

# Comparaison de la méthode de saisie dichotomique avec les méthodes d'épellation traditionnelles

ÉQUIPE DICOTOMIX ENS de Lyon  
dicotomix@listes.ens-lyon.fr

## Résumé

*Les méthodes traditionnelles permettant de communiquer en se limitant à une entrée de la forme oui ou non sont basées sur une épellation lettre par lettre. Ici, une approche différente est considérée, où les mots sont épelés un à un par dichotomie [11] sur l'ensemble du dictionnaire français. Cette méthode est comparée aux méthodes traditionnelles en vigueur dans le milieu médical. Nous montrons que notre méthode permet d'écrire les mots beaucoup plus rapidement, et ce même si l'on présuppose que quatre lettres suffisent à deviner le mot dans les méthodes classiques.*

## I. INTRODUCTION

**L**A communication chez des patients atteints de locked-in syndrom (LIS) ou de la maladie de Charcot est difficile. Elle se fait avec un aide soignant, qui interprète les signaux que le patient parvient à exprimer, que ce soit par des techniques d'eye-tracking [4], de Brain-Computer interface [10] [3] ou autrement. [6]

Les méthodes de saisie traditionnelles utilisées dans le milieu médical reposent toutes sur le même principe. [16] Elles consistent à épeler les mots lettre à lettre, en parcourant l'alphabet pour chacune des lettres. Ce type de méthode peut s'avérer frustrant [16] pour l'utilisateur, qui décide parfois même de renoncer à la communication, tant elle peut être fastidieuse. L'outil informatique s'avère ainsi être capital [7] pour permettre à l'aide-soignant d'aider le patient à communiquer. Une autre approche propose d'utiliser plus efficacement une entrée binaire afin d'écrire non pas une lettre à la fois, mais un mot à la fois. Ainsi, nous mettons ici en avant l'intérêt d'une telle méthode, en les comparant sur l'ensemble du dictionnaire.

## II. MÉTHODES TRADITIONNELLES

L'essentiel des méthodes traditionnelles repose sur une division de l'alphabet en une table en deux dimensions, où la recherche de la lettre voulue se fait en deux temps. [1] Dans un premier temps, on recherche la ligne dans laquelle se situe la lettre voulue, en posant la question "est-ce que votre lettre est dans la première ligne", ou dans la seconde, jusqu'à arriver à la bonne ligne. Ensuite, on itère de la même manière sur les colonnes, en proposant chacune des lettres de la ligne jusqu'à trouver celle voulue.

Code "EJASINT" conseillé par ALIS						
1	E	A	N	R	C	V
2	J	I	L	P	H	W
3	S	U	D	G	K	
4	T	M	B	Z		
5	O	F	X			
6	Q	Y				

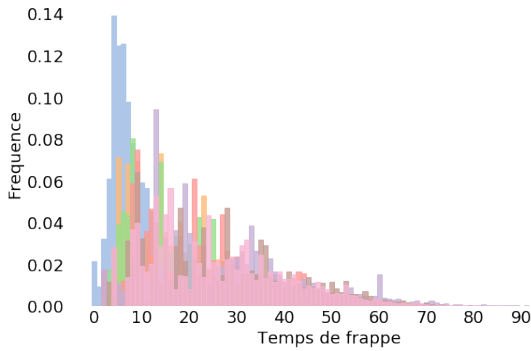
**Figure 1** – Alphabet EJASINT, majoritairement utilisé pour la saisie lettre à lettre par l'association ALIS

Une des méthodes d'épellation les plus répandues utilise l'alphabet EJASINT [5] décrit dans la figure 1. Les autres méthodes envisagées ici sont décrites en annexe, et reposent toutes sur ce même principe de grille de lettres, avec une répartition différente de l'alphabet.

## III. MÉTHODE DICHOTOMIQUE

Notre méthode repose sur un parcours du dictionnaire français par dichotomie, en tenant compte des fréquences des mots dans l'usage courant de la langue française. L'on présente un mot au patient et la question posée est de la forme "votre mot est-il avant ou après dans l'alphabet ?"

Une seconde méthode que nous évoquerons est la méthode de l'épellation dichotomique, reprenant le principe de la dichotomie mais cette fois pour une épellation lettre par lettre : "la lettre recherchée est-elle avant ou après dans l'alphabet ?".



**Figure 2** – Distribution du temps de frappe en fonction de la longueur des mots pour différentes méthodes. Bleu - recherche dichotomique, orange : épellation dichotomique, vert : alphabet EJASINT, rouge : 3, violet : 4, marron : 5, rose pâle : 6

L'intérêt de l'approche dichotomique est très simple : à chaque question, la moitié des possibilités sont éliminées. Cela permet de se rapprocher rapidement de la solution, même dans un grand lexique.

#### IV. ÉPELLATION COMPLÈTE

Nous nous référons, dans cette partie, aux différentes méthodes au travers des numéros suivants : 0 - recherche dichotomique (implémentée dans Dicotomix), 1 - épellation dichotomique, 2 - alphabet EJASINT, les suivants correspondant aux numéros des alphabets en annexe.

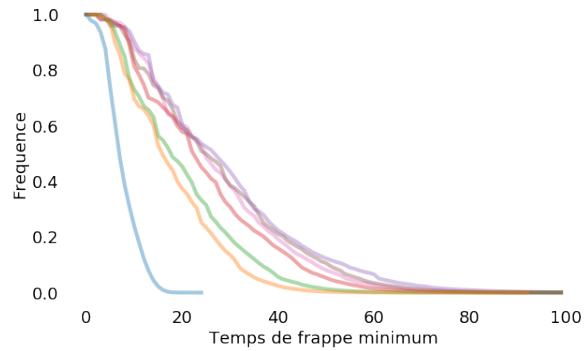
##### i. Distribution des temps de frappe

Pour mesurer l'efficacité de chaque méthode, nous comptons le nombre de questions nécessaire pour trouver les mots. Ainsi, nous considérons chaque mot du dictionnaire et calculons le temps qu'il faut pour le trouver.

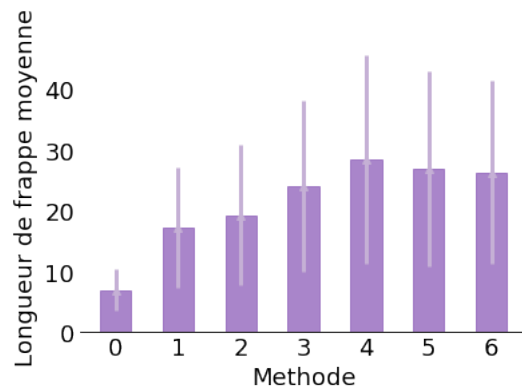
La distribution de ces longueurs de frappe, ajustée à l'aide de la fréquence d'apparition des mots dans la langue française, est un bon indicateur de la rapidité d'utilisation des différentes méthodes. La figure 2 montre que la probabilité d'utiliser peu de coups pour épeler un mot est très élevée avec la méthode Dicotomix. Les autres méthodes sont distribuées sur des temps d'épellation beaucoup plus élevés.

La figure 3 montre la probabilité pour qu'un mot ait une longueur de frappe supérieure à la longueur de frappe donnée en abscisse. Elle confirme l'efficacité théorique de notre méthode car la fréquence d'apparition de mots nécessitant au moins 20 coups pour être tapés est presque nulle, contre plus de 50% pour les autres méthodes.

La méthode d'épellation dichotomique, quant à elle, donne des résultats proches de l'alphabet EJASINT.



**Figure 3** – Proportion de mots ayant un temps de frappe supérieur à un temps donné. Bleu - recherche dichotomique, orange : épellation dichotomique, vert : alphabet EJASINT, rouge : 3, violet : 4, marron : 5, rose pâle : 6



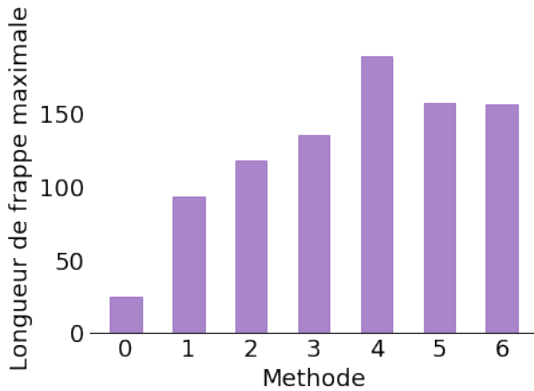
**Figure 4** – Longueur de frappe moyenne sur l'ensemble du dictionnaire (barre violettes), avec l'écart-type de la distribution (petites barres violettes claires). 0 - recherche dichotomique, 1 - épellation dichotomique, 2 - alphabet EJASINT, 3 à 6 - méthodes 3 à 6

##### ii. Moyenne et maximum

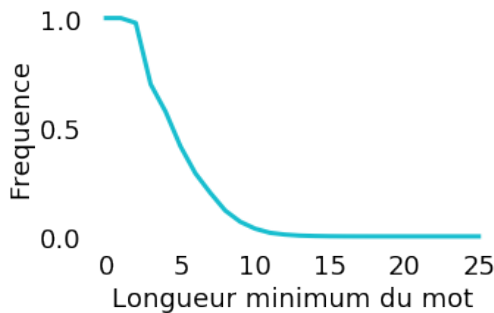
Les distributions donnent beaucoup d'informations sur les méthodes, mais les graphes sont difficilement lisibles. D'un point de vue pratique, on s'intéresse principalement au temps moyen nécessaire pour écrire un mot, et le temps maximal garanti.

La figure 4 montre que la recherche dichotomique a une longueur de frappe moyenne significativement plus faible que toutes les autres méthodes considérées, ce qui signifie que les mots sont en moyenne tapés trois à cinq fois plus vite. De plus, son écart-type est plus faible donc il est peu probable d'avoir à écrire un mot "long".

Même si l'intervention de l'aidant permet souvent de trouver les mots sans épeler toutes les lettres, on peut s'intéresser au pire des cas, c'est-à-dire au mot du dictionnaire le plus long à trouver. On remarque sur



**Figure 5** – Temps de frappe du mot le moins accessible. 0 - recherche dichotomique, 1 - épellation dichotomique, 2 - alphabet EJAINT, 3 à 6 - méthodes 3 à 6

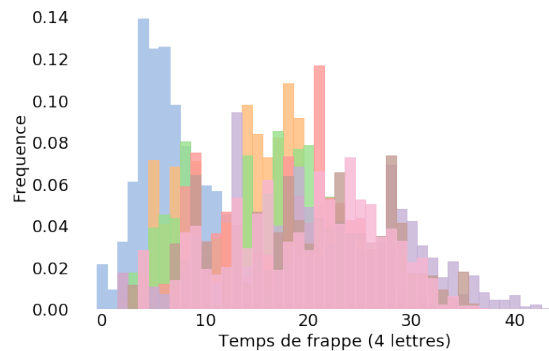


**Figure 6** – Fréquence des mots de longueur supérieure à une longueur donnée dans le dictionnaire

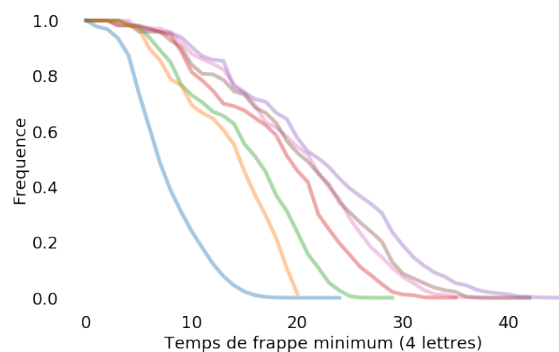
la figure 14 que la dichotomie sur le lexique offre une performance drastiquement meilleure que les autres méthodes, avec un facteur variant de quatre à sept.

## V. ÉPELLATION DES 4 PREMIÈRES LETTRES

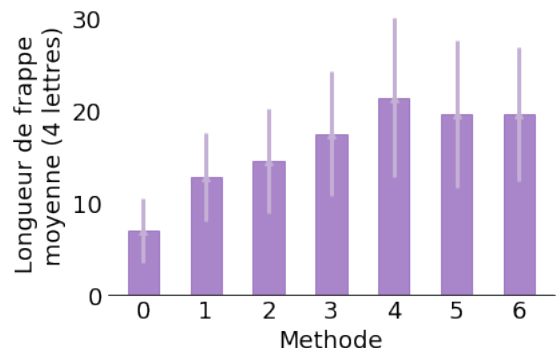
Remarquons que, en fréquence, plus de la moitié des mots du français contiennent moins de 4 lettres (cf figure 6). Par ailleurs, faisons l'hypothèse qu'en pratique, l'on arrive à savoir quel est le mot tapé à partir de la quatrième lettre. Nous reprenons donc dans cette partie la même comparaison des différentes méthodes, cette fois en se limitant à l'épellation des 4 premières lettres de chaque mot. Notons par ailleurs, que la méthode de recherche dichotomique continue, en revanche, à épeler les mots complètement.



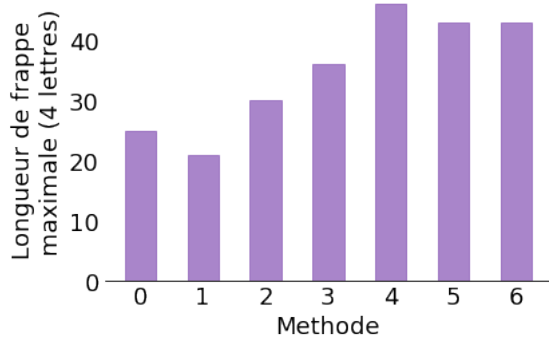
**Figure 7** – Distribution du temps de frappe en fonction de la longueur des mots pour différentes méthodes (épellation des 4 premières lettres uniquement). Bleu - recherche dichotomique, orange : épellation dichotomique, vert : alphabet EJAINT, rouge : 3, violet : 4, marron : 5, rose pâle : 6



**Figure 8** – Proportion de mots ayant un temps de frappe donné (épellation des 4 premières lettres uniquement). Bleu - recherche dichotomique, orange : épellation dichotomique, vert : alphabet EJAINT, rouge : 3, violet : 4, marron : 5, rose pâle : 6



**Figure 9** – Longueur de frappe moyenne sur l'ensemble du dictionnaire (barre violette), avec l'écart-type de la distribution (petites barres violettes claires) (épellation des 4 premières lettres uniquement). 0 - recherche dichotomique, 1 - épellation dichotomique, 2 - alphabet EJAINT, 3 à 6 - méthodes 3 à 6



**Figure 10** – Temps de frappe du mot le moins accessible (épellation des 4 premières lettres uniquement). 0 - recherche dichotomique, 1 - épellation dichotomique, 2 - alphabet EJA SINT, 3 à 6 - méthodes 3 à 6

### i. Distribution des temps de frappe

La distribution des temps de frappe (figure 7) continue à donner la même tendance que précédemment, et la méthode de recherche dichotomique permet toujours de trouver plus rapidement les mots les plus fréquents.

Cette tendance est également confirmée par la fréquence des mots ayant une longueur de frappe supérieure à la longueur donnée en abscisse (figure 8).

### ii. Moyenne et maximum

La longueur moyenne de frappe (figure 9) reste aussi avantageuse par rapport aux autres méthodes. Cependant, la longueur maximale (10) est plus élevée que pour l'épellation dichotomique. Malgré tout, la recherche dichotomique reste meilleure, de ce point de vue, que toutes les méthodes à présent utilisées dans le milieu médical.

## VI. CONCLUSION

La méthode de recherche dichotomique utilisée par le projet Dicotomix est donc une méthode permettant, d'un point de vue théorique, d'écrire de façon significativement plus rapide que les méthodes actuelles. Il faut noter que cette étude théorique ne prend pas en compte :

- Le protocole expérimental qui sera effectivement implémenté avec chaque patient et qui pourrait potentiellement rajouter des questions à la recherche.
- Le coût cognitif supérieur que notre méthode nécessite par rapport aux autres méthodes.
- L'impact plus grand d'une erreur commise par le patient sur la recherche d'un mot ainsi que la nécessité d'être à l'aise avec son orthographe.

Malgré ces obstacles pratiques, notre étude montre son potentiel pour aider les patients à communiquer. Charge à nous de faire diminuer son coût cognitif pour proposer une phase de test stimulante aux patients.

## RÉFÉRENCES

- [1] ALIS. Codes de communication conseillés par alis. 2014.
- [2] D. Morlet P. Ruby O. Bertrand S. Ciancia N. Gallifet J. Luauté J. Mattout E. Maby, M. Perrin. Evaluation in a locked-in patient of the openvibe p300-speller. 2011.
- [3] B. Sorger et al. A real-time fmri-based spelling device immediately enabling robust motor-independent communication. 2012.
- [4] SW Park et al. Augmentative and alternative communication training using eye blink switch for locked-in syndrome patient. 2012.
- [5] V. Gaudeul. Communiquer sans la parole ? guide pratique des techniques et des outils disponibles. 2015.
- [6] A. Goldstein. Combating locked-in syndrome : New methods of communication for als patients. 2006.
- [7] C. Haristoy. L'outil de communication technologique : un moyen de reprendre des décisions pour les personnes atteintes du locked-in syndrome. *Médecine humaine et pathologie*, 2015.
- [8] R. Srinivasan J. Ding, G. Sperling. Attentional modulation of ssvep power depends on the network tagged by the flicker frequency. 2012.
- [9] J. Mattout J. Luauté, D. Morlet. Bci in patients with disorders of consciousness : Clinical perspectives. 2014.
- [10] N. Birbaumer J. Perelmouter. A binary spelling interface with random errors. 2000.
- [11] D. Knuth. *The Art of Computer Programming*, volume 3. Searching and Sorting. Addison-Wesley, 1973.
- [12] H. Ney M. Sundermeyer, R. Schluter. Lstm neural networks for language modeling. 2012.
- [13] S. Taouza N. Karour, S. Ouled Jaafri. Report about akinator. 2014.
- [14] F. Nijboer. Technology transfer of bci as assistive technology : Barriers and opportunities. 2015.
- [15] V. Pidoux. Expérimentation et clinique électroencéphalographiques entre physiologie, neurologie et psychiatrie. 2010.
- [16] J. Terroir. La communication : source d'anxiété ? *IFSI Croix-Rouge de la Métropole Tourcoing*, 2004-2007.

## ANNEXES

Code "EJASINT" conseillé par ALIS						
1	E	A	N	R	C	V
2	J	I	L	P	H	W
3	S	U	D	G	K	
4	T	M	B	Z		
5	O	F	X			
6	Q	Y				

Figure 11 – Alphabet EJASINT, majoritairement utilisé pour la saisie lettre à lettre par l'association ALIS

A	E	I	O	U	Y					
B	C	D	F	G	H	J	K	L		
M	N	P	Q	R	S	T	V	W	X	Z

Figure 12 – Alphabet de la méthode d'écriture 3.

A	Z	E	R	T	Y	U	I	O	P
Q	S	D	F	G	H	J	K	L	M
	W	X	C	V	B	N			

Figure 13 – Alphabet de la méthode d'écriture 4.

Consonnes 1 <sup>re</sup> partie			
B	C	D	F
G	H	J	K

Consonnes 2 <sup>e</sup> partie			
L	M	N	P
Q	R	S	T
V	W	X	Z

Voyelles		
A	E	I
O	U	Y

Figure 14 – Alphabet de la méthode d'écriture 5.

B	C	D	F	G
H	J	K	L	M

N	P	Q	R	S
T	V	W	X	Z

A	E	I	O	U	Y
---	---	---	---	---	---

Figure 15 – Alphabet de la méthode d'écriture 6.