Une disposition universelle des ponctuations sur les claviers des ordinateurs : est-elle possible ?

Sur la standardisation de la pratique des ponctuations sur les claviers des ordinateurs

Mahmoud Fawzi MAMMERI
ESC Koléa, Algérie

Résumé : Le scripteur est perdu entre les différents claviers (QWERTY, AZERTY, arabe), pourtant les ponctuations sont par essence universelles.

Mots-clés : ponctuation, clavier, disposition, QWERTY, AZERTY

1 Introduction

Les méthodes d'entrées utilisées sur les ordinateurs sont basées sur le clavier hérité des machines à écrire. Combien même les claviers des machines à écrire, en premier, et ceux des ordinateurs, par la suite, ont fait l'objet de plusieurs améliorations et normalisations, qui ont abouti à leurs perfectionnements, le clavier moderne n'a pas échappé à diverses critiques de la part des chercheurs, des écrivains ou même des utilisateurs ordinaires. Pour s'en convaincre, on peut citer le cas du clavier AZERTY qui a suscité beaucoup de débats qui ont donné lieu à plusieurs versions de claviers alternatifs à l'AZERTY traditionnel, de la disposition ZHJAY d'Albert Navarre au nouvel AZERTY récemment proposé par l'AFNOR en 2019. Les améliorations apportées au clavier ont touché plusieurs aspects : ergonomie, vitesse de frappe, facilité de trouver les caractères, choix entre saisie de texte « formel » ou de texte émanant du langage courant, facilité pour les manipulations des programmeurs informatiques, facilité pour les utilisateurs de médias sociaux…

L'intérêt que nous portons au clavier concerne la facilité de trouver les caractères. Certes, ce type d'amélioration a été déjà discuté dans divers claviers dans les deux cadres monolingue et plurilingue. Cependant, dans le cadre plurilingue, les améliorations sont restées à un niveau régional et non universel : clavier international *US*, clavier canadien multilingue standard, clavier finnois multilingue, etc. Pour un scripteur multilingue, il n'est pas toujours évident d'apprendre par cœur les dispositions des ponctuations de toutes les langues qu'il utilise. Un exemple assez frappant est celui du point-virgule qui possède des positions différentes dans quatre claviers différents : le QWERTY standard, l'AZERTY traditionnel, le nouvel AZERTY et le clavier arabe 101 touches. Ce qui nous intéresse est non seulement de faciliter l'accès aux ponctuations sur un clavier précis mais de faciliter leurs accès sur l'ensemble des claviers des ordinateurs. Ceci passe par une disposition universelle. Cette idée trouve son origine dans le fait qu'il existe des signes qui sont considérés comme universels par l'ensemble des scripteurs : ponctuations, chiffres, etc. Il serait donc intéressant de reconsidérer les caractères à disposer sur les claviers selon deux sortes de signes : (i) des signes spécifiques à une langue ou groupe de langues (graphèmes alphabétiques : alphabets, diacritiques, etc.) et (ii) des signes communs à l'ensemble des langues du monde (signes de ponctuation, chiffres, opérateurs arithmétiques, etc.). Les premiers sont alors à disposer selon les contraintes de la ou les langues concernées. Les seconds, en revanche, sont universels et sont, par conséquent, à disposer selon un modèle universel commun à toutes les langues : un point, par exemple, serait saisi en arabe, anglais, français… selon une même méthode d'entrée. C'est le cas des chiffres qui sont considérés comme des signes universels et ont bénéficié de positions fixes depuis les premiers claviers des machines à écrire.

La particularité la plus frappante qui caractérise les signes de ponctuation est le fait qu'ils ont en commun d'appartenir à tous les scripts et sont ainsi reconnus par tous les scripteurs sans exception. Cette universalité devrait par conséquent être reflétée dans la pratique de ces signes sur les claviers. Autrement dit, il serait question de chercher les possibilités sur les claviers pour concrétiser cette propriété. Pour répondre à une telle question, nous commençons tout d'abord par admettre une hypothèse, constatée par tous les scripteurs multilingues, qui consiste à dire que les ponctuations n'occupent pas des positions fixes sur les claviers des ordinateurs. Nous allons donc en premier lieu affirmer cette hypothèse et la prouver par des constats et des détails pratiques. Et ce sont ces constats et détails qui vont nous ouvrir la voie vers ce qui est possible sur les claviers des ordinateurs en faveur des ponctuations. Un point qui pourrait aller à l'encontre d'une disposition universelle des ponctuations est le fait que certains scripts ou langues ont déjà développé des instanciations spécifiques pour certaines ponctuations. En arabe, par exemple, on utilise le glyphe miroir « ؟ » pour le point d'interrogation et il faudrait dans une telle situation trouver des moyens pour absorber ces différences dans la forme. Nous verrons que cette question fera intervenir un autre matériel qui est le codage informatique des caractères. C'est aussi le cas de l'espagnol qui utilise deux glyphes « ¿ » et « ¡ » en plus des points d'interrogation et d'exclamation universels, mais ce problème influe plutôt sur le nombre des ponctuations à gérer.

La problématique que nous venons d'exposer est concernée principalement par deux notions : les ponctuations et les claviers des ordinateurs. Nous les introduisons dans

Published in : JAMBOROVA LEMAY, Diana ; PEREIRA, Christophe ; ŠMILAUER, Ivan (dir.), *Des langues calculables à l'homme incalculable. Hommage à Patrice Pognan*, Editions des archives contemporaines, ISBN : 9782813004260, doi : 10.17184/eac.9782813004260

ce qui suit avec trois autres notions qui leur sont associées : caractère, disposition et méthode d'entrée.

1.1 Caractère

Pour écrire dans une langue, on utilise un système d'écriture, qui utilise un ensemble de signes appelés caractères. Un caractère peut être une lettre de l'alphabet, un chiffre (0, 1…), une ponctuation (. : …, ; ! ?), un symbole (= + * – /), un élément de parenthésage ([{ ( ) }]) ou encore des caractères tels que l'espace, la tabulation, le retour à la ligne, etc. Dans une langue bicamérale et utilisant des diacritiques, comme le français, *A*, *a*, *Â*, *â*, *À* et *à* sont six caractères différents.

1.2 Ponctuation

Pour définir la ponctuation, nous distinguons les définitions destinées au grand public des définitions établies dans le cadre des recherches en théorie de la ponctuation. Pour les besoins du texte, une définition grand public nous suffit. Plusieurs définitions existent : le *Bon usage*, le Trésor *TLFi*, la *Grammaire méthodique du français*, etc. Nous retenons celle de la *Grammaire méthodique du français*, qui définit la ponctuation dans un cadre indépendant de toute théorie linguistique :

> La ponctuation est le système des signes graphiques qui contribuent à l'organisation d'un texte écrit et qui apportent des indications prosodiques, marquent des rapports syntaxiques ou véhiculent des informations sémantiques.
>
> (Riegel *et al.* 1996, p. 83)

On retrouve la plupart des signes de ponctuation dans toutes les langues :

> Les signes de ponctuation sont presque tous interlinguaux, encore que chaque langue en fasse un emploi un peu particulier.
>
> (Josette Rey-Debove, 1978 : 46 ; cité dans Ponge, 2011 : 129)

L'ensemble de tous les signes (graphiques) de ponctuation cités par les différents auteurs rassemble : point, virgule, point-virgule, deux-points, point d'exclamation, point d'interrogation, parenthèses, guillemets, crochets, points de suspension, tiret, trait d'union ou division, alinéa, astérisque, paragraphe, point à la ligne, tiret double, apostrophe, *et* commercial, barre transversale, croix, gros points, verset, répons, pied de mouche, main, traits de rappel, blanc, oppositions majuscule/minuscule, gras/italique/romain, titres/intertitres et texte courant, souligné/non souligné, etc.

Cependant, tous les systèmes d'écriture restreignent les signes de ponctuation à un nombre réduit. C'est ce que confirme C. Mulvey :

> Those twelve [full stop, semicolon, colon, comma, slash, hyphen, parenthesis, exclamation, apostrophe, quotation mark, question mark and ellipsis]

Published in : JAMBOROVA LEMAY, Diana ; PEREIRA, Christophe ; ŠMILAUER, Ivan (dir.), *Des langues calculables à l'homme incalculable. Hommage à Patrice Pognan*, Editions des archives contemporaines, ISBN : 9782813004260, doi : 10.17184/eac.9782813004260

> provide the fundamentals of English [all] language punctuation, and all of them do more than one job. If we had a complete and unambiguous set of punctuation marks, we might need as many as 50, but the writing world does not want the trouble of such precision. […] Punctuation marks are ambiguous therefore. They suggest rather than define. They rely on context and the quick wittedness of the reader.
>
> (Mulvey, 2016)

Dans ce qui suit, nous retenons quinze signes universels : blanc, trait d'union, point, virgule, point-virgule, deux-points, point d'exclamation, point d'interrogation, points de suspension, parenthèses, guillemets, crochets, tiret, barre oblique, astérisque.

1.3 Clavier

Dans le domaine de la saisie des caractères, le terme clavier est souvent utilisé pour désigner (i) un dispositif physique, comme périphérique d'entrée standard de l'ordinateur (ii) une disposition des touches sur un clavier physique, mais aussi (iii) une méthode d'entrée pour une langue donnée. Il conviendrait de distinguer chacun de ces termes.

1.3.1 Clavier

Le clavier en tant que dispositif physique est un matériel connecté à un appareil (ordinateurs, smartphones, distributeurs automatiques, calculatrices, télécommandes, machines à écrire, etc.) pour l'alimenter avec des données et en contrôler le fonctionnement : en permettant de saisir des caractères (lettres, chiffres, symboles, etc.) ou des commandes. Malgré l'existence d'une multitude de dispositions (c.-à-d. d'arrangements) des touches du clavier pour différentes langues et différents pays, tous les claviers actuels sur le marché ont une structure identique. Les claviers des ordinateurs ont hérité des claviers des machines à écrire. Mais si la machine à écrire est destinée à la saisie dans une seule langue, l'ordinateur, étant muni d'un langage universel qui est l'électronique, est censé manipuler non seulement n'importe quelle langue mais aussi plusieurs langues à la fois. Ainsi, les claviers des ordinateurs sont censés être plurilingues. C'est ce qui a donné naissance à la notion de disposition logique.

1.3.2 Disposition

Pour saisir un texte dans une langue, on passe plus de temps à saisir des lettres de l'alphabet qu'autre chose (ponctuations, chiffres, etc.). C'est la raison pour laquelle il est important de disposer ces lettres sur les touches du clavier de la manière la plus optimale, c.-à-d. la plus accessible pour le scripteur. Plusieurs paramètres sont pris en compte pour disposer les lettres sur les touches d'un clavier : fréquence des lettres, groupements de lettres dans des digraphes et trigraphes, caractère bicaméral de la langue, etc. La disposition la plus ancienne et la plus connue est la QWERTY. Cette disposition, héritée des claviers des machines à écrire, a été développée pour l'anglais, mais est aussi utilisée dans plusieurs langues utilisant le script latin. Plusieurs dispositions sont apparues par la suite pour remplir les besoins des différents

Published in : JAMBOROVA LEMAY, Diana ; PEREIRA, Christophe ; ŠMILAUER, Ivan (dir.), *Des langues calculables à l'homme incalculable. Hommage à Patrice Pognan*, Editions des archives contemporaines, ISBN : 9782813004260, doi : 10.17184/eac.9782813004260

systèmes d'écriture. Ces dispositions peuvent être classées de différentes manières : selon les langues, les scripts, etc. Selon leur rapport avec la disposition QWERTY, on distingue trois types de dispositions :

- certaines, toujours, en QWERTY, mais, étendues de différentes manières pour remplir des besoins spécifiques à différentes langues (utilisant le latin ou un autre script) : dispositions *US* internationale, anglaise du Canada, Royaume-Uni et Irlande, espagnole, russe, chinoise, etc. ;
- d'autres qui héritent du QWERTY mais qui s'en distinguent par des différences mineures : QWERTZ (claviers germanophones), AZERTY (claviers francophones), ĄŽERTY (clavier lituanien), etc. ;
- d'autres enfin non issues de QWERTY : arabe, grec, arménien, thaï, hébreu, chinois, etc.

1.3.3 Méthode d'entrée

Les méthodes d'entrée ont été introduites dans la période des pages de code, où l'on avait besoin d'introduire des caractères d'une langue autre que la langue de saisie sans changer de page de code. Une fois qu'un alphabet est disposé sur un clavier, tout ce qui reste à disposer se fait selon une méthode qui, en principe, garantit de fournir tout ce qui est nécessaire pour écrire dans une langue donnée : ponctuations, symboles, diacritiques, chiffres, translittérations, etc., mais aussi, dans le cas des textes plurilingues, des caractères utilisés par d'autres alphabets. Par exemple, une méthode d'entrée pour le français va permettre, en plus de tout de ce qu'on peut obtenir directement par un simple appui sur les touches du clavier, de déterminer :

- comment passer des minuscules aux lettres capitales (Shift ou CapsLk) et vise versa ;
- comment obtenir des lettres accentuées (AltGr, touches mortes, touches modificatrices, etc.) ;
- comment affecter plusieurs caractères à une touche unique et comment les saisir (simple appui sur une touche, appui simultané Shift+touche ou appui simultané AltGr+touche).

La notion de méthode d'entrée trouve son origine dans la saisie des caractères du chinois et d'autres langues dont les jeux de caractères, qui se comptent en milliers, excèdent de loin celui dont disposent les périphériques d'entrée et ne peuvent donc pas tous être disposés sur les touches d'un clavier normal. De nombreuses méthodes d'entrée ont été conçues pour permettre de mieux encoder les caractères chinois avec des claviers standards. La méthode la plus connue est le codage du *pinyin* : un logiciel d'entrée permettant de taper les caractères chinois en pinyin[1] et un clavier QWERTY ou AZERTY habituel.

[1] Le pinyin est une transcription du chinois mandarin qui utilise l'alphabet latin.

Published in : JAMBOROVA LEMAY, Diana ; PEREIRA, Christophe ; ŠMILAUER, Ivan (dir.), *Des langues calculables à l'homme incalculable. Hommage à Patrice Pognan*, Editions des archives contemporaines, ISBN : 9782813004260, doi : 10.17184/eac.9782813004260

Avec l'introduction d'Unicode et le développement des systèmes d'exploitation, l'installation de plusieurs méthodes d'entrée et le basculement entre ces méthodes sont de plus en plus facilités. Ainsi, lorsque nous travaillons avec une méthode d'entrée, on peut facilement saisir des caractères dans d'autres langues : (i) soit en invoquant la méthode d'entrée de la langue choisie (ii) soit en utilisant les possibilités offertes par la méthode d'entrée courante sans changer de méthode.

2 Matériels et méthodes

Pour répondre à notre problématique, nous avons opté pour une approche descriptive analytique. Pour pouvoir décrire les claviers actuels ensuite analyser les dispositions proposées, nous avons besoin de consulter trois types d'outils : les dispositions standards, les codages informatiques et les normes établies pour organiser l'utilisation des claviers des ordinateurs dans le cadre de la saisie des textes. Ceci va nous permettre d'extraire des données sur lesquels nous pourrons développer nos descriptions.

2.1 Dispositions logiques

Les dispositions des touches des claviers d'ordinateurs actuels ont hérité de la disposition des touches des machines à écrire, malgré le fait que les problématiques liées à la mécanique des machines à écrire ne soient pas les mêmes que celles liées aux claviers des ordinateurs modernes. Étant intéressé par la saisie des textes plurilingues, nous introduisons trois dispositions importantes par le nombre de scripteurs qui les utilisent et par leur influence sur les autres dispositions utilisées : la QWERTY, l'AZERTY et la disposition arabe.

2.1.1 Clavier QWERTY

Cette disposition date de la fin du XIX^{e} siècle. Elle a été utilisée dans les claviers des machines à écrire. Depuis, elle a résisté à plusieurs tentatives de changement (voir Howells 2005 ; David & Zeitlin 1998 ; Liebowitz & Margolis 1998.) Plusieurs dispositions ont été proposées, chacune ayant apporté son lot d'arguments : DHIATENSOR (an), ZHJAYSCPG (fr), Dvorac (en), Colemak (en), Bépo (fr), Turque en F, etc., mais aucune d'elles n'a été prise au sérieux par les constructeurs qui sont toujours pour la QWERTY (Gardey 1998). La disposition QWERTY a été adaptée à plusieurs langues ou groupes de langues à script latin pour produire des dispositions dérivées : les QWERTZ (langues germanophones), les AZERTY (langues francophones) et plusieurs autres variantes QWERTY pour des versions nationales.

2.1.2 Clavier AZERTY

Pure dérivée du QWERTY, l'AZERTY est la disposition de base utilisée pour l'écriture du français. Cette disposition souffre de plusieurs déficiences dont les capitales accentuées, les ligatures (minuscules ou capitales), les guillemets typographiques, etc. Ceci a motivé l'apparition de plusieurs dispositions concurrentes, même si elles n'ont pas toutes bénéficié d'une reconnaissance commerciale ou d'un intérêt de la part des utilisateurs : ZHJAY, Claude Marsan, Dvorak-fr et Bépo (Pineau, 2011 ; Marsan, 1991). Le dernier né est le nouvel AZERTY proposé par l'AFNOR en 2019.

Published in : JAMBOROVA LEMAY, Diana ; PEREIRA, Christophe ; ŠMILAUER, Ivan (dir.), *Des langues calculables à l'homme incalculable. Hommage à Patrice Pognan*, Editions des archives contemporaines, ISBN : 9782813004260, doi : 10.17184/eac.9782813004260

2.1.3 Clavier arabe

Le caractère cursif de l'arabe n'a pas facilité la tâche des concepteurs de claviers, de codages informatiques ou d'algorithmes de visualisation, et a longtemps inquiété les imprimeurs puis les inventeurs des machines à écrire ; c'est d'ailleurs la raison pour laquelle plusieurs tentatives de simplification du script arabe ont vu le jour pour aboutir à une meilleure prise en charge. Ceci est essentiellement dû à la multitude des formes contextuelles pour chaque lettre arabe, ce qui a donné lieu, au début de l'imprimerie du caractère arabe, à une casse très compliquée dépassant le millier de caractères. Avec l'arrivée d'Unicode, l'essentiel de cette distinction n'est plus reflété dans le codage informatique : elle est purement graphique ; les moteurs de rendu se chargent de choisir la bonne forme selon la position de la lettre dans le mot. La disposition arabe a été héritée, elle aussi, du clavier de la machine à écrire arabe. La disposition de ce clavier a été basée sur l'ordre alphabétique[2] de l'alphabet arabe (*tartīb hiǧāʾī*) qui est un ordre moderne introduit au VIII^{e} siècle et basé sur deux principes : regrouper les lettres selon les ressemblances de leurs formes (*rasm*), ensuite les ordonner selon le nombre de points souscrits ou suscrits (*ʾiʿǧām*). Au début du XX^{e} siècle, l'ordre alphabétique a été repris par Selim S. Haddad pour mettre au point une disposition pour le clavier des machines à écrire de Monarch Corp. Après avoir perfectionné la mécanique du clavier arabe pour le compte de Royal McBee Corp. en 1957, Seyed Khalil[3] a légèrement modifié la disposition initiale de Selim S. Haddad. La nouvelle disposition a été adoptée par l'industrie des machines à écrire et, par la suite, retenue pour la conception des claviers des ordinateurs. De ces deux dispositions historiques, ont découlé deux dispositions principales : l'une pour l'IBM PC, l'autre pour le Mac. Celle d'IBM est disponible en plusieurs variantes : 101 touches, 102 et 102 AZERTY pour le Maghreb et la 101 pour les pays du Mashreq.

2.2 Codage des caractères

Le caractère est un élément important dans la composition des textes. C'est en réunissant des caractères entre eux qu'on obtient des mots ; en insérant certains marqueurs, dits ponctuations, à des endroits précis qu'on obtient des propositions et des phrases… Dans l'imprimerie, les caractères étaient mis dans des casses et il fallait s'occuper de la gestion de ces casses ; c'était la seule contrainte. Dans les machines à écrire, les choses étaient plus contraignantes puisqu'il fallait faire fonctionner une mécanique très compliquée dans un espace très réduit qu'est le clavier : le nombre de caractères était alors très réduit et les caractères étaient sélectionnés minutieusement. Avec l'arrivée des ordinateurs, même si les claviers étaient les mêmes que ceux utilisés dans les machines à écrire, et donc l'espace alloué aux caractères était là aussi toujours aussi réduit, l'électronique des ordinateurs a permis, avec le développement des mémoires, des possibilités pour les caractères qui n'existaient pas auparavant. Au départ, le nombre des caractères codés était de 128 dont 94 seulement sont imprimables[4] ; c'était le code ASCII apparu en 1961. Vingt ans après, ce nombre a doublé avec l'apparition de la page de code CP437 de MS-DOS qui utilise un codage sur 8 bits. Le codage sur 8 bits

[2] L'arabe dispose aussi d'un ordre dit levantin (*tartīb ʾabǧadīy*), dérivé de l'ordre de l'alphabet phénicien.
[3] https://patents.google.com/patent/US2940575A/en
[4] Lettres, chiffres et symboles (ponctuations, opérateurs arithmétiques et caractères spéciaux).

Published in : JAMBOROVA LEMAY, Diana ; PEREIRA, Christophe ; ŠMILAUER, Ivan (dir.), *Des langues calculables à l'homme incalculable. Hommage à Patrice Pognan*, Editions des archives contemporaines, ISBN : 9782813004260, doi : 10.17184/eac.9782813004260

a donné naissance à une multiplication des codages spécifiques et propriétaires pour différentes langues et scripts qui a donné lieu à des incompatibilités. Certes, l'intervention de l'ISO a permis de remettre de l'ordre pour assurer la compatibilité entre différents codages, mais le nombre de caractères codés est resté toujours insuffisant. Et ce n'est que vers le début des années 1990, avec l'apparition du standard Unicode, que les barrières du codage ont été affranchies.

2.3 La norme ISO/CEI 9995 (ISO/CEI, 1994)

L'évolution de l'informatique a facilité la conception de claviers et a donné naissance à diverses dispositions pour différents scripts et langues. Cette prolifération des dispositions n'a pas été sans conséquences sur les scripteurs et notamment les plurilingues qui souvent trouvent du mal à passer d'un clavier à un autre. Ceci a nécessité la normalisation de la disposition des touches sur les claviers. La norme internationale ISO/CEI 9995, présentée sous le titre général *Technologies de l'information — Disposition des claviers conçus pour la bureautique*, est alors apparue pour permettre de standardiser une certaine organisation des claviers des ordinateurs. Parmi les dix parties que comprend l'ISO/CEI 9995 deux intéressent notre propos : la *Partie 1 — Principes généraux pour la disposition des claviers* et la *Partie 2 — Module alphanumérique*. L'ISO/CEI 9995-1 fournit une description des claviers adaptés aux systèmes bureautiques avec identification des parties et définition des concepts de base utilisés dans toute la norme. L'ISO/CEI 9995-2, concerne la géométrie du clavier harmonisé à 48 touches alphanumériques : elle décrit « le module alphanumérique d'un clavier et donne des indications sur la division de ce module en zones. Elle définit également l'agencement, le nombre et l'emplacement des touches de la zone alphanumérique […] ».

Selon cette norme, la distribution des blocs de touches est toujours la même :

- un clavier alphanumérique principal formé de six rangées (de A à F, de bas en haut), la rangée A pour des touches de commande standards, les trois suivantes pour l'alphabet et d'autres besoins du système d'écriture adopté (autres que les chiffres), la cinquième pour les chiffres et autres besoins et la sixième pour les touches de fonctions ;
- un pavé numérique dédié principalement aux chiffres (absent dans certains ordinateurs portables).

Cette norme constitue en quelque sorte un modèle type pour la planification d'une disposition universelle. En effet, avec le clavier normalisé il est possible de mesurer l'écart entre les dispositions actuelles des ponctuations des claviers et ce qui pourrait être une norme universelle.

3 Résultats

Nous allons énumérer ci-dessous toutes les données nécessaires à la discussion de notre problématique. Ces données sont issues des lectures et examens sur les matériels que nous venons de présenter.

Published in : JAMBOROVA LEMAY, Diana ; PEREIRA, Christophe ; ŠMILAUER, Ivan (dir.), *Des langues calculables à l'homme incalculable. Hommage à Patrice Pognan*, Editions des archives contemporaines, ISBN : 9782813004260, doi : 10.17184/eac.9782813004260

3.1 Description du clavier selon la norme ISO/CEI 9995

Selon l'ISO/CEI 9995-1, le clavier est divisé en modules : (i) module alphanumérique (ii) module numérique et (iii) module d'édition et de fonctions. Ces modules sont eux-mêmes sous-divisés en zones. Le module alphanumérique (voir la figure 1) correspond au bloc composé des lettres, ponctuations et symboles spéciaux et chiffres ainsi que des touches modificatrices et de commande. Il est divisé en deux zones :

- zone alphanumérique : toutes les touches produisant des caractères imprimables à l'exception des deux touches Tabulation et Entrée (toutes les touches en fond blanc de la figure 1).
- zone de fonctions : ce sont toutes les touches de fonction plus les deux touches Tabulation et Entrée (touches en fond gris de la figure 1).



Figure 1 : Module alphanumérique d'un clavier harmonisé

Si on veut détailler davantage la zone alphanumérique, on peut la décomposer en trois parties :

- La partie chiffres (E00-E10) : positions fixes pour produire les chiffres ;
- La partie lettres (B01-B07 ; C01-C10 ; D01-D10) : positions fixes réservées aux alphabets ;
- La partie non affectée (B08-B11 ; C11-C12 ; D11-D12 ; E11-E12) : touches non fixées.[5]

Pour ce qui concerne la prise en charge des ponctuations, la norme ISO/CEI 9995-2 normalise leur affectation au module alphanumérique, plus spécialement à la zone alphanumérique, mais ne précise pas des positions particulières à l'intérieur de cette zone. Elle précise plutôt que « l'affectation des caractères graphiques est déterminée par une norme nationale ou l'usage » (ISO/CEI, 2002 : 6). Dans le cas de l'alphabet latin, un certain nombre de prescriptions sont à appliquer :(ISO/CEI, 2002 : 6,7)

[5]Nous n'avons pas pris en compte dans cette partie les touches B00, D13 et E13 puisqu'elles sont dépendantes du type de clavier.

Published in : JAMBOROVA LEMAY, Diana ; PEREIRA, Christophe ; ŠMILAUER, Ivan (dir.), *Des langues calculables à l'homme incalculable. Hommage à Patrice Pognan*, Editions des archives contemporaines, ISBN : 9782813004260, doi : 10.17184/eac.9782813004260

1. Les 83 caractères graphiques invariables de l'ISO 646 doivent être affectés. Il s'agit des caractères suivants : les 26 lettres minuscules [...] ; les 26 lettres majuscules [...] ; les dix chiffres [...] ; le point d'exclamation ; les guillemets ; [...] l'apostrophe ; l'astérisque ; la parenthèse gauche ; la parenthèse droite ; [...] la virgule ; le trait d'union [...] ; le point ; la barre oblique ; les deux points ; le point-virgule ; [...] le point d'interrogation ; [...] l'espace.

2. Les chiffres 0 à 9 doivent être affectés aux touches de la rangée E.

3. L'espace doit correspondre à la barre d'espacement de la rangée A.

3.2 Dispositions des ponctuations sur le QWERTY, l'AZERTY et les claviers arabes

En analysant plusieurs claviers, nous constatons que la plupart des ponctuations sont situées sur deux parties du clavier harmonisé : la partie non affectée[6] et la partie réservée aux chiffres. Le tableau 1 ci-dessous donne les positions des ponctuations sur les claviers que nous avons retenus : QWERTY standard, AZERTY traditionnel, nouvel AZERTY, Arabe 101, Arabe 102 et Arabe 102 AZERTY.

Ponctuations	Dispositions					
	AZERTY Trad.	Nouvel AZERTY	QWERTY Standard	Arabe (101)	Arabe (102)	Arabe (102) AZERTY
Blanc	A03	A03	A03	A03	A03	A03
Apostrophe	-	E08	-	-	-	-
Trait d'union	E06	D11	E11	E11	E11	E11
Point	B08	B07	B09	B09	B09	B09
Point virgule	B08	B10	C10	D10	D10	D10
Virgule	B07	B08	B08	C08	C08	C08
Deux points	B09	B09	C10	C10	C10	C10
Points de suspension	-	B09	-	-	-	-
Point d'interrogation	B07	B07	B10	B10	B10	B10
Point d'exclamation	B10	B08	E01	E01	E01	-
Guillemet dactylog. double	E03	E11	C11	C11	C11	E03, C11
Guillemet dactylog. simple	E04	E11	C11	-	-	E04
Parenthèse ouvrante	E05	E05	E09	E09	E09	E05
Parenthèse fermante	E11	E06	E10	E10	E10	E11
Crochet ouvrant	E05	E05	D11	C03	C03	C03
Crochet fermant	E11	E06	D12	C04	C04	C04
Accolade ouvrante	E03	D05	D11	B03	D11	D11
Accolade fermante	E12	D06	D12	B04	D12	D12
Tirets	-	D11	-	-	-	-
Barre oblique	B09	C11	B10	C09	C09	C09

Tableau 1 : Disposition des ponctuations sur certains claviers (pour un utilisateur plurilingue)

[6] Ceci est le cas général de tous les claviers (que nous n'avons pu citer à cause de leur nombre très important), à une exception près, le Bépo, qui réserve toutes les positions des rangées B-D de la zone alphanumérique aux graphèmes alphabétiques et ainsi ne suit pas le clavier harmonisé établi par la norme ISO/CEI 9995.

Published in : JAMBOROVA LEMAY, Diana ; PEREIRA, Christophe ; ŠMILAUER, Ivan (dir.), *Des langues calculables à l'homme incalculable. Hommage à Patrice Pognan*, Editions des archives contemporaines, ISBN : 9782813004260, doi : 10.17184/eac.9782813004260

Selon les positions occupées par les ponctuations sur les claviers retenus, nous pouvons établir que :

- Certaines ponctuations ne figurent pas sur la disposition logique du clavier harmonisé. C'est le cas des tirets cadratin et demi-cadratin et des guillemets qui sont obtenus par un algorithme de correction automatique. Dans MS Word français, lorsqu'on tape (i) deux traits d'union successifs, ils sont remplacés par un tiret demi-cadratin (ii) quatre traits d'union, par un tiret cadratin et (iii) les guillemets droits, par les guillemets en chevrons doubles. Ces méthodes d'entrée sont généralement paramétrables dans les logiciels de traitement de textes tels que MS Word à partir des options de correction automatique (correction en cours de frappe).
- Certaines autres sont toujours situées soit sur la partie droite de la zone alphanumérique (point, deux points, virgule, point-virgule, etc.) soit sur la partie réservée aux chiffres (parenthèses, trait d'union) même si elles n'occupent pas les mêmes positions dans les différents claviers.
- Certaines ponctuations ont des positions qui ne sont pas les mêmes sur tous les claviers. Le point d'exclamation en est un exemple. Sur les QWZERTY (*US*, *US International*, Espagne, finnois multilingue, etc.), le point d'exclamation se trouve toujours dans la partie réservée aux chiffres en position E01. Sur les AZERTY (standard, nouvel, etc.), il est sur la partie non affectée pour les versions Windows (en B10 sur le standard et en B08 sur le nouvel) et sur la partie réservée aux chiffres en position E08 pour les versions Mac ; une exception vient de l'AZERTY Belge qui le met dans la partie réservée aux chiffres en position E08. Pour les claviers arabes deux cas se présentent : sur les 101 (ANSI) et 102 (ISO), il se trouve sur la partie réservée aux chiffres, alors que le 102 AZERTY ne dispose pas d'une touche dédiée au point d'exclamation !

3.3 Propriétés des ponctuations

Les ponctuations ont certaines propriétés qui pourraient être indicatives quant à leurs dispositions sur les claviers des ordinateurs :

- Les ponctuations sont universelles et en nombre réduit.
- Certaines langues ont des formes différentes des formes latines pour certaines ponctuations. En arabe (et dans d'autres langues qui utilisent le script arabe comme le perse ou l'ourdou), le point d'interrogation, la virgule, le point-virgule, les parenthèses, les crochets et les guillemets ont des formes légèrement différentes de leurs équivalents latins. Par exemple, l'arabe adopte un signe d'interrogation avec effet miroir (؟) par rapport à son équivalent latin. Historiquement, c'est Ahmad Zaki qui a recommandé un ajustement dans la forme de certaines ponctuations universelles pour qu'elles puissent être assorties avec le sens d'écriture de l'arabe (Zaki, 2013 : 11). Un autre exemple vient de l'espagnol qui emploie des signes doubles dans les interrogatives (¿ ?) et les exclamatives (¡ !)

Published in : JAMBOROVA LEMAY, Diana ; PEREIRA, Christophe ; ŠMILAUER, Ivan (dir.), *Des langues calculables à l'homme incalculable. Hommage à Patrice Pognan*, Editions des archives contemporaines, ISBN : 9782813004260, doi : 10.17184/eac.9782813004260

(Parkes, 1992 : 56-57). Dans la tradition typographique de l'espagnol, le premier signe (culbuté) est utilisé pour l'ouverture et le second pour la fermeture.

- Avec le développement des codages, notamment avec l'arrivée d'Unicode, plusieurs ponctuations sont nouvellement codées. Certaines sont reniées. C'est le cas par exemple des espaces (espace justifiante insécable, espace insécable fine, etc.), tirets (demi-cadratin, cadratin, etc.), traits d'union (trait d'union insécable, etc.), guillemets (français, anglais britanniques, anglais américains, allemands…) etc. Ces ponctuations ont déjà existé en typographie de la période de l'imprimerie traditionnelle. Certaines autres ponctuations sont des nouvelles formes pour des ponctuations anciennes, comme c'est le cas du point de suspension qui vient remplacer les trois points. Enfin, d'autres sont carrément nouvelles ; c'est le cas de ponctuations utilisées dans des textes spécifiques telles que l'arobase utilisée dans les adresses mail et le croisillon typographique utilisé comme hashtag pour marquer des contenus sur Internet (principalement sur les réseaux sociaux).

- Dans la manipulation des textes, les ponctuations sont, au même titre que les lettres de l'alphabet, plus sollicitées que les chiffres, alors que ces derniers ont des positions universelles mais non les premiers qui varient d'un clavier à un autre.

3.4 Qu'y a-t-il d'intéressant dans Unicode pour les ponctuations ?

Unicode est bien plus qu'une table de correspondance entre codes et caractères codés. Il identifie chaque caractère par une sémantique. Unicode fournit des tableaux de propriétés de caractère que les programmeurs peuvent utiliser dans différents algorithmes (tri, analyse, etc.) pour différents besoins : quelle est la casse d'un caractère (maj., min., casse de titre Dz) ? Quelle est sa directionalité (gauche, droite, faible, neutre) ? Est-il à symétrie miroir ? Est-il à caractère combinatoire ? *Etc.* (Andries, 2002)

La donnée qui nous intéresse ici est la propriété dite « symétrie miroir » (ou miroir bidirectionnel). L'idée derrière cette propriété est que certains caractères, qui sont utilisés de gauche à droite, s'ils sont réfléchis, peuvent être utilisés de droite à gauche, et inversement. Le cas le plus, mais non le mieux, connu est celui des parenthèses : les seules parenthèses présentes en Unicode ont les numéros U+0028 (ouvrante) et U+0029 (fermante). Or, le glyphe de la parenthèse ouvrante (resp. fermante) n'est pas le même selon que le sens d'écriture est de gauche à droite ou de droite à gauche. Ainsi, le choix de l'ordre d'affichage de ces caractères en arabe, par exemple, est délégué au logiciel qui, muni de la propriété miroir du caractère, sait afficher le bon glyphe dans le bon sens.

4 Discussion

D'après les données collectées ci-dessus, résultats des investigations menées sur les ponctuations, leurs codages et leurs dispositions sur certains claviers représentatifs, il est clair que la propriété d'universalité qui caractérise les ponctuations n'est pas concrétisée sur les claviers des ordinateurs. C'est ce qui explique que pour un scripteur

Published in : JAMBOROVA LEMAY, Diana ; PEREIRA, Christophe ; ŠMILAUER, Ivan (dir.), *Des langues calculables à l'homme incalculable. Hommage à Patrice Pognan*, Editions des archives contemporaines, ISBN : 9782813004260, doi : 10.17184/eac.9782813004260

multilingue, il n'est pas toujours évident de s'habituer aux différentes dispositions des ponctuations des langues qu'il manipule sur son clavier.

Pour connaître les causes de cette déficience au niveau des claviers des ordinateurs, il faut remonter un peu plus loin dans l'histoire des codages informatiques. En effet, les codages des caractères ont une influence directe sur la prise en charge des ponctuations sur les claviers des ordinateurs. Lorsqu'il était impossible de coder plus de 128 caractères (caractères de contrôle compris) sur les ordinateurs à l'époque de l'ASCII, il n'était pas possible, par exemple, de manipuler sur ordinateur des ponctuations comme les tirets, les guillemets typographiques, les points de suspension ou l'apostrophe typographique. Les guillemets typographiques doubles étaient alors remplacés par les guillemets droits (dits dactylographiques) ou confectionnés avec les signes mathématiques *inférieur* et *supérieur* en doubles ; les guillemets typographiques simples étaient remplacés par des apostrophes droites ou confectionnés avec les signes mathématiques *inférieur* et *supérieur* ; les tirets étaient produits avec des traits d'union, etc. Un peu plus tard, les guillemets en chevrons doubles, qui n'ont pas été pris en charge dans le code ASCII, ont été ajoutés dans la page de code 437.[7] Dans Unicode, enfin, les limites du codage n'existent plus : il suffit seulement que le caractère soit pris en charge par les polices de caractères utilisées. Historiquement, le code ASCII a été le premier codage des caractères utilisé sur ordinateur. L'ASCII, malgré le fait qu'il soit très réduit, contient presque tout ce dont on a besoin dans un clavier universel : alphabet, chiffres, ponctuations et symboles. Pour les ponctuations, l'essentiel de ces signes a été fixé depuis l'apparition du code ASCII. Avec l'apparition de la famille des codages 8859-n, les choses ont commencé à basculer avec l'intégration de nouvelles ponctuations :

- universelles, comme les guillemets en chevrons doubles, mais uniquement dans certains codages 8859-n (7 parmi les 15 codages de la famille) ;
- spécifiques à certaines langues ou scripts, c'est le cas de l'arabe (virgule, point-virgule et point d'interrogation) et l'espagnol (point d'interrogation et point d'exclamation culbutés) ;

La seule exception concerne l'espace insécable (NBSP) qui a été fixée universellement à la position A0 dans tous les codages de la famille 8859.

Tout ce qui vient d'être remarqué et expliqué est senti, la plupart du temps, par tout scripteur multilingue à chaque fois qu'il bascule d'un clavier à un autre. La solution la plus simple serait de disposer les ponctuations universelles dans des positions fixes et accessibles comme le sont les lettres et les chiffres, et ceci de façon harmonisée dans tous les claviers d'ordinateurs. Si cette idée d'unifier les positions des ponctuations est séduisante, est-elle faisable techniquement ? En d'autres termes, a-t-on les moyens de la réaliser sans bouleverser les habitudes des scripteurs ? La réponse à cette question fait appel à deux sous questions qui en découlent directement : d'une part, les ponctuations sont-elles toujours en nombre réduit malgré l'existence des ponctuations

[7] La page de code 437 est le premier codage 8 bits proposé par IBM.

Published in : JAMBOROVA LEMAY, Diana ; PEREIRA, Christophe ; ŠMILAUER, Ivan (dir.), *Des langues calculables à l'homme incalculable. Hommage à Patrice Pognan*, Editions des archives contemporaines, ISBN : 9782813004260, doi : 10.17184/eac.9782813004260

spécifiques ? Et d'autre part, y a-t-il des positions libres et en nombre suffisant ? Les ponctuations sont par essence universelles et donc toujours en nombre très réduit : quinze au total. Ceci est confirmé par plusieurs auteurs comme nous l'avons vu ci-dessus. Le problème des ponctuations spécifiques doit en principe être réglé ailleurs que dans le codage : s'il n'est pas possible, du moins dans l'état actuel des logiciels de correction automatique, de manipuler des ponctuations supplémentaires telles que les points d'interrogation et d'exclamation ouvrants en espagnol autrement que par leur codage, ceci ne veut nullement cautionner la prise en charge dans le codage de nouvelles ponctuations pour l'arabe uniquement à cause de leurs glyphes légèrement différents de leurs correspondants latins. En d'autres termes, c'est une erreur que de coder deux fois la même ponctuation : un code pour le latin et un autre pour l'arabe. De plus, certaines ponctuations n'ont généralement pas des touches dédiées sur le clavier harmonisé et sont déterminées par des algorithmes de correction automatique. C'est le cas, par exemple, des tirets (en dash, em dash), des guillemets typographiques et des espaces fines insécables qui sont obtenus avec des corrections automatiques sûres, respectivement, le trait d'union, les guillemets droits et l'espace justifiante ordinaire. Qu'est-ce qui empêchera donc de laisser les moteurs de rendu déterminer le glyphe correspondant au point d'interrogation dans la langue de saisie. Il serait alors plus commode, pour un caractère tel que le point d'interrogation de n'utiliser que le codage universel avec en complément une propriété « symétrie miroir » indiquant que le glyphe du caractère suit le sens d'écriture. Ceci peut être généralisé à d'autres caractères. Ainsi, est justifiée l'hypothèse que les ponctuations universelles sont toujours en nombre très réduit.

Si les ponctuations sont en nombre réduit, il reste à s'assurer que le clavier harmonisé dispose de suffisamment positions libres. Ceci est vérifié sur le tableau 1 ci-dessus qui donne toutes les positions non affectées dans le cadre du clavier harmonisé. De plus, toutes ces positions sont situées dans la partie droite du module alphanumérique (y compris la zone affectée aux chiffres), la partie dans laquelle sont généralement situées toutes les ponctuations universelles. Par conséquent, il n'y a aucun risque de bouleverser les habitudes des scripteurs.

Bibliographie

ISO/CEI. (2002), *Technologies de l'information — Disposition des claviers conçus pour la bureautique — Partie 2 : Module alphanumérique* (ISO/CEI 9995-2). Disponible sur : http://hapax.qc.ca/ISO9995/ISO-CEI\%209995-2-2003.pdf

ISO/CEI. (1994), *Technologies de l'information — Disposition des claviers conçus pour la bureautique — Partie 1 : Principes généraux pour la disposition de claviers* (ISO/CEI 9995-1). Disponible sur : http://hapax.qc.ca/ISO9995/ISO-CEI\%209995-1-1994.pdf

Zaki, A. (2013). *Al-tarqīmu wa 'alāmātuhu fi al-luġati al-'arabiyyat*, Le Caire, Kalimāt li-al-ṭarğamat wa al-našr.

Parkes, M.B. (1992). *Pause and effect. An introduction to the history of punctuation in the West*, Aldershot, Scolar press.

Andries, P. (2002). « Introduction à Unicode et à l'ISO 10646 », *Document numérique*, vol. 6(3), 51-88. https://doi.org/10.3166/dn.6.3-4.51-88

Ponge, M. (2011). « Pertinence linguistique de la ponctuation en traduction (français-espagnol) », *La linguistique*, *47*(2), 121-136. DOI : https://doi.org/10.3917/ling.472.0121

Published in : JAMBOROVA LEMAY, Diana ; PEREIRA, Christophe ; ŠMILAUER, Ivan (dir.), *Des langues calculables à l'homme incalculable. Hommage à Patrice Pognan*, Editions des archives contemporaines, ISBN : 9782813004260, doi : 10.17184/eac.9782813004260

Riegel M., Pellat J.-C. & Rioul R. (2018). *Grammaire méthodique du français*, 7e édition, Paris, Presses Universitaires de France, xxiii+646 p.

Mulvey, C. (2016). The English Project's History of English Punctuation : The Revised Text of a Lecture Given on 13 October 2015, English Language Day, at the University of Winchester, UK. *English Today, 32*(3), 45-51. DOI : https://doi.org/10.1017/S0266078416000110

Andries, P. (2002). « Entretien avec Ken Whistler, directeur technique du consortium Unicode », *Document numérique*, *6*(3), 329-351. DOI : https://doi.org/10.3166/dn.6.3-4.329-351

David Paul A., Zeitlin Edith. (1998). « Comprendre les aspects économiques de QWERTY : la contrainte de l'histoire », *Réseaux*, volume 16, n° 87. Les claviers. p. 9-21. DOI : https://doi.org/10.3406/reso.1998.3160

Margolis, S. & Liebowitz, S. (1998). « La fable du clavier », *Réseaux*, 87(1), 23-44. URL : https://www.cairn-int.info/revue-reseaux1-1998-1-page-23.htm

John Howells. (2005). *The Management of Innovation and Technology. The Shaping of Technology and Institutions of the Market Economy*, SAGE Publications.

https://pdfs.semanticscholar.org/4246/00772ccf5615d37fb61952942ca40975a508.pdf

Gardey, D. (1998). « La standardisation d'une pratique technique : la dactylographie (1883-1930) », *Réseaux*, 87(1), 75-103. URL : https://www.cairn-int.info/revue-reseaux1-1998-1-page-75.htm

Martine Pineau. (2011). « La main et le clavier : histoire d'un malentendu », ILCEA, *14*. DOI : https://doi.org/10.4000/ilcea.1067

Marsan Claude. (1991). « La fin de l'AZERTY ! ? », *Linx*, hors-série n° 4. Texte et ordinateur. Les Mutations du Lire-Écrire. p. 297-302. DOI : https://doi.org/10.3406/linx.1991.1206

Published in : JAMBOROVA LEMAY, Diana ; PEREIRA, Christophe ; ŠMILAUER, Ivan (dir.), *Des langues calculables à l'homme incalculable. Hommage à Patrice Pognan*, Editions des archives contemporaines, ISBN : 9782813004260, doi : 10.17184/eac.9782813004260