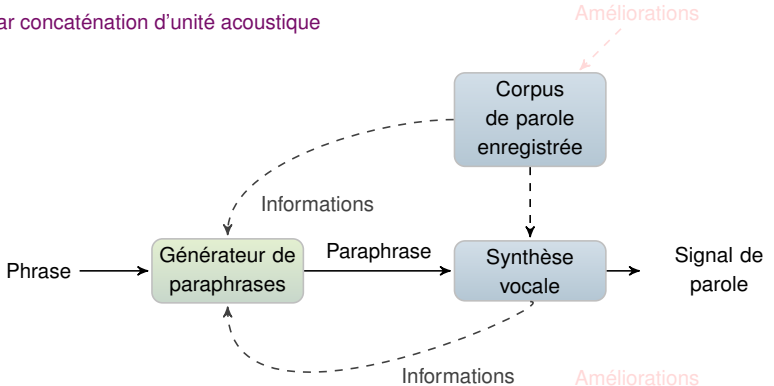
Par concaténation d'unité acoustique



- La qualité d'une voie dépend de l'adéquation entre le corpus de synthèse et les phrases à synthétiser.
- Le coût de développement d'une voie est dépendant de la taille du corpus de synthèse.

La synthèse vocale

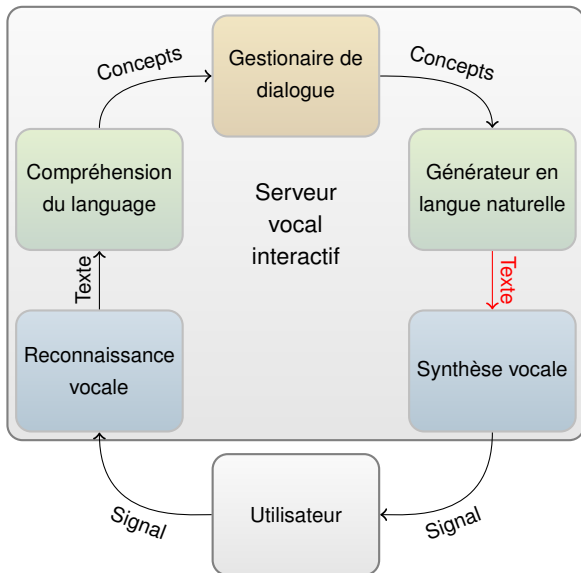
Par concaténation d'unité acoustique



- La qualité d'une voie dépend de l'adéquation entre le corpus de synthèse et les phrases à synthétiser.
- Le coût de développement d'une voie est dépendant de la taille du corpus de synthèse.



Serveurs vocaux interactifs



- Qualité acoustique
- Amorçage syntaxique
- Cohérence syntaxique
- Améliorer le dialogue

Définir

Systèmes vocaux

Paraphrases

Étudier

Évaluation classique

Production classique

Limites

Proposer

Une autre évaluation

Une autre production

Définir

Systèmes vocaux

Paraphrases

Étudier

Évaluation classique

Production classique

Limites

Proposer

Une autre évaluation

Une autre production

- **Définir : Une tâche et un contexte**
Systèmes vocaux humain-machine
Paraphrases
- **Étudier : le cadre classique de la paraphrase**
Évaluation des paraphrases
Production de paraphrases
Les limites de la production et de l'évaluation classique
- **Proposer : un autre modèle plus adapté**
Une autre façon d'évaluer les paraphrases
Une autre façon de produire des paraphrases



orange

Définitions

- Quatre notions clefs : phrase ; différentes ; sens ; ressemblant.

Définir

Systèmes vocaux

Paraphrases

Étudier

Évaluation classique

Production classique

Limites

Proposer

Une autre évaluation

Une autre production

Moyens **alternatifs** de transmettre la **même information** [Barzilay et McKeown, 2001]

Moyens **alternatifs** de transmettre le **même sens** [Zhao et coll, 2009]

Un ensemble de **phrases** qui exprime la **même chose** ou le **même évènement** [Sekine, 2005]

Opération de reformulation aboutissant à un énoncé contenant le **même signifié** (ou encore ayant une même structure profonde), mais dont le signifiant est **différent** [TLFi, 2010]

Générateur de paraphrase

Outil qui, à partir d'une **phrase**, produit au moins une **phrase différente** qui a un **sens ressemblant**.

- Paraphrases pour les systèmes vocaux :
 - [Kaji et coll., 2004], [Nakatsu et White, 2006], [Cahill et coll., 2009] ;
- Paraphrases pour remplacer une phrase :
 - [Knight et Marcu, 2000], [Nasr, 1996], [Max, 2008] ;
- Paraphrases pour améliorer des systèmes de traitement automatique de texte :
 - [Duclaye et coll., 2003], [Sekine, 2005], [Callison-Burch et coll., 2006] ;
- Paraphrases pour aider l'évaluation :
 - [Lepage et Denoual, 2003]



Production de paraphrases

- Approches nombreuses et dépendantes du type de ressources.

		Type de production		
		Schémas de transformations	Représentation sémantique	Approche statistique
		Spécifique	Emprunté au dialogue	Emprunté à la traduction statistique
Type de ressources	Bases de connaissances linguistiques	[McKeown, 1983] ; [Kaji et coll., 2002] ; [Hassan et coll., 2007]	[Nasr, 1996] ; [Fujita et coll., 2005] ; [Power et Scott, 2005]	
	Bi-corpus monolingue aligné	[Lin et Pantel, 2001]		[Quirk et coll., 2004] ; [Lepage et Denoual, 2005]
	Bi-corpus monolingue comparable	[Barzilay et Lee, 2003]		
	Bi-corpus bilingue aligné	[Max, 2008] ; [Zhao et coll., 2008]		[Bannard et Callison-Burch, 2005] ; [Zhao et coll., 2009] ; [Max, 2009]

Définir

Systèmes vocaux

Paraphrases

Étudier

Évaluation classique

Production classique

Limites

Proposer

Une autre évaluation

Une autre production

Définir

Systèmes vocaux

Paraphrases

Étudier

Évaluation classique

Production classique

Limites

Proposer

Une autre évaluation

Une autre production

- Problème peu traité
- Trois approches :
 - pré-filtrage [Zhao et coll., 2009]
 - post-sélection [Cahill et coll. 2009]
 - optimisation conjointe [Callison-Burch et coll. 2009]



Évaluation des paraphrases

- Pas de mesure automatique.
- Pas de consensus sur le protocole d'évaluation.

Définir

Systèmes vocaux

Paraphrases

Étudier

Évaluation classique

Production classique

Limites

Proposer

Une autre évaluation

Une autre production

	Nombre de juges	Nombre de questions	Taille de l'échelle	Originale identifiée	Synthèse des évaluations
Barzilay et Lee [2003]	2	1	2	Non	Non
Dolan et Brockett [2005]	2 + 1	1	2	Non	Oui
Callison-Burch [2007]	2	2	5	Oui	Oui
Zhao et coll. [2009]	2	2	3	?	Non
Zhao et coll. [2010]	2	2	5	?	Non

- Systèmes vocaux humain-machine :
 - le texte de la phrase à produire est une variable ;
 - adapter la phrase à une tâche.

- Paraphrase : phrase ; différentes ; sens ; ressemblant.

- Problématiques :
 - **production** ;
 - utilisation ;
 - **évaluation.**

Définir

Systèmes vocaux

Paraphrases

Étudier

Évaluation classique

Production classique

Limites

Proposer

Une autre évaluation

Une autre production

Définir

Systèmes vocaux
Paraphrases

Étudier

Évaluation classique
Production classique
Limites

Proposer

Une autre évaluation
Une autre production

- Définir : Une tâche et un contexte
Systèmes vocaux humain-machine
Paraphrases
- Étudier : le cadre classique de la paraphrase
Évaluation des paraphrases
Production de paraphrases
Les limites de la production et de l'évaluation classique
- Proposer : un autre modèle plus adapté
Une autre façon d'évaluer les paraphrases
Une autre façon de produire des paraphrases



orange

Réalisation

- Nous avons mis en œuvre une plateforme d'évaluation en analysant les différents choix de l'état de l'art.
- Objectif : simplifier les campagnes d'évaluations.

Eichier Édition Affichage Historique Marque-pages Outils Aide

http://pro.jonathan.chevelu.fr/reformulations/

Paraphrases: Accueil

PARAPHRASES

Identifiant :

Mot de passe :

Connexion

[S'enregistrer](#)
[Mot de passe perdu](#)

Accueil

Accueil

Présentation :

Bonjour, et bienvenue sur le site des reformulations. L'objectif de ce site est d'évaluer grâce à vous des reformulations de phrases afin de faire avancer la recherche. De plus, vous pourrez vous comparer aux autres évaluateurs afin de savoir si vous êtes le plus efficace et si votre jugement est le plus proche de l'opinion générale. Pour commencer il suffit de suivre les indications disponibles ci-dessous. Alors merci d'avance pour votre participation.

Comment commencer ?

1. La première étape consiste à vous enregistrer sur le site grâce au lien dans le menu de gauche ou en cliquant [ici](#).
2. Une fois votre inscription réalisée, vous devez valider votre compte grâce au lien présent dans l'e-mail qui vous aura été envoyé.

Terminé



orange

Réduire la charge cognitive des juges

- Décomposer une question complexe en deux questions plus simples.
- N'avoir que le nombre minimum de réponses possibles.
- Mettre en lumière les mots potentiellement problématiques.

Les deux phrases suivantes veulent-elles dire la même chose ?

"Le sauvage *chat* mange le *rongeur*. " "*Le chat* sauvage mange la *souris*. "

Oui

Non

Ne sais pas

La phrase suivante est-elle en bon français ?

Le sauvage chat mange le *rongeur*.

Oui

Non

Ne sais pas

- Interface la plus simple possible pour l'utilisateur :
≈ 45 secondes par question.

Définir

Systèmes vocaux

Paraphrases

Étudier

Évaluation classique

Production classique

Limites

Proposer

Une autre évaluation

Une autre production



orange

Des consignes et des exemples

● Harmonisation des stratégies de réponses.

Evaluation sémantique :

Vous allez commencer un exercice d'évaluation sémantique. Deux phrases vont vous être présentées et il vous sera demandé de répondre à une question par "Oui", "Non" ou "Ne sais pas".

Dans le but de vous aider, les différences entre les deux phrases seront marquées en rouge. De plus, des exemples seront présents en bas de la page pour vous guider vers les réponses attendues.

Il est conseillé de répondre relativement vite et en donnant votre première impression.

Il n'est pas demandé d'évaluer la grammaire des phrases mais si vous jugez qu'une des deux phrases présente des fautes qui empêchent sa compréhension alors vous devez répondre "Non".

Commencer l'évaluation

Evaluation syntaxique :

Vous allez commencer un exercice d'évaluation syntaxique. Une phrase va vous être présentée et il vous sera demandé de répondre à une question par "Oui", "Non" ou "Ne sais pas".

Dans le but de vous aider, des exemples seront présents en bas de la page pour vous guider vers les réponses attendues. De plus, les régions où des erreurs peuvent être présentes sont marquées en rouge. Attention tout de même, des erreurs (en particulier des mots manquants) peuvent ne pas être colorées.

Il est conseillé de répondre relativement vite et en donnant votre première impression.

Commencer l'évaluation

Par exemple, si les deux phrases sont :

Cliquez : Car :

Cette question est extrêmement importante !	Ce point est d'une importance capitale !	<input type="radio"/> Oui	Les deux phrases disent la même chose.
Je rentre tard.	Je rentre tôt.	<input type="radio"/> Non	Les deux phrases ne disent pas la même chose.
C'est un joli petit chat.	C'est un joli petit chat noir.	<input type="radio"/> Non	La seconde phrase est beaucoup plus précise car avec la première, on ne peut pas savoir que le chat est noir.
Je la trouve mignonne.	Je la trouve jolie.	<input type="radio"/> Oui	Même si les deux phrases ne sont pas identiques, elles disent "à peu près" la même chose.
Il verse de l'argent sur son compte.	Il remplit de l'argent sur son compte.	<input type="radio"/> Non	Dans ce contexte, "remplit" n'est pas synonyme de "verse".
Il me remplit un verre de vin.	Il me verse un verre de vin.	<input type="radio"/> Oui	Les phrases disent la même chose.

Par exemple, si la phrase est :

Cliquez : Car :

La vie est belle !	<input type="radio"/> Oui	Cette phrase est en bon français.
La vie est beaux.	<input type="radio"/> Non	Problème d'accord entre "vie" et "beaux".
Le ciel est beau. Et bleu	<input type="radio"/> Non	Le point est mal placé ou alors il manque un second point à la fin.
La vie) est belle " !!!	<input type="radio"/> Non	Il y a des symboles de ponctuations en trop.
Moi aime tomate.	<input type="radio"/> Non	La phrase n'est pas en bon français.
En conséquence, j' aime le bleu .	<input type="radio"/> Oui	Les règles de positionnement des espaces (en particulier pour la ponctuation) ne sont pas à prendre en compte.

Définir

Systèmes vocaux

Paraphrases

Étudier

Évaluation classique

Production classique

Limites

Proposer

Une autre évaluation

Une autre production

Définir

Systèmes vocaux
Paraphrases

Étudier

Évaluation classique
Production classique
Limites

Proposer

Une autre évaluation
Une autre production

- **Définir** : Une tâche et un contexte
Systèmes vocaux humain-machine
Paraphrases
- **Étudier** : le cadre classique de la paraphrase
Évaluation des paraphrases
Production de paraphrases
Les limites de la production et de l'évaluation classique
- **Proposer** : un autre modèle plus adapté
Une autre façon d'évaluer les paraphrases
Une autre façon de produire des paraphrases

Définir

Systèmes vocaux
Paraphrases

Étudier

Évaluation classique
Production classique

Limites

Proposer

Une autre évaluation
Une autre production

- **Nous avons mis en œuvre un générateur de paraphrases à partir des outils librement disponibles.**
- Paradigme utilisé production statistique de paraphrases :

Production de paraphrases = Traduction monolingue
↓

Traduction automatique statistique

Définir

Systèmes vocaux
Paraphrases

Étudier

Évaluation classique
Production classique

Limites

Proposer

Une autre évaluation
Une autre production

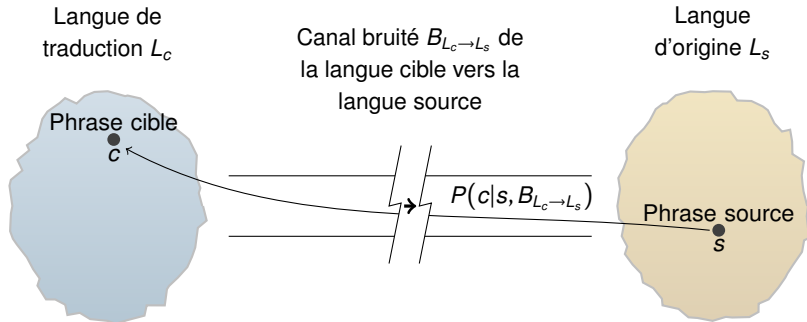
- **Nous avons mis en œuvre un générateur de paraphrases à partir des outils librement disponibles.**
- Paradigme utilisé production statistique de paraphrases :

Production de paraphrases = Traduction monolingue
↓

Traduction automatique statistique



Modèle du canal bruité utilisé en traduction automatique statistique



$$c^* = \arg \max_c P(c|s, B_{L_c \rightarrow L_s}) \quad (1)$$

$$c^* \approx \underbrace{\arg \max_c \max_l}_\text{Décodeur} \underbrace{P(c)}_\text{Modèle de langue} \prod_{i \in l} \underbrace{P(s'_i | c'_i, B_{L_c \rightarrow L_s})}_\text{Table de traduction} \quad (2)$$

- Définir
- Systèmes vocaux
- Paraphrases
- Étudier
- Évaluation classique
- Production classique**
- Limites
- Proposer
- Une autre évaluation
- Une autre production



orange

Composants du générateur de paraphrases

- Modèle de langue n -grammes :
 - prédit la qualité syntaxique d'une phrase ;
 - estime la probabilité d'un mot en fonction des mots qui le précède.
approche locale ;
 - SRILM [Stolcke, 2002].
- Table de paraphrases :
 - méthode par langue pivot [Bannard et Callison-Burch, 2005] ;
 - corpus : Europarl français ; **anglais comme langues pivots** ;
- Décodeur :
 - fondé sur l'algorithme de Viterbi [1967] ;
 - programmation dynamique : complexité linéaire en la longueur de la phrase ;
 - **contraintes de la fonction de score : incrémentale ; calculable avec un historique restreint** ;
 - Moses [Koehn et coll., 2007].

Définir

Systèmes vocaux

Paraphrases

Étudier

Évaluation classique

Production classique

Limites

Proposer

Une autre évaluation

Une autre production

Exemple de décodage

- Modèle de langue d'ordre 3 ; table de paraphrase d'ordre 2.

Le ||| Le ||| 1
sauvage ||| sauvage ||| 0,74
chat ||| chat ||| 0,74
sauvage chat ||| chat sauvage ||| 0,30
méchant chat ||| chat sauvage ||| 0,15
mange ||| mange ||| 0,5
mange ||| dévore ||| 0,5



Mot 1
Le



Mot 2
chat



Mot 3
sauvage



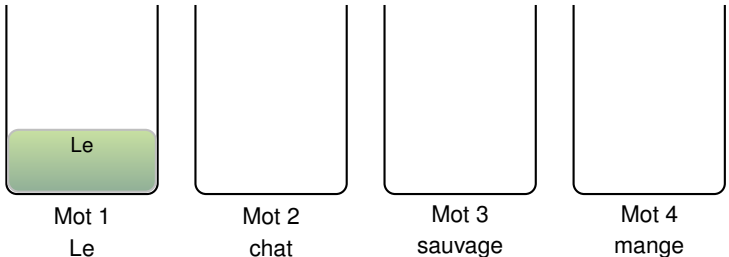
Mot 4
mange

Exemple de décodage

- Construction des hypothèses de traduction.

Le ||| Le ||| 1
sauvage ||| sauvage ||| 0,74
chat ||| chat ||| 0,74
sauvage chat ||| chat sauvage ||| 0,30
méchant chat ||| chat sauvage ||| 0,15

mange ||| mange ||| 0,5
mange ||| dévore ||| 0,5

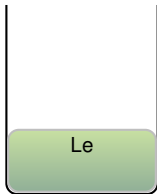




Exemple de décodage

- Deux mots conservés pour le modèle de langue.

Le ||| Le ||| 1
sauvage ||| sauvage ||| 0,74
chat ||| chat ||| 0,74
sauvage chat ||| chat sauvage ||| 0,30
méchant chat ||| chat sauvage ||| 0,15
mange ||| mange ||| 0,5
mange ||| dévore ||| 0,5



Mot 1
Le



Mot 2
chat



Mot 3
sauvage



Mot 4
mange

Définir

Systèmes vocaux
Paraphrases

Étudier

Évaluation classique
Production classique
Limites

Proposer

Une autre évaluation
Une autre production



orange

Exemple de décodage

- Étendre les hypothèses de la pile 2.

Le ||| Le ||| 1

sauvage ||| sauvage ||| 0,74

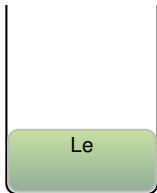
chat ||| chat ||| 0,74

sauvage chat ||| chat sauvage ||| 0,30

méchant chat ||| chat sauvage ||| 0,15

mange ||| mange ||| 0,5

mange ||| dévore ||| 0,5



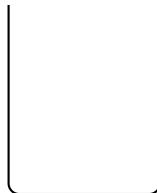
Mot 1
Le



Mot 2
chat



Mot 3
sauvage



Mot 4
mange

Définir

Systèmes vocaux

Paraphrases

Étudier

Évaluation classique

Production classique

Limites

Proposer

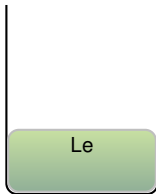
Une autre évaluation

Une autre production

Exemple de décodage

- Étendre les hypothèses de la pile 1.

Le ||| Le ||| 1
 sauvage ||| sauvage ||| 0,74
 chat ||| chat ||| 0,74
 sauvage chat ||| chat sauvage ||| 0,30
 méchant chat ||| chat sauvage ||| 0,15
 mange ||| mange ||| 0,5
 mange ||| dévore ||| 0,5



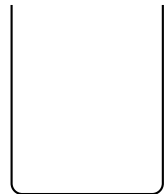
Mot 1
Le



Mot 2
chat



Mot 3
sauvage

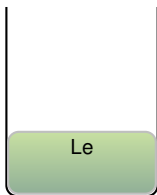


Mot 4
mange

Exemple de décodage

- Calcul du score des états.

Le ||| Le ||| 1
sauvage ||| sauvage ||| 0,74
chat ||| chat ||| 0,74
sauvage chat ||| chat sauvage ||| 0,30
méchant chat ||| chat sauvage ||| 0,15
mange ||| mange ||| 0,5
mange ||| dévore ||| 0,5



Mot 1
Le



Mot 2
chat



Mot 3
sauvage

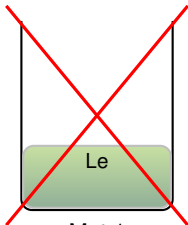


Mot 4
mange

Exemple de décodage

- Supprimer l'historique inutile.

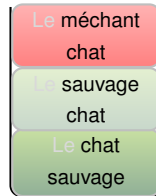
Le ||| Le ||| 1
 sauvage ||| sauvage ||| 0,74
 chat ||| chat ||| 0,74
 sauvage chat ||| chat sauvage ||| 0,30
 méchant chat ||| chat sauvage ||| 0,15
 mange ||| mange ||| 0,5
 mange ||| dévore ||| 0,5



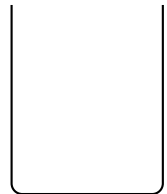
Mot 1
Le



Mot 2
chat



Mot 3
sauvage

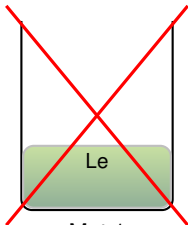


Mot 4
mange

Exemple de décodage

- Étendre les hypothèses de la pile 2 et 3.

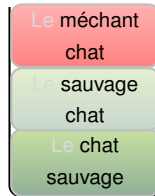
Le ||| Le ||| 1
 sauvage ||| sauvage ||| 0,74
 chat ||| chat ||| 0,74
 sauvage chat ||| chat sauvage ||| 0,30
 méchant chat ||| chat sauvage ||| 0,15
 mange ||| mange ||| 0,5
 mange ||| dévore ||| 0,5



Mot 1
Le



Mot 2
chat



Mot 3
sauvage

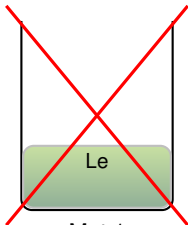


Mot 4
mange

Exemple de décodage

- Calcul du score des états.

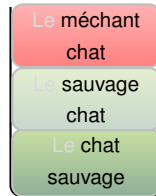
Le ||| Le ||| 1
 sauvage ||| sauvage ||| 0,74
 chat ||| chat ||| 0,74
 sauvage chat ||| chat sauvage ||| 0,30
 méchant chat ||| chat sauvage ||| 0,15
 mange ||| mange ||| 0,5
 mange ||| dévore ||| 0,5



Mot 1
Le



Mot 2
chat



Mot 3
sauvage

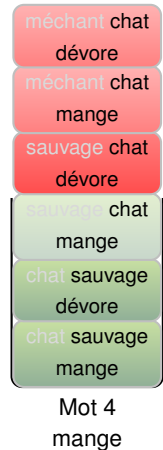
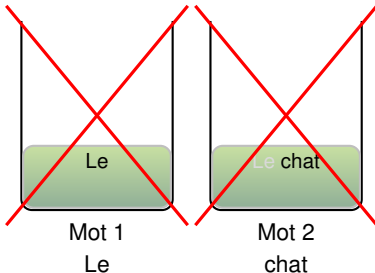


Mot 4
mange

Exemple de décodage

- Supprimer l'historique inutile.

Le ||| Le ||| 1
 sauvage ||| sauvage ||| 0,74
 chat ||| chat ||| 0,74
 sauvage chat ||| chat sauvage ||| 0,30
 méchant chat ||| chat sauvage ||| 0,15
 mange ||| mange ||| 0,5
 mange ||| dévore ||| 0,5

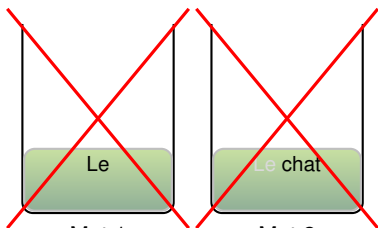




Exemple de décodage

- Fusion des états identique : conservation du meilleur chemin.

Le ||| Le ||| 1
sauvage ||| sauvage ||| 0,74
chat ||| chat ||| 0,74
sauvage chat ||| chat sauvage ||| 0,30
méchant chat ||| chat sauvage ||| 0,15
mange ||| mange ||| 0,5
mange ||| dévore ||| 0,5

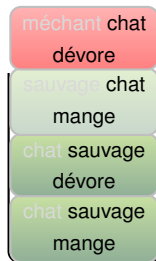


Mot 1
Le

Mot 2
chat



Mot 3
sauvage

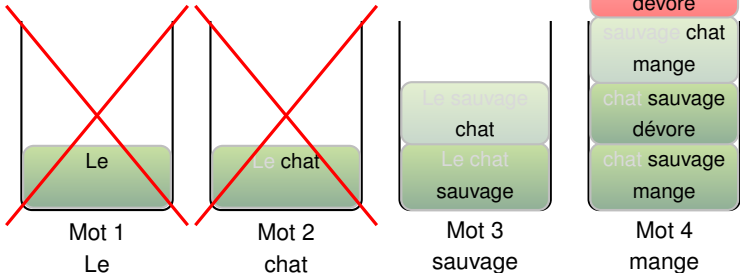


Mot 4
mange

Exemple de décodage

- Complexité au pire : $\mathcal{O}(N \times V^n \times L)$.

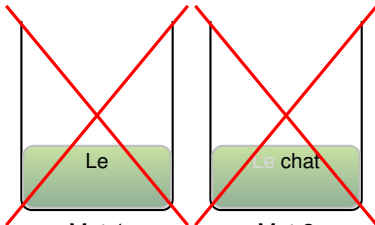
Le ||| Le ||| 1
 sauvage ||| sauvage ||| 0,74
 chat ||| chat ||| 0,74
 sauvage chat ||| chat sauvage ||| 0,30
 méchant chat ||| chat sauvage ||| 0,15
 mange ||| mange ||| 0,5
 mange ||| dévore ||| 0,5



Exemple de décodage

- Recherche par faisceau : limitation de la taille de la pile.

Le ||| Le ||| 1
 sauvage ||| sauvage ||| 0,74
 chat ||| chat ||| 0,74
 sauvage chat ||| chat sauvage ||| 0,30
 méchant chat ||| chat sauvage ||| 0,15
 mange ||| mange ||| 0,5
 mange ||| dévore ||| 0,5

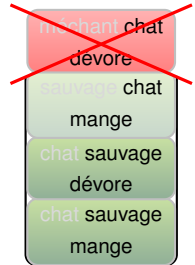


Mot 1
Le

Mot 2
chat



Mot 3
sauvage



Mot 4
mange

- Définir : Une tâche et un contexte
Systèmes vocaux humain-machine
Paraphrases
- Étudier : le cadre classique de la paraphrase
Évaluation des paraphrases
Production de paraphrases
Les limites de la production et de l'évaluation classique
- Proposer : un autre modèle plus adapté
Une autre façon d'évaluer les paraphrases
Une autre façon de produire des paraphrases

Définir

Systèmes vocaux

Paraphrases

Étudier

Évaluation classique

Production classique

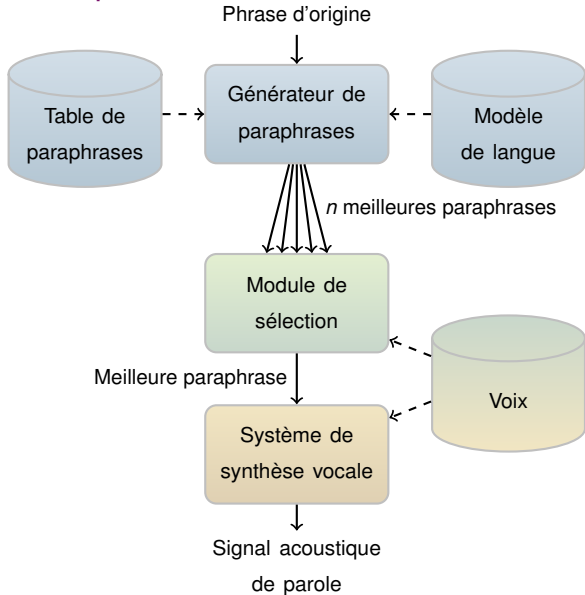
Limites

Proposer

Une autre évaluation

Une autre production

Système complet



Définir

Systèmes vocaux

Paraphrases

Étudier

Évaluation classique

Production classique

Limites

Proposer

Une autre évaluation

Une autre production



Performances intermédiaires

	Syntaxe correcte		Sens préservé		Paraphrase correcte	
	non	oui	non	oui	non	oui
non	49	8	48	10	65	11
oui	3	40	10	32	2	22
Kappa	0,69 (p -valeur $< 10^{-3}$) Accord substantiel					

- **Nous avons montré que les performances sont trop faibles pour une utilisation industrielle.**



orange

Utilisation d'un synonyme :

- O : La Commission soutient activement la mise au point au niveau mondial et de façon **uniformisée** d'un cycle de tests des motocycles.
- P : La Commission soutient activement la mise au point au niveau mondial et de façon **standardisée** d'un cycle de tests des motocycles.

Définir

Systèmes vocaux

Paraphrases

Étudier

Évaluation classique

Production classique

Limites

Proposer

Une autre évaluation

Une autre production

Suppression d'une virgule :

- O : Mais la Commission, non plus, n'a pas joué son rôle et **cela est** grave, Monsieur le Président.
- P : Mais la Commission non plus, n'a pas joué son rôle et **c'est un problème** grave, Monsieur le Président.

Plusieurs modifications en continu :

- O : Je **voudrais également mettre en avant une préoccupation liée** à la question de l'immigration.
- P : Je **souligne d'ailleurs l'inquiétude quant** à la question de l'immigration.



orange

Définir

Systèmes vocaux
Paraphrases

Étudier

Évaluation classique
Production classique

Limites

Proposer

Une autre évaluation
Une autre production

Problème d'accord en genre et en nombre :

- O : Et, quoi qu'il en soit, cela **exige** un examen **sectoriel** préalable.
- P : * Et, quoi qu'il en soit, cela **requiert** un examen **sectorielles** préalable.

Problème de dépendance à longue distance :

- O : Il est tombé gravement malade et n'a personne pour prendre soin de lui.
- P : * Il est tombé gravement malade, et **elle** n'a personne pour prendre soin de lui.

- **Nous avons montré les limites du modèle de langue.**

Problème de contexte inconnu :

- O : Pour des raisons pratiques, je considère donc que la proposition **de M.** Fayot est bonne.
- P : * Pour des raisons pratiques, je considère donc que la proposition **du président** Fayot est bonne.



Performances finales : évaluation linguistique

		Syntaxe correcte		Sens préservé		Paraphrase correcte	
		non	oui	non	oui	non	oui
non	60	1	39	4	62	1	
oui	3	36	16	41	7	30	
kappa	0,76 (p -valeur $< 10^{-3}$) Accord substantiel						

Définir

Systèmes vocaux

Paraphrases

Étudier

Évaluation classique

Production classique

Limites

Proposer

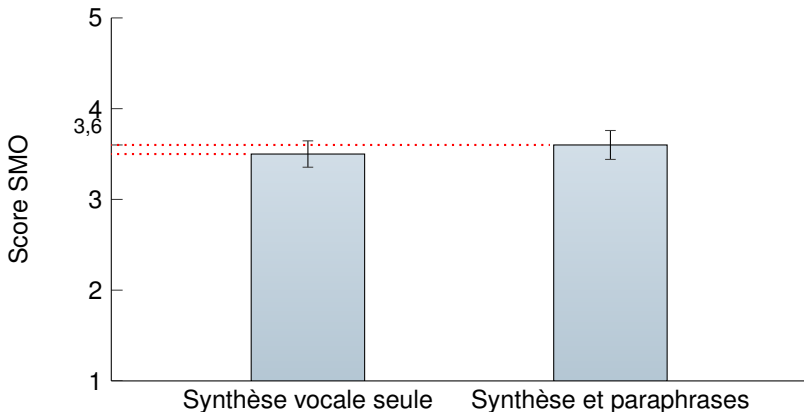
Une autre évaluation

Une autre production



Évaluation acoustique

- Score de synthèse amélioré de $15 \pm 20\%$



- **Nous avons montré le besoin de prendre en compte la tâche lors de la production.**



orange

Limite du décodeur : Problématiques

Définir

Systèmes vocaux
Paraphrases

Étudier

Évaluation classique
Production classique

Limites

Proposer

Une autre évaluation
Une autre production

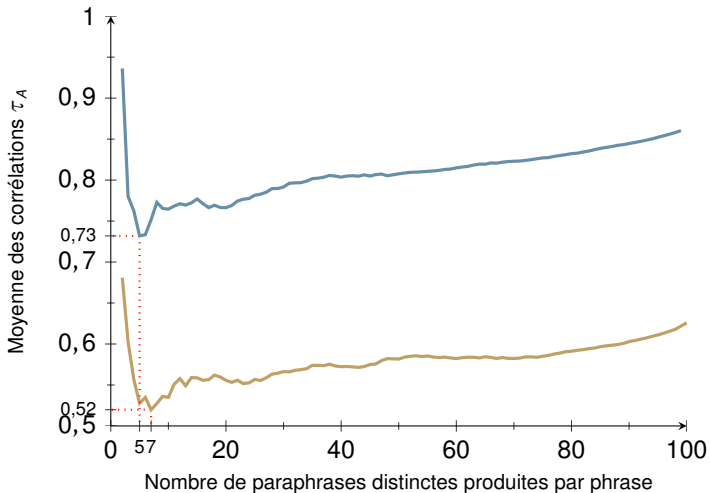
- Plusieurs heuristiques sont introduite dans le modèle de la paraphrase pour pouvoir utiliser un décodeur de type Viterbi :
 - recherche par faisceaux : ne teste pas tous les découpages d'une paraphrase pour calculer sont score.
 - approximation de la somme sur tous les découpage par un maximum.
- Mesure du coût de ces heuristiques en comparant avec l'ordre optimal *a posteriori* des meilleurs paraphrases de Moses.



orange

Résultats

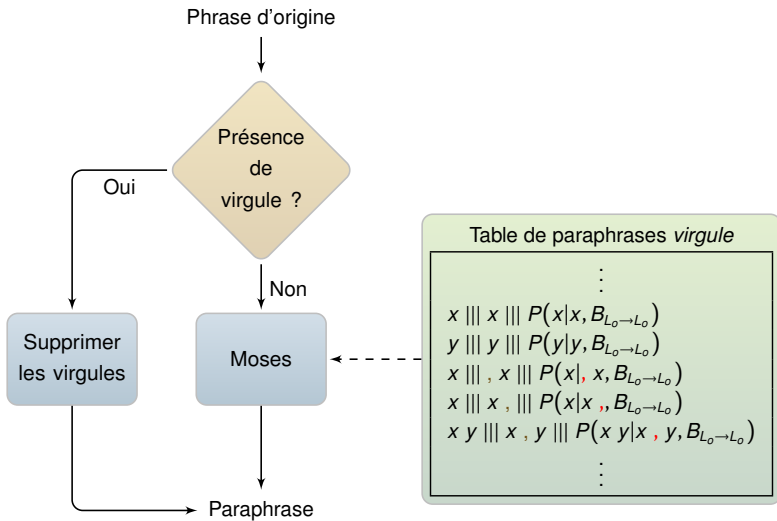
Coût des approximations du modèle



- **Nous avons montré que les différentes heuristiques ont un impacte important sur l'ordre des n meilleures paraphrases.**



Limite de l'évaluation



Définir

Systèmes vocaux

Paraphrases

Étudier

Évaluation classique

Production classique

Limites

Proposer

Une autre évaluation

Une autre production



Système	Référence	Virgule
Sens préservé	40%	76%
Syntaxe correcte	32%	56%
Syntaxe correcte et sens préservé	22 %	55%
Kappa	0,63 (p -valeur $< 10^{-3}$) Accord substantiel	

- Une évaluation des paraphrases fondée sur la naturalité et la conservation du sens est insuffisante.
- **Nous avons montré le besoin de prendre en compte une nouvelle dimension lors de l'évaluation.**



orange

Un cadre de la production de paraphrases

Conclusion

Définir

Systèmes vocaux

Paraphrases

Étudier

Évaluation classique

Production classique

Limites

Proposer

Une autre évaluation

Une autre production

- Développement d'un générateur de paraphrases :
 - les performances trop faibles pour une utilisation industrielle ;
 - le modèle de langue ne corrige pas suffisamment la syntaxe ;
 - la nécessité de prendre en compte la tâche lors de la production ;
 - les contraintes du décodeur limitent le modèle.
- Développement d'une plateforme d'évaluation :
 - insuffisance d'une évaluation fondé uniquement sur le sens et la syntaxe ;
 - besoin d'une nouvelle dimension pour l'évaluation.

Définir

Systèmes vocaux
Paraphrases

Étudier

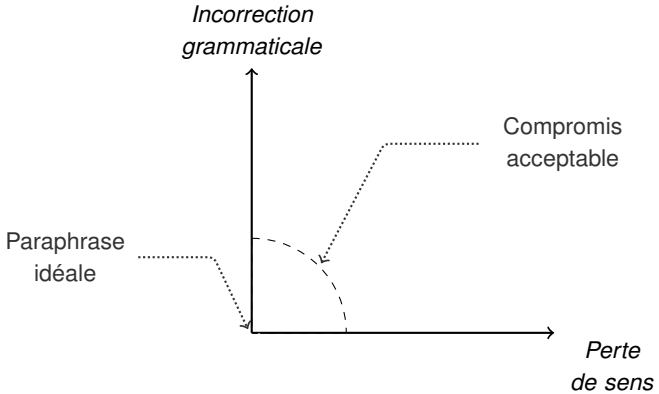
Évaluation classique
Production classique
Limites

Proposer

Une autre évaluation
Une autre production

- **Définir** : Une tâche et un contexte
Systèmes vocaux humain-machine
Paraphrases
- **Étudier** : le cadre classique de la paraphrase
Évaluation des paraphrases
Production de paraphrases
Les limites de la production et de l'évaluation classique
- **Proposer** : un autre modèle plus adapté
Une autre façon d'évaluer les paraphrases
Une autre façon de produire des paraphrases

Les trois objectifs des paraphrases



Définir

Systèmes vocaux

Paraphrases

Étudier

Évaluation classique

Production classique

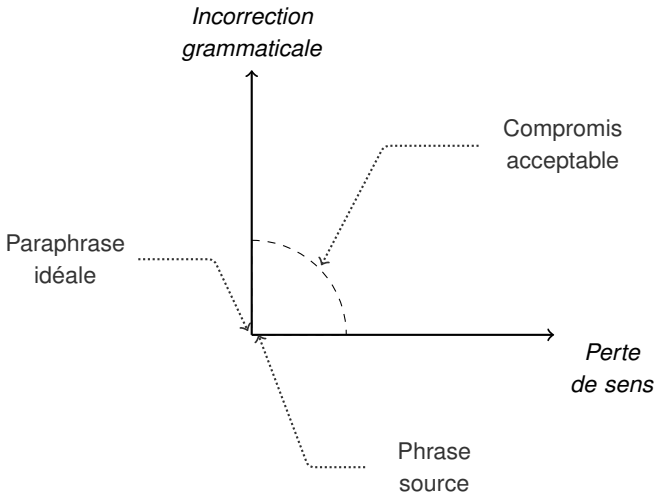
Limites

Proposer

Une autre évaluation

Une autre production

Les trois objectifs des paraphrases



Définir

Systèmes vocaux

Paraphrases

Étudier

Évaluation classique

Production classique

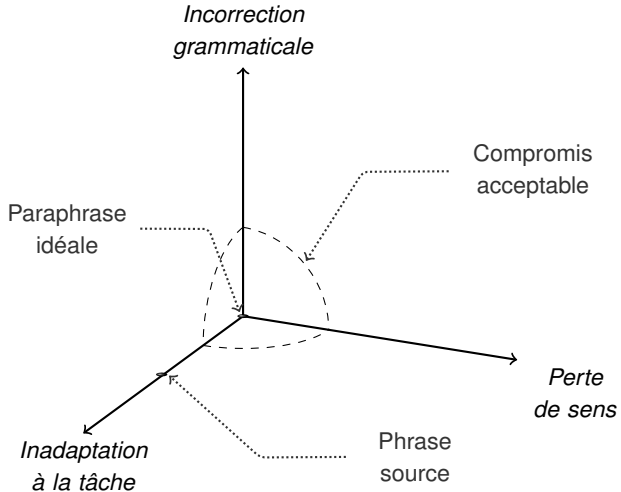
Limites

Proposer

Une autre évaluation

Une autre production

Les trois objectifs des paraphrases



Définir

Systèmes vocaux

Paraphrases

Étudier

Évaluation classique

Production classique

Limites

Proposer

Une autre évaluation

Une autre production



Comparer des générateurs

Définir

Systèmes vocaux

Paraphrases

Étudier

Évaluation classique

Production classique

Limites

Proposer

Une autre évaluation

Une autre production

- Critère : ?
- Mesure : ?

Définir

Systèmes vocaux

Paraphrases

Étudier

Évaluation classique

Production classique

Limites

Proposer

Une autre évaluation

Une autre production

Conjecture

Il est, en général, plus difficile de construire un générateur capable de produire une paraphrase correcte très différente de la phrase d'origine qu'une paraphrase très proche.

- Critère : la variabilité.
- Mesure : ?



Définir

Systèmes vocaux

Paraphrases

Étudier

Évaluation classique

Production classique

Limites

Proposer

Une autre évaluation

Une autre production

Corollaire

Un générateur capable de produire des paraphrases correctes très différentes de la phrase d'origine est tout aussi capable, en général, de produire des paraphrases correctes moins différentes.

- Critère générique : la variabilité.
- Mesure : ?



Définir

Systèmes vocaux

Paraphrases

Étudier

Évaluation classique

Production classique

Limites

Proposer

Une autre évaluation

Une autre production

Corollaire

Un générateur capable de produire des paraphrases correctes très différentes de la phrase d'origine est tout aussi capable, en général, de produire des paraphrases correctes moins différentes.

- Critère générique : la variabilité.
- Mesure : distance d'édition sur les caractères.



Conjecture

Il existe, en général, plus de paraphrases acceptables pour une phrase longue que pour une phrase courte.

- Critère générique : la variabilité.
- Mesure : distance d'édition sur les caractères normalisé.

$$\mu_{\text{var}}(\text{Paraphrase} \mid \text{Origine}) = \min \left(\frac{\text{distance d'édition}(\text{Paraphrase}, \text{Origine})}{|\text{Origine}|}, 1 \right)$$



Comparer des générateurs

Résultats

Système	Référence	Virgule
Sens préservé	32 %	56 %
Syntaxe correcte	40 %	76 %
μ_{var} moyen	20,8% ±12,1	4,0% ±3,6
Moyenne harmonique des trois objectifs	9,6%	3,7%
Syntaxe correcte et sens préservé	22 %	55 %
μ_{var} moyen des paraphrases correctes	17,8% ±7,4	3,8% ±2,1
Moyenne harmonique des paraphrases correctes	9,8%	3,6%
Kappa	0,63 (p -valeur $< 10^{-3}$) Accord substantiel	

Définir

Systèmes vocaux

Paraphrases

Étudier

Évaluation classique

Production classique

Limites

Proposer

Une autre évaluation

Une autre production



orange

Une autre façon de d'évaluer les paraphrases

Récapitulatif

- Nous avons montré que évaluation classique focaliser sur le sens n'est pas pertinente.
- Nous avons montré qu'**il est nécessaire de prendre en compte la tâche lors de l'évaluation**
- Nous avons proposé :
 - **une tâche générique pour comparer les générateurs : la variabilité.**
 - **une mesure pour cette tâche : la distance d'édition normalisé.**

Définir

Systèmes vocaux

Paraphrases

Étudier

Évaluation classique

Production classique

Limites

Proposer

Une autre évaluation

Une autre production

- **Définir** : Une tâche et un contexte
Systèmes vocaux humain-machine
Paraphrases
- **Étudier** : le cadre classique de la paraphrase
Évaluation des paraphrases
Production de paraphrases
Les limites de la production et de l'évaluation classique
- **Proposer** : un autre modèle plus adapté
Une autre façon d'évaluer les paraphrases
Une autre façon de produire des paraphrases

Définir

Systèmes vocaux
Paraphrases

Étudier

Évaluation classique
Production classique
Limites

Proposer

Une autre évaluation
Une autre production



orange

Une autre approche de la production

Principe

- Nous avons montré que :
 - le modèle de langue ne corrige pas suffisamment la syntaxe ;
 - la nécessité de prendre en compte la tâche lors de la production ;
 - les contraintes du décodeur limitent le modèle.
- Production de paraphrases et traduction sont deux activités différentes :
 - pas besoin de tout transformer ;
 - pas besoin d'ordonner le décodage
- Nous souhaitons :
 - un modèle adapté aux spécificités de la paraphrase ;
 - pouvoir évaluer globalement une paraphrase (approche holistique).

Définir

Systèmes vocaux

Paraphrases

Étudier

Évaluation classique

Production classique

Limites

Proposer

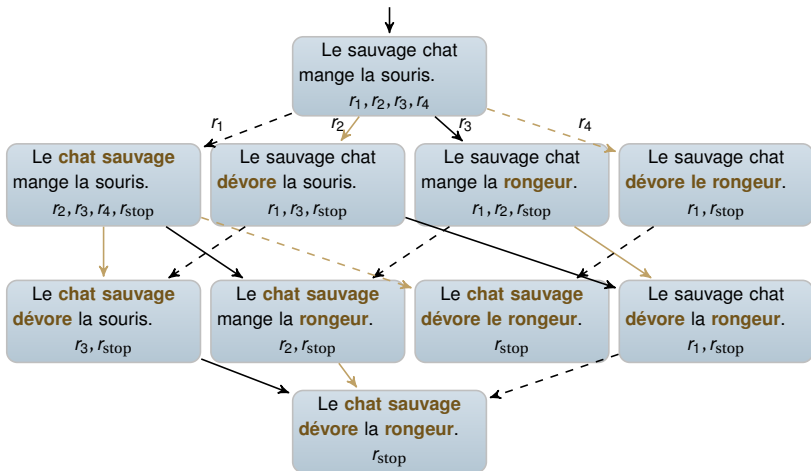
Une autre évaluation

Une autre production



Production par règles de transformation

r_1	$\{([2; 3], \text{chat sauvage})\}$	\dashrightarrow
r_2	$\{([4; 4], \text{dévore})\}$	\rightarrow
r_3	$\{([6; 6], \text{rongeur})\}$	\rightarrow
r_4	$\{([4; 6], \text{dévore le rongeur})\}$	\dashrightarrow



Définir

Systèmes vocaux

Paraphrases

Étudier

Évaluation classique

Production classique

Limites

Proposer

Une autre évaluation

Une autre production

Définir

Systèmes vocaux
Paraphrases

Étudier

Évaluation classique
Production classique
Limites

Proposer

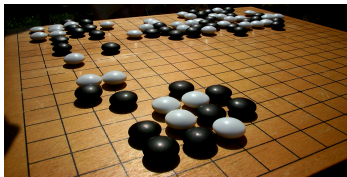
Une autre évaluation
Une autre production

- **Avantages :**
 - Évaluation globale des paraphrases.
 - Pas besoin de tout transformer.
 - Gestion intrinsèque des mots inconnus.
 - Pas besoin d'ordonner le décodage
- **Inconvénient :**
 - Inadapté pour l'algorithme de Viterbi.



orange

Jeu de go



Définir

Systèmes vocaux

Paraphrases

Étudier

Évaluation classique

Production classique

Limites

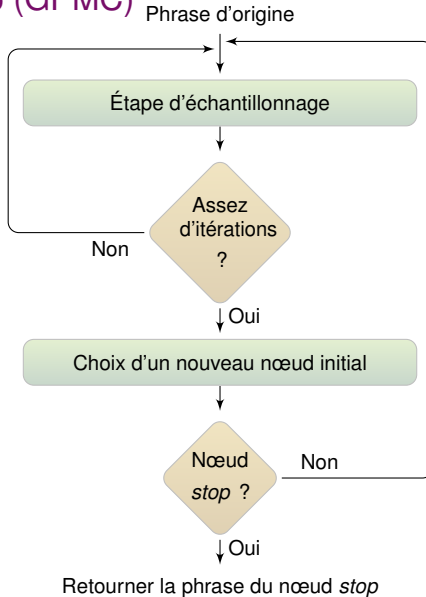
Proposer

Une autre évaluation

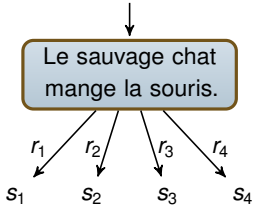
Une autre production

- Un des derniers jeux à résister aux ordinateurs :
 - Difficulté à évaluer une position intermédiaire.
 - Arbre des coups possibles très large.
- Une nouvelle classe d'algorithmes par échantillonnage de Monte-Carlo a très fortement amélioré les performances des programmes de jeux de Go.
- **Nous avons adapté cet algorithme pour la production de paraphrases.**

Générateur de Paraphrases par échantillonnage de Monte-Carlo (GPMC)



Étape d'échantillonnage



Définir

Systèmes vocaux
Paraphrases

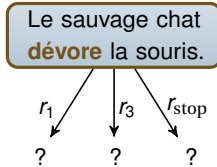
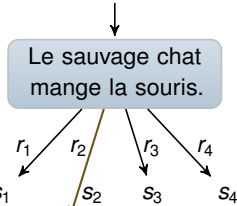
Étudier

Évaluation classique
Production classique
Limites

Proposer

Une autre évaluation
Une autre production

Étape d'échantillonnage



Définir

Systèmes vocaux

Paraphrases

Étudier

Évaluation classique

Production classique

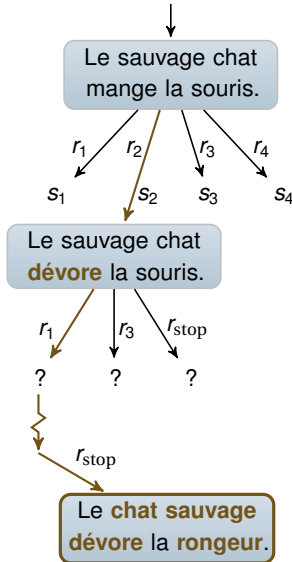
Limites

Proposer

Une autre évaluation

Une autre production

Étape d'échantillonnage



Définir

Systèmes vocaux

Paraphrases

Étudier

Évaluation classique

Production classique

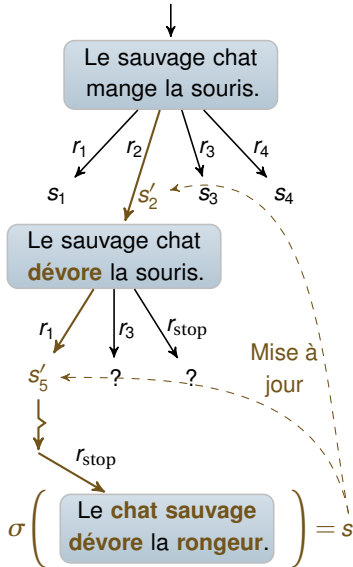
Limites

Proposer

Une autre évaluation

Une autre production

Étape d'échantillonnage



Définir

Systèmes vocaux

Paraphrases

Étudier

Évaluation classique

Production classique

Limites

Proposer

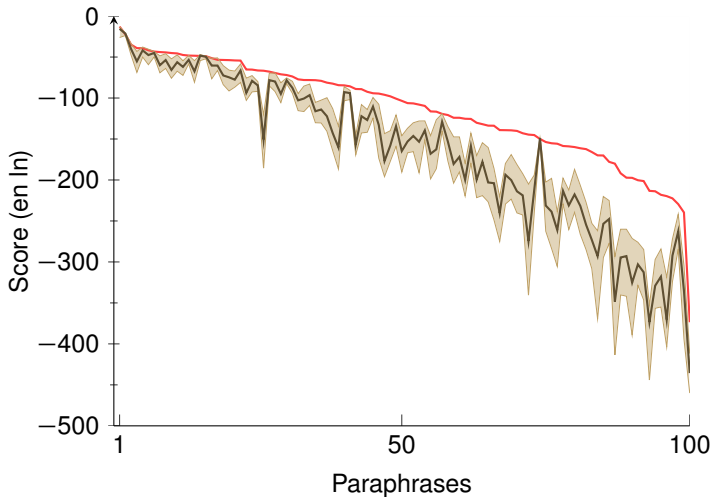
Une autre évaluation

Une autre production



Résultats préliminaires

- GPMC est moins stable et moins performant que Moses.



Définir

Systèmes vocaux

Paraphrases

Étudier

Évaluation classique

Production classique

Limites

Proposer

Une autre évaluation

Une autre production



Ajout d'heuristiques

Principe

Définir

Systèmes vocaux

Paraphrases

Étudier

Évaluation classique

Production classique

Limites

Proposer

Une autre évaluation

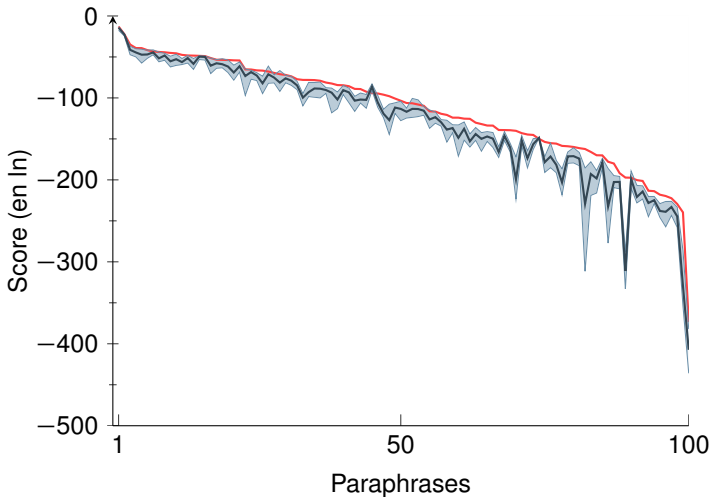
Une autre production

- Score de découpage optimal :
 - Fonction de score holistique.
 - Plus besoin de tout transformer.
- Réduction de l'espace d'exploration :
 - Empêcher d'explorer un sous-arbre complètement défini.
 - Amélioration de l'efficacité des explorations.



Nouvelles performances

- Une forte amélioration des performances de GPMC. Elles sont désormais proches de celles de Moses.



Définir

Systèmes vocaux

Paraphrases

Étudier

Évaluation classique

Production classique

Limites

Proposer

Une autre évaluation

Une autre production



Performances finales

Système	Moses	GPAC
Sens préservé	32 %	34 %
Syntaxe correcte	40 %	33 %
μ_{var} moyen	20,8 % \pm 12,1	17,1 % \pm 11,9
Moyenne harmonique des trois objectifs	9,6 %	8,4 %
Syntaxe correcte et sens préservé	22 %	22 %
μ_{var} moyen des paraphrases correctes	17,8 % \pm 7,4	11,4 % \pm 8,0
Moyenne harmonique des paraphrases correctes	9,8 %	8,0 %
Kappa	0,71 (p -valeur $< 10^{-3}$) Accord substantiel	

Définir

Systèmes vocaux

Paraphrases

Étudier

Évaluation classique

Production classique

Limites

Proposer

Une autre évaluation

Une autre production



Définir

Systèmes vocaux

Paraphrases

Étudier

Évaluation classique

Production classique

Limites

Proposer

Une autre évaluation

Une autre production

Conclusion

Définir

Systèmes vocaux
Paraphrases

Étudier

Évaluation classique
Production classique
Limites

Proposer

Une autre évaluation
Une autre production

- Démontrer les spécificités de la production de paraphrases :
 - Nécessité de prendre en compte la tâche lors de l'évaluation et de la production.
 - L'inutilité d'une production de gauche à droite cherchant à tout transformer.
- Proposer les outils prenant en compte ces spécificités :
 - Un critère d'évaluation générique.
 - Une plateforme d'évaluation.
 - Un générateur holistique innovant fondé sur l'échantillonnage de Monte-Carlo.

Définir

Systèmes vocaux
Paraphrases

Étudier

Évaluation classique
Production classique
Limites

Proposer

Une autre évaluation
Une autre production

- Faire évoluer le modèle de la paraphrase :
 - modèle de langue ;
 - intégration de modèle des systèmes-vocaux ;
 - pondération des différentes composantes.
- Étudier et améliorer GPMC :
 - fonction de compromis exploration/exploitation ;
 - biaiser l'échantillonnage avec un score *a priori* ;



orange

Merci



Putois G, Chevelu J, Boidin C.

Paraphrase generation to improve Text-To-Speech Synthesis.

Dans : Proceedings of the 11th Annual Conference of the International Speech Communication Association (Interspeech). Tokyo ; 2010.



Chevelu J, Putois G, Lepage Y.

The True Score of Statistical Paraphrase Generation.

Dans : Proceedings of the 23rd Conference on Computational Linguistics (Coling) : Posters. Pékin ; 2010.



Chevelu J, Putois G, Lepage Y, Moudenc T.

L'évaluation des paraphrases : pour une prise en compte de la tâche.

Dans : Actes de Traitement Automatique des Langues Naturelles (TALN). Montréal ; 2010.



Lardilleux A, Chevelu J, Lepage Y, Putois G, Gosme J.

Lexicons or phrase tables ? An investigation in sampling-based multilingual alignment.

In : Proceedings of the 3rd Workshop on Example-Based Machine Translation (EBMT). Dublin : CNGL ; 2009.



Boidin C, Reiser V, Van Der Plas L, Lemon O, Chevelu J.

Predicting how it sounds : Re-ranking dialogue prompts based on TTS quality for adaptive Spoken Dialogue Systems.

Dans : Proceedings of the Interspeech Special Session : Machine Learning for Adaptivity in Spoken Dialogue. Brighton ; 2009.



Chevelu J, Lavergne T, Lepage Y, Moudenc T.

Transformation Rules and Monte-Carlo Sampling : a Different Approach for Statistical Paraphrase Generation.

Dans : Proceedings of the Conference of the Pacific Association for Computational Linguistics (PACLING). Sapporo ; 2009.



Chevelu J, Lavergne T, Lepage Y, Moudenc T.

Introduction of a new paraphrase generation tool based on Monte-Carlo sampling.

Dans : Proceedings of the ACL-IJCNLP Conference Short Papers. Singapour ; 2009.

Définir

Systèmes vocaux

Paraphrases

Étudier

Évaluation classique

Production classique

Limites

Proposer

Une autre évaluation

Une autre production



orange

Contexte

Production
statistique de
paraphrases

Évaluation

GPMC

Bonus

Entrez à vos risques et périls !



Travaux antérieurs

- Paraphrases pour les systèmes vocaux :
 - transformer un texte « plutôt écrit » en texte « plutôt orale » sans évaluation acoustique [Kaji et coll., 2004]
 - tentative d'améliorer un SVI lorsque des paraphrases sont disponibles [Nakatsu et White, 2006]
 - production de paraphrases pour la synthèse mais sans évaluation acoustique ni sémantique [Cahill et coll., 2009]
- Paraphrases pour remplacer une phrase :
 - compression de texte [Knight et Marcu, 2000]
 - normalisation de texte [Nasr, 1996]
 - aide à la rédaction [Max, 2008]
- Paraphrases pour améliorer des systèmes de traitement automatique de texte :
 - systèmes de question-réponse [Duclaye et coll., 2003]
 - systèmes d'extraction d'information [Sekine, 2005]
 - systèmes de traduction automatique [Callison-Burch et coll., 2006]
- Paraphrases pour aider l'évaluation :
 - augmenter le nombre de référence [Lepage et Denoual, 2003]

Contexte

Production
statistique de
paraphrases

Évaluation

GPMC



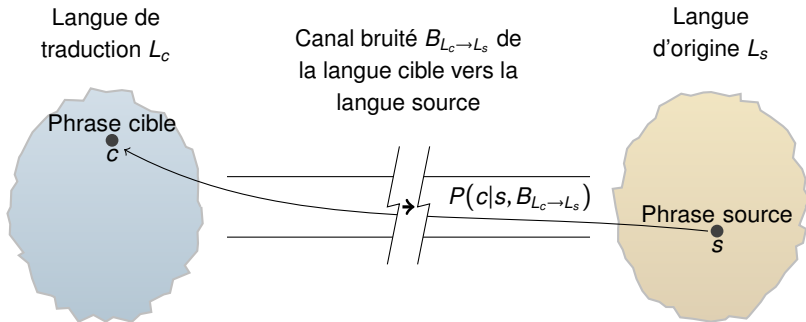
Modèle du canal bruité utilisé en traduction automatique statistique

Contexte

Production
statistique de
paraphrases

Évaluation

GPMC



$$c^* = \arg \max_c P(c|s, B_{L_c \rightarrow L_s}) \quad (3)$$

$$c^* = \arg \max_c P(c) \times P(s|c, B_{L_c \rightarrow L_s}) \quad (4)$$



Modèle de langue

- Même problème qu'en traduction automatique.
- Objectif : prédire la qualité syntaxique d'une phrase.
- Modèle n -grammes : limitation de l'historique.

$$P(m_1 m_2 m_3 \dots m_k) = \prod_{j=1}^k P(m_j | m_{j-1} \dots m_{j-n+1}) \quad (8)$$

- Estimation de la probabilité des segments rares :
 - heuristiques de lissage et de retrait [Katz, 1987].
- Utilisation de SRILM [Stolcke, 2002].

Contexte

Production
statistique de
paraphrases

Évaluation

GPNC



Table de paraphrases : table de traduction monolingue

merci ||| miséricorde ||| 0.571429
merci ||| la miséricorde ||| 0.571429
merci ||| eleison ||| 0.571429
merci ||| remerciements que j' ai ||| 0.333333
merci ||| remerciements du ||| 0.333333
merci ||| et grâce à vous ||| 0.333333
:

- Traduction :
 - Bi-corpus bilingue aligné au niveau phrase
 - Aligneur sous-phrastique (Giza++ [Och et Ney, 2003], Anymalign [Lardilleux et Lepage, 2008])
- Production de paraphrases :
 - Rareté des corpus monolingues aligné
 - Problème de l'acquisition de paraphrases
 - [Barzilay et McKeown, 2001] [Dolan et Brockett, 2005] [Lepage et Denoual, 2005] [Bouamor et coll., 2010]

Contexte

Production
statistique de
paraphrases

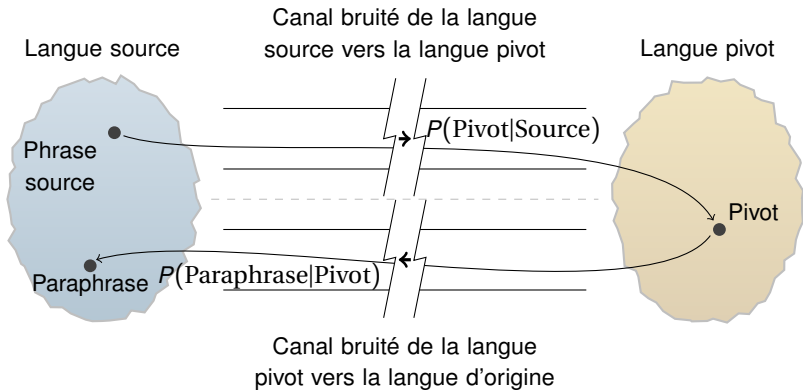
Évaluation

GPMC



Table de paraphrases par langue pivot

Bannard et Callison-Burch [2005]



Contexte

Production statistique de paraphrases

Évaluation

GPIC



Table de paraphrases par auto-jointure d'une table de traduction

Table de traduction
bilingue

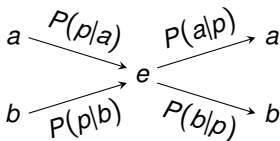


Table intermédiaire
non triée

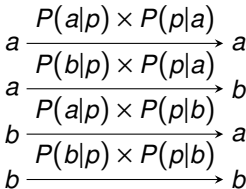
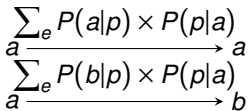
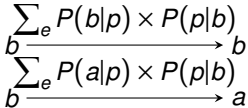
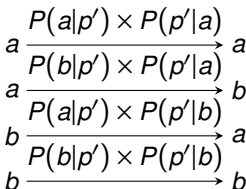
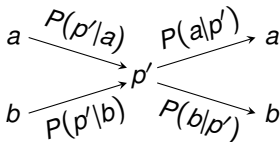


Table de paraphrase



\otimes

Σ





Treillis de production

Le chat sauvage mange la souris.

Le ||| Le ||| $P(\text{Le}|\text{Le},B)$

sauvage ||| sauvage ||| $P(\text{sauvage}|\text{sauvage},B)$

chat ||| chat ||| $P(\text{chat}|\text{chat},B)$

mange ||| mange ||| $P(\text{mange}|\text{mange},B)$

la ||| la ||| $P(\text{la}|\text{la},B)$

souris ||| souris ||| $P(\text{souris}|\text{souris},B)$

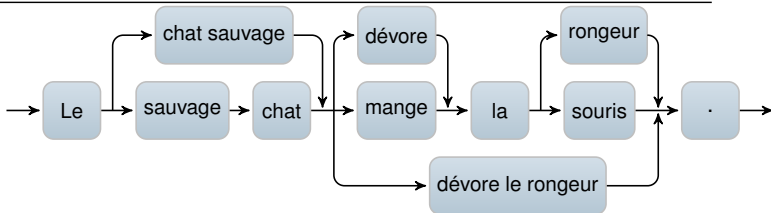
. ||| . ||| $P(\cdot|\cdot,B)$

sauvage chat ||| chat sauvage ||| $P(\text{sauvage chat}|\text{chat sauvage},B)$

mange ||| dévore ||| $P(\text{mange}|\text{dévore},B)$

souris ||| rongeur ||| $P(\text{souris}|\text{rongeur},B)$

mange la souris ||| dévore la souris ||| $P(\text{mange la souris}|\text{dévore la souris},B)$





Décodeur

- Décodeur fondé sur l'algorithme de Viterbi [1967].
- Programmation dynamique :
 - complexité linéaire en la longueur de la phrase
- Contraintes : la fonction de score doit être
 - incrémentale ;
 - calculable avec un historique restreint.
- Heuristique de recherche par faisceau :
 - réduction du nombre d'hypothèse par état ;
- Utilisation de Moses [Koehn et coll., 2007].

Contexte

Production
statistique de
paraphrases

Évaluation

GPMC



Complexité du décodeur

- Complexité au pire : $\mathcal{O}(N \times V^n \times L)$.
- Pire des cas : la table de paraphrase contient V^j entrées pour chaque j -grammes jusqu'aux N -grammes.
- Pour calculer jusqu'au mot k à partir du calcul jusqu'à $k - 1$:
 - V entrées portant sur un 1-gramme applicable depuis la pile $k - 1$ pour arriver au mot k ; la pile $k - 1$ contient les V^{n-1} états nécessaires au modèle de langue ; soit $V^{n-1} \times V$ opérations ;
 - V^2 entrées portant sur un 2-grammes applicable depuis la pile $k - 2$ pour arriver au mot k ; la pile $k - 1$ contient les V^{n-2} états nécessaires au modèle de langue ; soit $V^{n-2} \times V^2$ opérations ;
 - \vdots
 - V^N entrées portant sur un N -grammes applicable depuis la pile $k - N$ pour arriver au mot k ; la pile $k - N$ contient les V états nécessaires au modèle de langue ; soit $V \times V^{n-1}$ opérations ;
- En fait :

$$\mathcal{O} \left(\sum_{i=0}^{L-1} \min(N, i) \times V^{\min(n, i+1)} \right)$$



Cas de la somme sur tous les découpages

- Modèle de langue d'ordre 3 ; table de paraphrase d'ordre 2.

Le ||| Le ||| 1
sauvage ||| sauvage ||| 0,74
chat ||| chat ||| 0,74
sauvage chat ||| chat sauvage ||| 0,30
méchant chat ||| chat sauvage ||| 0,15
mange ||| mange ||| 0,5
mange ||| dévore ||| 0,5



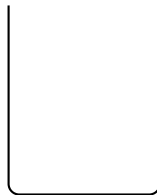
Mot 1
Le



Mot 2
chat



Mot 3
sauvage



Mot 4
mange

Contexte

Production
statistique de
paraphrases

Évaluation

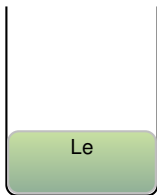
GPMC



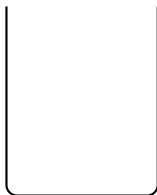
Cas de la somme sur tous les découpages

- Construction des hypothèses de traduction.

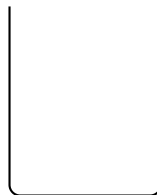
Le ||| Le ||| 1
sauvage ||| sauvage ||| 0,74
chat ||| chat ||| 0,74
sauvage chat ||| chat sauvage ||| 0,30
méchant chat ||| chat sauvage ||| 0,15
mange ||| mange ||| 0,5
mange ||| dévore ||| 0,5



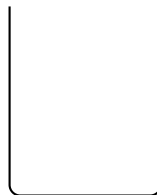
Mot 1
Le



Mot 2
chat



Mot 3
sauvage



Mot 4
mange

Contexte

Production
statistique de
paraphrases

Évaluation

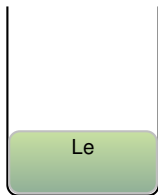
GPMC



Cas de la somme sur tous les découpages

- Deux mots conservés pour le modèle de langue.

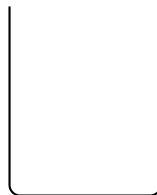
Le ||| Le ||| 1
sauvage ||| sauvage ||| 0,74
chat ||| chat ||| 0,74
sauvage chat ||| chat sauvage ||| 0,30
méchant chat ||| chat sauvage ||| 0,15
mange ||| mange ||| 0,5
mange ||| dévore ||| 0,5



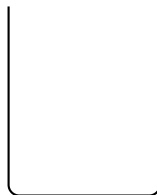
Mot 1
Le



Mot 2
chat



Mot 3
sauvage



Mot 4
mange

Contexte

Production
statistique de
paraphrases

Évaluation

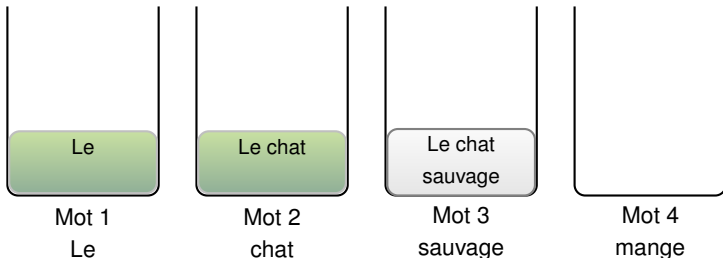
GPMC



Cas de la somme sur tous les découpages

- Étendre les hypothèses de la pile 2.

Le ||| Le ||| 1
sauvage ||| sauvage ||| 0,74
chat ||| chat ||| 0,74
sauvage chat ||| chat sauvage ||| 0,30
méchant chat ||| chat sauvage ||| 0,15
mange ||| mange ||| 0,5
mange ||| dévore ||| 0,5

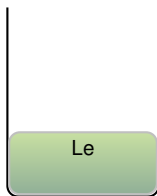




Cas de la somme sur tous les découpages

- Étendre les hypothèses de la pile 1.

Le ||| Le ||| 1
sauvage ||| sauvage ||| 0,74
chat ||| chat ||| 0,74
sauvage chat ||| chat sauvage ||| 0,30
méchant chat ||| chat sauvage ||| 0,15
mange ||| mange ||| 0,5
mange ||| dévore ||| 0,5



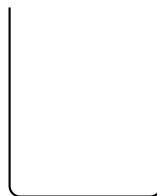
Mot 1
Le



Mot 2
chat



Mot 3
sauvage



Mot 4
mange

Contexte

Production
statistique de
paraphrases

Évaluation

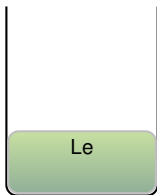
GPMC



Cas de la somme sur tous les découpages

- Calcul du score des états.

Le ||| Le ||| 1
sauvage ||| sauvage ||| 0,74
chat ||| chat ||| 0,74
sauvage chat ||| chat sauvage ||| 0,30
méchant chat ||| chat sauvage ||| 0,15
mange ||| mange ||| 0,5
mange ||| dévore ||| 0,5



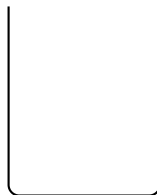
Mot 1
Le



Mot 2
chat



Mot 3
sauvage



Mot 4
mange

Contexte

Production
statistique de
paraphrases

Évaluation

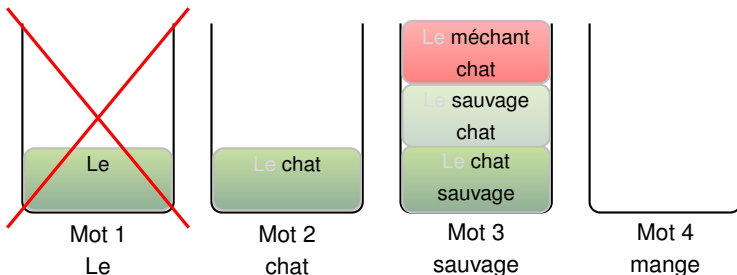
GPMC



Cas de la somme sur tous les découpages

- Supprimer l'historique inutile.

Le ||| Le ||| 1
sauvage ||| sauvage ||| 0,74
chat ||| chat ||| 0,74
sauvage chat ||| chat sauvage ||| 0,30
méchant chat ||| chat sauvage ||| 0,15
mange ||| mange ||| 0,5
mange ||| dévore ||| 0,5





Cas de la somme sur tous les découpages

- Étendre les hypothèses de la pile 2 et 3.

Le ||| Le ||| 1

sauvage ||| sauvage ||| 0,74

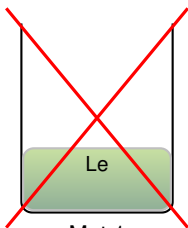
chat ||| chat ||| 0,74

sauvage chat ||| chat sauvage ||| 0,30

méchant chat ||| chat sauvage ||| 0,15

mange ||| mange ||| 0,5

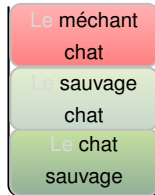
mange ||| dévore ||| 0,5



Mot 1
Le



Mot 2
chat



Mot 3
sauvage



Mot 4
mange

Contexte

Production
statistique de
paraphrases

Évaluation

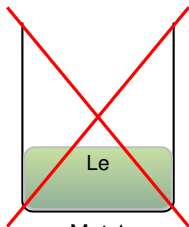
GPMC



Cas de la somme sur tous les découpages

- Calcul du score des états.

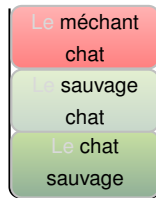
Le ||| Le ||| 1
sauvage ||| sauvage ||| 0,74
chat ||| chat ||| 0,74
sauvage chat ||| chat sauvage ||| 0,30
méchant chat ||| chat sauvage ||| 0,15
mange ||| mange ||| 0,5
mange ||| dévore ||| 0,5



Mot 1
Le



Mot 2
chat



Mot 3
sauvage



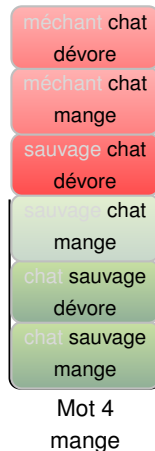
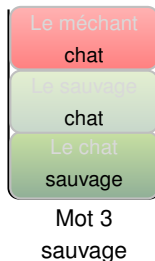
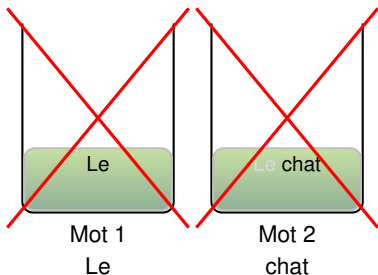
Mot 4
mange



Cas de la somme sur tous les découpages

- Supprimer l'historique inutile.

Le ||| Le ||| 1
 sauvage ||| sauvage ||| 0,74
 chat ||| chat ||| 0,74
 sauvage chat ||| chat sauvage ||| 0,30
 méchant chat ||| chat sauvage ||| 0,15
 mange ||| mange ||| 0,5
 mange ||| dévore ||| 0,5



Contexte

Production
statistique de
paraphrases

Évaluation

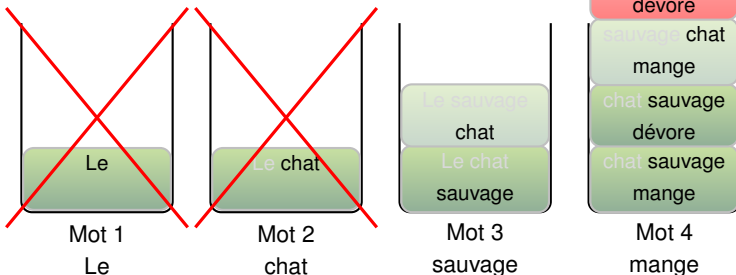
GPMC



Cas de la somme sur tous les découpages

- Fusion des états identique : **impossible !**
- Il faut un score pour chaque forme de surface possible !

Le ||| Le ||| 1
sauvage ||| sauvage ||| 0,74
chat ||| chat ||| 0,74
sauvage chat ||| chat sauvage ||| 0,30
méchant chat ||| chat sauvage ||| 0,15
mange ||| mange ||| 0,5
mange ||| dévore ||| 0,5

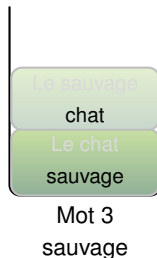
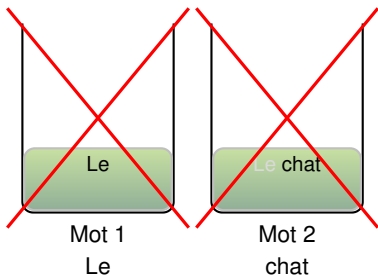




Cas de la somme sur tous les découpages

- Complexité au pire : $\mathcal{O}(V^L)$.

Le ||| Le ||| 1
 sauvage ||| sauvage ||| 0,74
 chat ||| chat ||| 0,74
 sauvage chat ||| chat sauvage ||| 0,30
 méchant chat ||| chat sauvage ||| 0,15
 mange ||| mange ||| 0,5
 mange ||| dévore ||| 0,5



Contexte

Production
statistique de
paraphrases

Évaluation

GPMC



Cas de la somme sur tous les découpages

Contexte

**Production
statistique de
paraphrases**

Évaluation

GPMC



orange

Contexte

Production
statistique de
paraphrases

Évaluation

GPMC

Utilisation d'un synonyme :

- O : La Commission soutient activement la mise au point au niveau mondial et de façon **uniformisée** d'un cycle de tests des motocycles.
- P : La Commission soutient activement la mise au point au niveau mondial et de façon **standardisée** d'un cycle de tests des motocycles.

Décomposition d'un acronyme :

- O : [...] tout a déjà été dit sur les **PME**.
- P : [...] tout a déjà été dit sur les **petites et moyennes entreprises**.

Introduction d'un anglicisme :

- O : [...] dans son **rectificatif** très appréciable.
- P : [...] dans son **corrigendum** très appréciable.



Suppression d'une virgule :

- O : Mais la Commission, non plus, n'a pas joué son rôle et **cela est** grave, Monsieur le Président.
- P : Mais la Commission non plus, n'a pas joué son rôle et **c'est un problème** grave, Monsieur le Président.

Interversion de mots :

- O : Je **vous remercie tous** et j'espère que nous nous reverrons.
- P : Je **voudrais tous vous remercier** et j'espère que nous nous reverrons.



orange

Contexte

Production
statistique de
paraphrases

Évaluation

GPMC

Plusieurs modifications disjointes :

- O : Nous devons demander aux Américains d'être plus actifs et **d'essayer** de remettre les **populations de la région** sur **le droit chemin**.
- P : Nous devons demander aux Américains d'être plus actifs et **de tenter** de remettre les **gens sur place** sur **la bonne voie**.

Plusieurs modifications en continu :

- O : Je **voudrais également mettre en avant une préoccupation liée** à la question de l'immigration.
- P : Je **souligne d'ailleurs l'inquiétude quant** à la question de l'immigration.



Problème d'accord en genre et en nombre :

- O : Et, quoi qu'il en soit, cela **exige** un examen **sectoriel** préalable.
- P : * Et, quoi qu'il en soit, cela **requiert** un examen **sectorielles** préalable.

Problème d'accord en genre et en nombre :

- O : À Vienne, la stratégie européenne pour l'emploi a reçu une nouvelle impulsion **ambitieuse**.
- P : * À Vienne, la stratégie européenne pour l'emploi a reçu une nouvelle impulsion **ambitieux**.

Problème de dépendance à longue distance :

- O : Il est tombé gravement malade et n'a personne pour prendre soin de lui.
- P : * Il est tombé gravement malade, et **elle** n'a personne pour prendre soin de lui.

- **Nous avons montré les limites du modèle de langue.**



orange

Contexte

Production
statistique de
paraphrases

Évaluation

GPMC

Problème de conservation du sens :

- O : Notre recommandation préconise de ne pas lever **son** immunité.
- P : * Notre recommandation préconise de ne pas lever **l'**immunité.

Problème de contexte inconnu :

- O : Pour des raisons pratiques, je considère donc que la proposition **de M.** Fayot est bonne.
- P : * Pour des raisons pratiques, je considère donc que la proposition **du** **président** Fayot est bonne.



orange

Contexte

Production
statistique de
paraphrases

Évaluation

GPMC

Jugement sévère :

O : **Merci beaucoup pour** votre **contribution** ici ce soir.

P : * **Je vous remercie de** votre **intervention** ici ce soir.

Un jugement clément :

O : Il est insensé de continuer à agir comme on le fait aujourd'hui, alors que **presque** 50% **des** 130 espèces de requin sont **aujourd'hui menacées**.

P : Il est insensé de continuer à agir comme on le fait aujourd'hui, alors que **près de 50% de** 130 espèces de requin sont **à présent mises en péril**.



Compromis exploration/exploitation

$$r_n^* = \underset{i}{\operatorname{argmax}} Q_p^\oplus(n, r_i) \quad (9)$$

$$Q_p^\oplus(n, r_i) = Q_p'(n, r_i) + \sqrt{\frac{2\log(s(n))}{s(n, r_i)}} \quad (10)$$

$$s(n_k) \leftarrow s(n_k) + 1 \quad (11)$$

$$s(n_k, r_k) \leftarrow s(n_k, r_k) + 1 \quad (12)$$

$$Q_p(n_k, r_k) \leftarrow \begin{cases} \max(Q_p(n_k, r_k), \sigma(n_S)) & \text{si } Q_p(n_k, r_k) \text{ est défini} \\ \sigma(n_S) & \text{sinon} \end{cases} \quad (13)$$

Contexte

Production
statistique de
paraphrases

Évaluation

GPMP



Compromis exploration/exploitation + rave

$$Q_r^\oplus(n, r_i) = Q_r'(n, r_i) + \sqrt{\frac{2\log(s_R(n))}{s_R(n, r_i)}} \quad (14)$$

$$Q_r'(n, r_i) = \begin{cases} Q_r(n, r_i) & \text{si } Q_r(n, r_i) \text{ est défini} \\ \max \sigma + \epsilon & \text{sinon} \end{cases} \quad (15)$$

$$s_R(n_k) \leftarrow s_R(n_k) + 1 \quad (16)$$

$$s_R(n_k, r_j) \leftarrow s_R(n_k, r_j) + 1 \quad (17)$$

$$Q_r(n_k, r_j) \leftarrow \begin{cases} \max(Q_r(n_k, r_j), \sigma(n_S)) & \text{si } Q_r(n_k, r_j) \text{ est défini} \\ \sigma(n_S) & \text{sinon} \end{cases} \quad (18)$$

$$r_n^* = \underset{r_i}{\operatorname{argmax}} \left(\alpha(n) \times Q_r^\oplus(n, r_i) + (1 - \alpha(n)) \times Q_p^\oplus(n, r_i) \right) \quad (19)$$

$$\alpha(n) = \sqrt{\frac{\beta}{3s(n) + \beta}} \quad (20)$$

Contexte

Production
statistique de
paraphrases

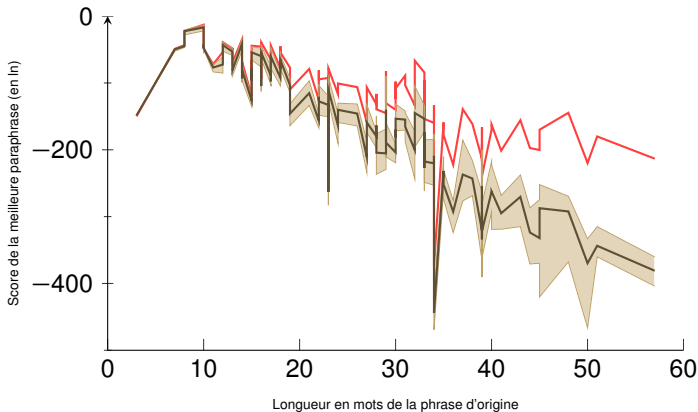
Évaluation

GPMC



Points faibles de la méthode

- Plus la phrase est longue et moins GPMC est capable d'optimiser.



Contexte

Production
statistique de
paraphrases

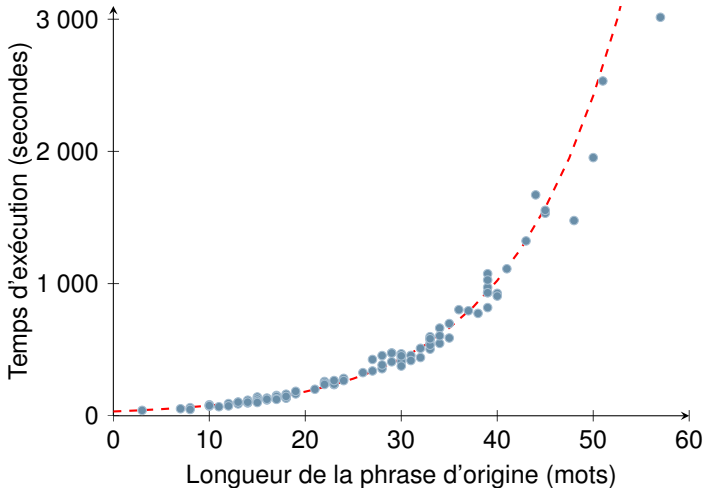
Évaluation

GPMC



Temps de calcul

- Un temps de calcul exponentiel en fonction de la longueur de la phrase.



Contexte

Production
statistique de
paraphrases

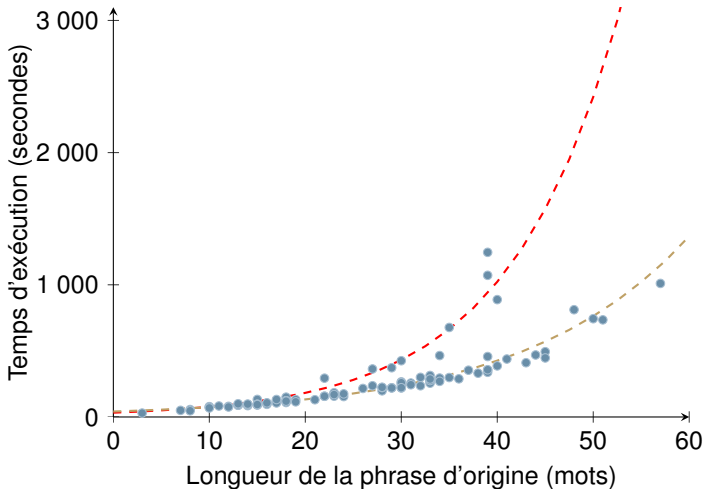
Évaluation

GPMC



Nouveau temps de calcul

- Un important gain de temps de calcul.



Contexte

Production
statistique de
paraphrases

Évaluation

GPMC



Récapitulatif

Systeme	Moses	GPMC	GPMC après améliorations
Moyenne des scores	-113,8	-163,9±1,9	-129,0±0,7
Écart-type moyen par phrase	-	13,4	6,2
Temps d'exécution moyen	1,0±1,1	477,1±531,5	259,7±236,7

Contexte

Production statistique de paraphrases

Évaluation

GPMC