

# Configuration et évaluation d'un système de prédiction de mots au sein d'un logiciel de communication améliorée et alternative

Magali Norré, Max De Wilde  
Université libre de Bruxelles

Les systèmes de Communication Améliorée et Alternative (CAA) :

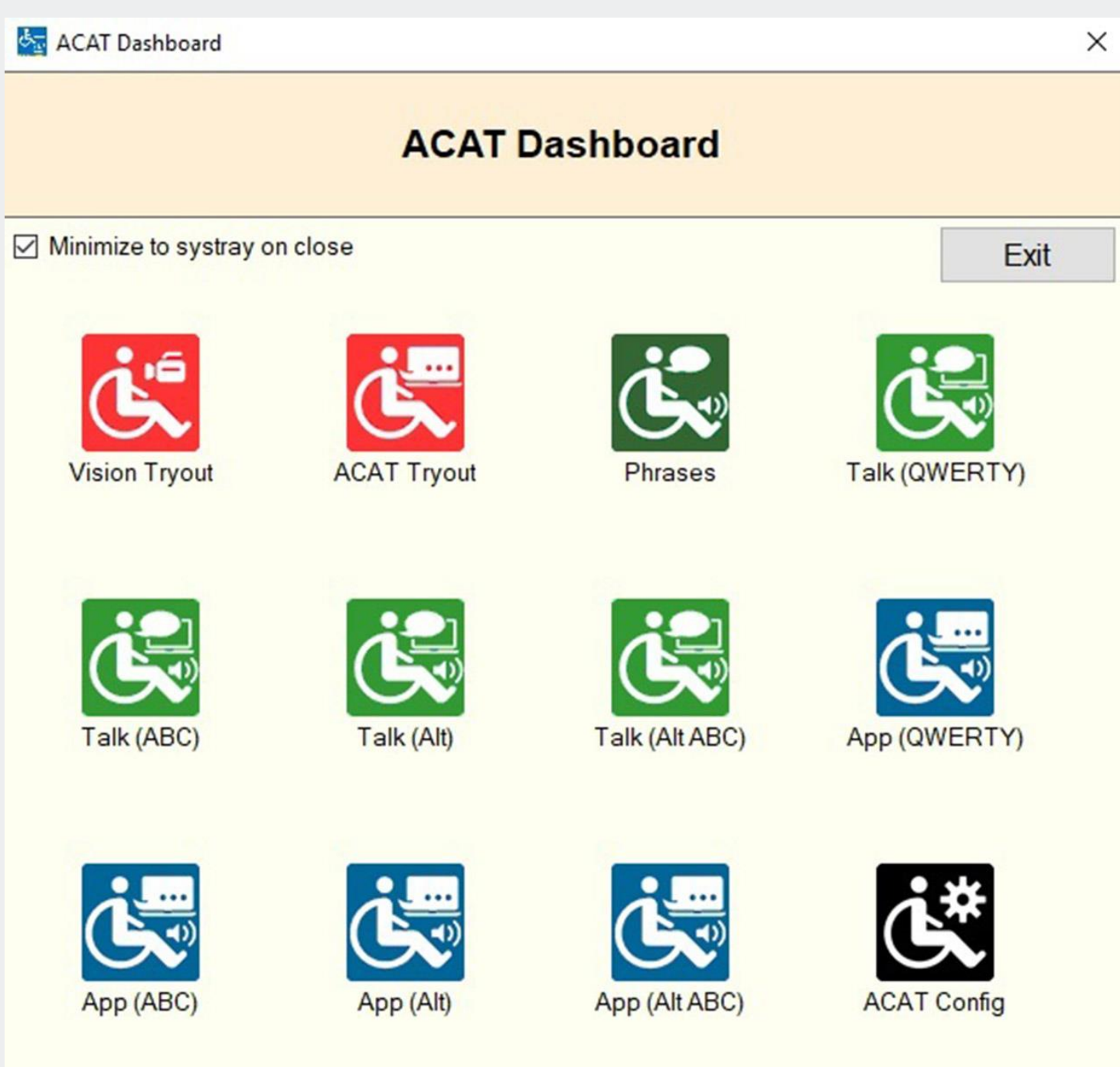
- sont disponibles gratuitement ou commercialisés ;
- permettent d'améliorer la communication orale des personnes en situation de handicap moteur et langagier, ex. : maladie de Charcot, Infirmité Motrice Cérébrale (IMC), etc. ;
- proposent une modalité de saisie adaptée à leurs capacités, une synthèse vocale, un clavier visuel, un système de défilement et/ou de prédiction de mots.

- La vitesse de la communication orale est de 150 à 200 mots par minute, celle via CAA ne dépasse pas 10 à 15.
- Certains utilisateurs ne peuvent pas se servir d'un clavier ou d'une souris, il existe donc des solutions améliorant la saisie.



## Presage & ACAT

### Outils



Deux systèmes *open source* ont été configurés et évalués :

- Presage est un moteur de prédiction de mots conçu par Vescovi (2004). Il est constitué de plusieurs modèles de langage n-grammes qui permettent d'afficher une liste de mots les plus probables.
- ACAT (Assistive Context-Aware Toolkit) est l'interface utilisateur développée par Intel pour Stephen Hawking (Denman et al. 2016). Il s'agit d'un logiciel de CAA en anglais qui comprend deux modes (*Talk* ou *App*) ainsi que plusieurs claviers virtuels accessibles à l'aide de différentes modalités de saisie. Le système de défilement permet de sélectionner les caractères et les mots avec un détecteur de mouvements des sourcils, des joues ou de la bouche via une webcam.

## Méthodologie

- L'objectif : évaluer diverses configurations de Presage afin d'analyser leur efficacité.
- L'impact de plusieurs paramètres a été étudié : (1) le corpus, (2) le modèle de langage n-gramme, (3) le nombre de prédictions à afficher ainsi que (4) l'interaction et la satisfaction de l'utilisateur concernant l'interface ACAT.
- Pour configurer Presage en français, trois corpus ont été évalués : (1) la version 2009 de *Google Books Ngram* (Michel et al. 2011), (2) le livre *À se tortre* et (3) *Perso* que nous avons constitué à partir d'une interview et d'articles de blog d'utilisateurs de CAA. Le corpus de test était constitué d'une centaine de phrases simples.
- Pour évaluer la prédiction, trois métriques ont été utilisées :

$$KSR = \left(1 - \frac{ki+ks}{kn}\right) * 100$$

$$\text{Taux d'économie} = \left(1 - \frac{ki}{kn}\right) * 100$$

$$\text{Hit ratio} = \frac{\text{nombre de mots prédits}}{\text{nombre total de mots}} * 100$$

- Pour évaluer l'interaction et la satisfaction d'un utilisateur ayant une IMC, nous l'avons enregistré en train de recopier une vingtaine de phrases avec ses pieds, il a répondu à un questionnaire et nous avons utilisé le SUS ou *System Usability Scale* (Brooke 1996).

## Résultats

### Évaluation quantitative

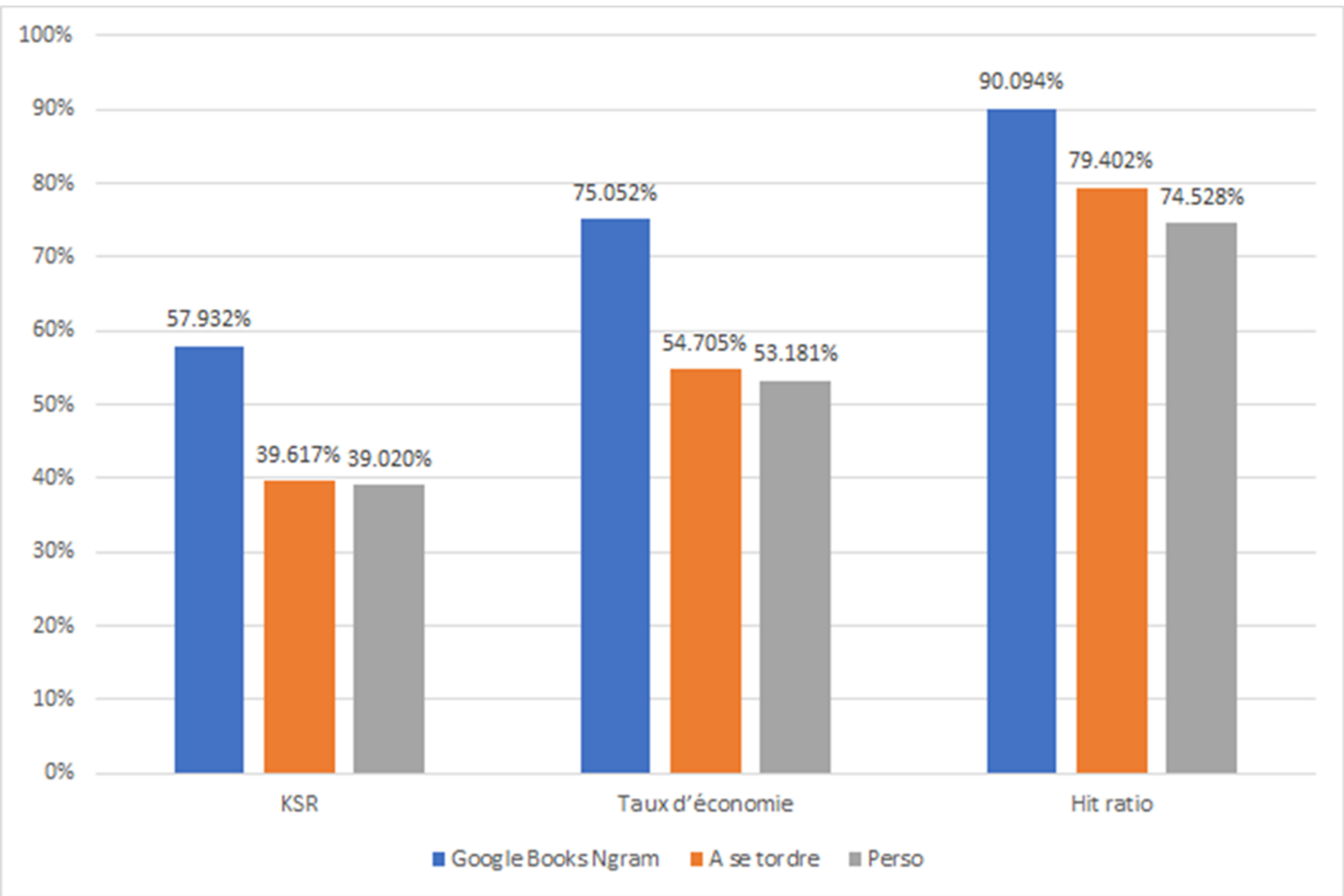


Figure 1 : Corpus

	1	2	3
KSR	43.800%	54.108%	57.932%
Taux d'économie	60.382%	71.287%	75.052%
Hit ratio	87.264%	90.408%	90.094%

Figure 2 : Nombre de n-grammes

	1	5	9
KSR	45.354%	57.932%	61.458%
Taux d'économie	60.501%	75.052%	78.996%
Hit ratio	79.716%	90.094%	92.295%

Figure 3 : Nombre de prédictions

### Évaluation qualitative

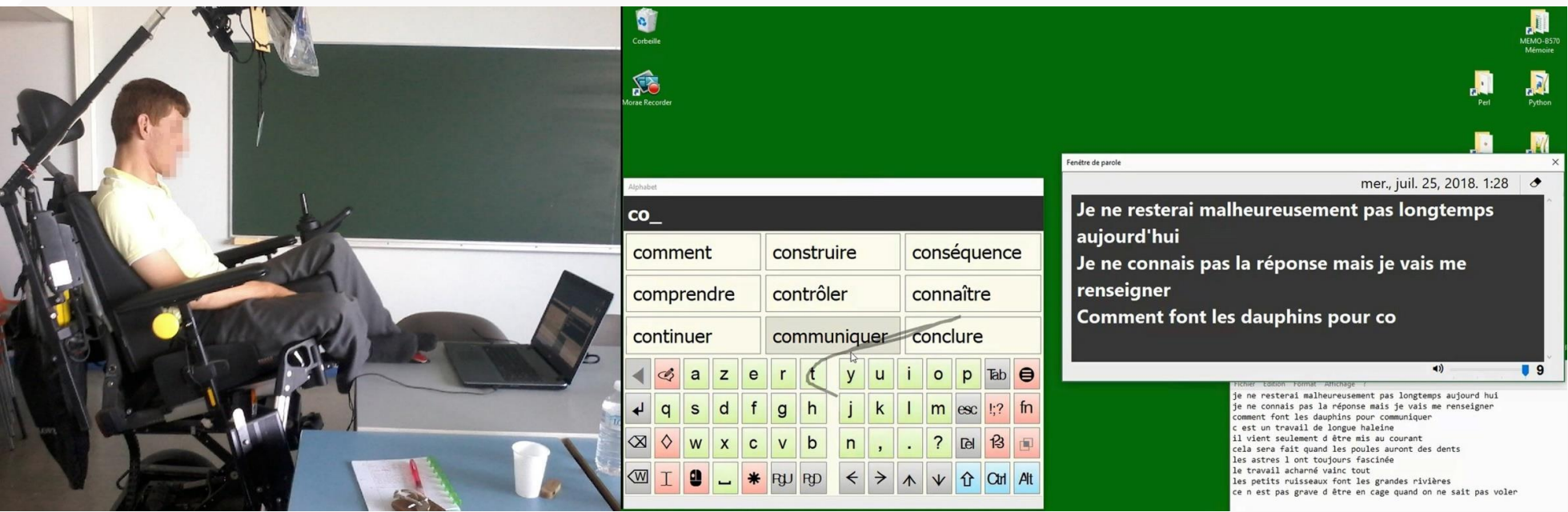


Figure 4 : Test utilisateur

## Conclusion

### 1) Corpus

- *Google Books Ngram* obtient de meilleures performances que les deux autres corpus et permet d'économiser plus d'une saisie sur deux (Fig. 1).

### 2) Prédicteurs

- Le modèle de langage trigramme lissé avec interpolation linéaire (entraîné avec *Google Books Ngram*) est plus performant qu'un modèle d'ordre inférieur (Fig. 2).

### 3) Prédictions

- Plus le nombre de prédictions est grand, meilleurs seront les résultats (Fig. 3).

### 4) Interface utilisateur

- La personne testée (Fig. 4) n'a pas toujours sélectionné la prédiction correcte dès qu'elle apparaissait dans la liste.
- Bien que des mots aient peu de chance d'être utilisés, elle a été satisfaite des systèmes (SUS : 70%), malgré certains bugs.

➡ Ces logiciels *open source* offrent des possibilités intéressantes. Notre étude s'est limitée à un test utilisateur. Néanmoins, elle a permis de relever une charge cognitive élevée et plusieurs problèmes ergonomiques. Il serait intéressant d'évaluer d'autres utilisateurs handicapés et de contribuer à l'amélioration ainsi qu'à l'accessibilité de ces systèmes.