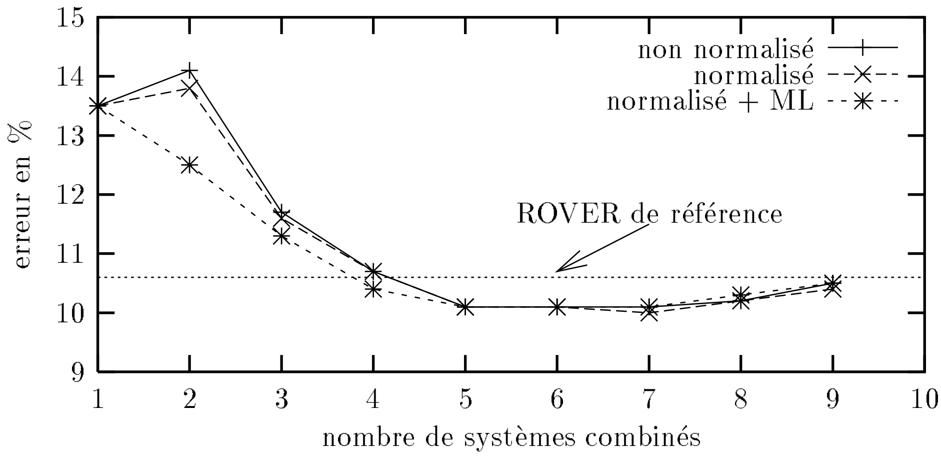


FIG. 1 – *Résumé des principaux résultats. On remarque que la combinaison des cinq meilleurs systèmes améliore les performances (10.1%) par rapport à la combinaison de tous les systèmes (neuf) pris dans l'ordre alphabétique (ROVER de référence : 10.6%). Les performances des systèmes individuels varient de 13.5 à 14.7% pour les cinq meilleurs et montent jusqu'à 25.7% pour le pire. La normalisation des transcriptions avant la combinaison entraîne une légère diminution du taux d'erreur de mots. L'utilisation d'un modèle de langage (ML) peut apporter des améliorations significatives, surtout lorsque peu de systèmes de reconnaissance sont combinés.*

référence :

`... statisticians have a hard time measuring its value. Take fast food. When the first ...`

système 1 (2 erreurs) :

`... statisticians have a hard time` <u>`mission`</u> `its value take fast food` <u>`in`</u> `the first ...`

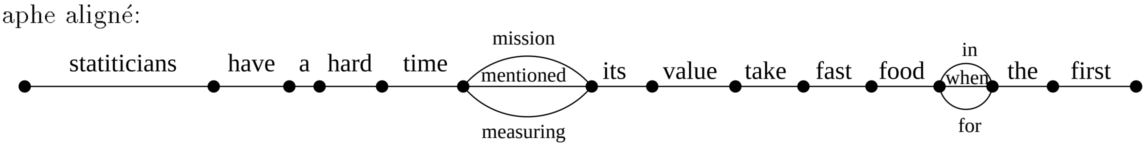
système 2 (1 erreur) :

`... statisticians have a hard time` <u>`mentioned`</u> `its value take fast food when the first ...`

système 3 (1 erreur) :

`... statisticians have a hard time measuring its value take fast food` <u>`for`</u> `the first ...`

graphe aligné:



ROVER de référence (2 erreurs) :

`... statisticians have a hard time` <u>`mission`</u> `its value take fast food` <u>`in`</u> `the first ...`

ROVER avec modèle de langage (1 erreur) :

`... statisticians have a hard time measuring its value take fast food` <u>`for`</u> `the first ...`

FIG. 2 – *Exemple de fonctionnement de l'approche ROVER. Les réponses des trois systèmes de reconnaissance de la parole sont erronées à deux endroits différents. L'alignement des trois transcriptions et la sélection du mot le plus fréquent séparément pour chaque nœud (approche ROVER de référence) ne permet pas de diminuer le taux d'erreur puisque les mots sont équiprobables. La dernière ligne montre comme l'utilisation d'un modèle de langage permet dans cet exemple d'éliminer une erreur (le mot "for" est faussement choisi puisque le système n'a pas la connaissance que le mot précédent "food" termine une phrase).*