

Pourquoi un dictionnaire a donné deux livres : la phonologie chinoise et la réforme de l'écriture dans le cadre du projet Demopaedia

par Nicolas Brouard, chercheur émérite, Institut national d'études démographiques (Ined)
15 décembre 2025

*Ce texte résume ce que nous avons du apprendre sur les révisions de l'écriture et de la phonologie chinoises sur une période de plus de 120 ans, et ainsi les raisons pour lesquelles le volume chinois de l'édition unifiée du **Dictionnaire démographique multilingue** est proposé deux variantes : l'une en caractères simplifiés et l'autre en caractères traditionnels.*

Il n'était pas initialement prévu de publier deux livres, mais au fur et à mesure de l'avancement du projet, il est devenu évident que deux versions seraient nécessaires pour produire des éditions imprimées du dictionnaire. Bien que les textes du dictionnaire soient identiques, à l'exception des variantes de caractères, la difficulté est apparue lors de la création de l'index. En raison de l'abandon progressif de l'écriture manuscrite au profit des ordinateurs et des smartphones, la classification traditionnelle par nombre de traits est tombée en désuétude. Nous avons dû trouver des méthodes de classification plus appropriées. Mais nous avons dû constater que chaque méthode de classification proposée était spécifique à la variante de caractère. Ce faisant, nous avons beaucoup appris sur le développement de deux systèmes différents : le pinyin, principalement utilisé en Chine continentale, et le zhuyin, utilisé par des communautés chinoises hors de Chine continentale.

Ce volume chinois de l'édition unifiée du Dictionnaire démographique multilingue diffère des première (1982) et deuxième (1992) éditions en ce que la traduction anglaise de chaque terme ou expression démographique chinois n'apparaît plus dans le texte des neuf chapitres du dictionnaire, mais uniquement dans l'index. À l'origine, et conformément à l'esprit de la Commission de terminologie des Nations Unies du milieu des années 1950, chaque paragraphe d'un volume devait être rédigé librement afin d'introduire les différents termes ou expressions démographiques propres à une édition particulière du dictionnaire (première édition en 1958, deuxième en 1975, et édition unifiée en 2013).

Le meilleur exemple de rédaction indépendante est l'édition allemande de 1987, où les auteurs ont rédigé des textes entièrement nouveaux à partir de l'édition de 1960. C'est à partir de cette seconde édition allemande, qui a introduit de nouveaux termes démographiques empruntés à l'épidémiologie, qu'a été construite l'édition dite unifiée, intégrant tous les termes omis dans les secondes éditions en anglais, français, espagnol et arabe, entre autres. À ce jour, l'édition allemande, tout comme les éditions coréenne, japonaise, thaïlandaise, malaise, arabe et russe, n'est toujours pas unifiée en raison de termes manquants dans d'autres langues.

Demopædia offre aux internautes un accès web à tous les dictionnaires publiés, permettant une navigation aisée d'une langue à l'autre et d'une édition à l'autre. Toutefois, il est également important de préserver ce qui a fait le succès de ce dictionnaire depuis ses premières éditions en 1958 : un **petit livre imprimé doté d'un index complet et facile à consulter** pour la recherche d'expressions démographiques. Cependant, avec la numérisation et l'abandon progressif de l'écriture manuscrite par les jeunes générations, nous avons rencontré un problème majeur pour la production de l'index du volume chinois. De nombreux dictionnaires chinois anciens utilisaient des index classés par nombre de traits du premier caractère, puis du deuxième, et ainsi de suite. Or, si l'on demande à un jeune Chinois de compter les traits d'un caractère de plus de dix traits, il y a de fortes chances qu'il se trompe, car il ne sait peut-être plus comment l'écrire

correctement. Des enquêtes du China Youth Daily (« De moins en moins d'opportunités », 2017) ont confirmé une crise de l'écriture manuscrite chez les jeunes Chinois, avec des chiffres proches de 60 à 80 % concerna¹nt les difficultés liées à l'oubli des caractères.

La saisie sur ordinateur ou smartphone dispense de la connaissance de l'écriture chinoise. C'est pourquoi le plus jeune auteur de cet ouvrage en chinois a proposé de classer les termes ou expressions démographiques selon leur transcription phonétique en caractères latins, le *pinyin*. Ainsi, l'expression 701-1 « croissance démographique » (人口增长), dont la transcription *pinyin* est « rén kǒu zēng zhǎng », est classée avant l'expression 105-2 « politique démographique » (人口政策), « rén kǒu zhèng cè », car la lettre « e » précède la lettre « h » dans l'alphabet latin.

Cependant, lors de la présentation du prototype de ce dictionnaire à un étudiant chinois en démographie et histoire, celui-ci nous a indiqué qu'il connaissait mal les caractères simplifiés et n'utilisait pas la transcription *pinyin*. Il avait plutôt recours à la transcription phonétique *zhuyin*, basée sur 37 caractères ordonnés inventés en 1913 en Chine continentale. Les quatre premiers caractères, ㄅ ㄆ ㄇ ㄈ, représentent les sons b, p, m et f, et sont communément appelés *bopomofo*.

La figure 1 présente un tel clavier, facilement paramétrable sur un smartphone (nous avons ajouté la phonétique latine). Pour saisir 人口, appuyez successivement sur les quatre touches : « ㄇ |r », « ㄣ |en », « ㄎ |k », « ㄣ |ou ». Après la saisie de seulement trois caractères, le dictionnaire sélectionné proposera 人口 (il peut être nécessaire de saisir les tons pour lever toute ambiguïté). De même, si vous choisissez un clavier *pinyin*, appuyez successivement sur « renk », et sans avoir besoin d'appuyer sur « o » et « u », 人口 vous sera proposé.

Figure 1 Clavier Zhuyin ou Bopomofo (iPhone) avec phonétique latine



¹ Feinuo Sun, actuellement professeure assistante de l'Université du Texas à Arlington, a ainsi proposé à Xiaochun Qiao, professeur honoraire de démographie à l'Université de Pékin et principal auteur de ce volume suite au décès de Yeuchung Yu, démographe chinois de la Division de la Population aux Nations Unies.

Ainsi, nous devons prendre en compte les différents lecteurs potentiels, en chinois simplifié (avec les index *pinyin*) et en chinois traditionnel (avec les index *zhuyin*).

Mais cette transcription phonétique ne servant qu'à la prononciation et non à la compréhension d'un mot ou expression chinoise, nous aurions pu les omettre. Cependant, si le logiciel commet une erreur de transcription, l'index trié selon une phonétique erronée ne sera pas correctement trié. Ni Google Traduction ni Word, par exemple, ne proposent des transcriptions d'une qualité suffisante. De plus, bien que chaque caractère corresponde à un code Unicode, il n'existe pas d'ordre de tri universel pour les caractères chinois, car la prononciation dépend du contexte.

Par exemple, 都会区 (307-4) a été initialement transcrit comme « dōu huì qū », mais la prononciation correcte est « dū huì qū » (ㄉㄨ ㄏㄨㄟˋ ㄑㄩˊ et non ㄉㄡ ㄏㄨㄟˋ ㄑㄩˊ en *zhuyin*), car la signification du caractère 都 dépend du contexte : « tous, les deux » ou, dans ce cas, « métropole, zone ».

Chyong Fang Ko, chercheur honoraire à l'Academia Sinica de Taïwan, a aimablement accepté de vérifier et de corriger les transcriptions phonétiques. Ces corrections ont été intégrées à un programme Python très long, spécialement conçu pour traiter tous les textes WIKI saisis par les auteurs en chinois simplifié et pour générer les tables SQL nécessaires aux index (voir le programme open source sur le [dépôt GitHub Demopaedia](#) de Joseph Larmarange).

Ainsi, le premier volume comprend une annexe avec un index des mots du dictionnaire en chinois simplifié, classés par *pinyin*. Afin d'aider le lecteur à prononcer les termes, nous avons ajouté leur transcription tonale *en pinyin*. Puisqu'il existe une correspondance stricte entre les deux transcriptions phonétiques, il aurait suffi d'ajouter la transcription *en zhuyin* après le *pinyin*. Cependant, le classement alphabétique en *pinyin*, qui ne prend en compte que les lettres, est totalement différent du classement phonétique en *zhuyin*. Par exemple, « croissance démographique » (人口增长) est classé avant « politique démographique » (人口政策) en *pinyin*. En *zhuyin*, « croissance démographique » est classé après « politique démographique » car le son ㄗ (zh) précède ㄗˊ (z).

Les caractères bopomofo sont triés en colonnes en partant de la gauche du clavier (et en utilisant la transcription latine au lieu des caractères zhuyin) :

- Initiales : bpmf, dtnl, gkh, jqx, zh-ch-sh-r, zcs
- Finales : aoe ê ; ai ei ao ou; un en ang eng; euh
- Médiales : i, u, ü

Zhuyin écrit toujours les médiales explicitement (兄 frère x+ü+eng: ㄒㄩˊㄣˊㄍㄨㄥˊ → initiale + médiale + finale), tandis qu'en pinyin, ces voyelles sont souvent fusionnées en une seule voyelle finale (par exemple, xiōng, écrit « x » + « iong », masque la voyelle ü ㄩˊ médiane). Cela s'explique aussi par le fait que le pinyin est limité aux 26 lettres latines, contre 37 pour le zhuyin.

Une solution aurait consisté à créer un second index trié en *zhuyin*, mais le volume en chinois devait rester de taille limitée. De plus, comme les utilisateurs *de zhuyin* lisent et préfèrent souvent les caractères chinois traditionnels, il semblait préférable de « fabriquer » et de publier un second ouvrage en caractères traditionnels. Nous disons « fabriquer » car, bien qu'il soit possible de spécifier la variante pour une meilleure lecture à l'écran (cyrillique ou latine pour le serbo-croate, ou traditionnelle ou simplifiée pour le chinois), en ajoutant l'option Wikipédia « &variant=_zh_hant » à la fin de chaque URL, « _zh_hans » étant la valeur par défaut (voir par exemple <http://zh-ii.demopaedia.org/w/index.php?title=10&variant=zh-hant>), une telle solution n'existe pas pour un livre imprimé. Et comme nous proposons un second ouvrage, plus long et destiné à un public de sinophones beaucoup plus restreint, nous y avons également ajouté un index « traditionnel », classé par nombre de traits. Le nombre maximal de traits dans

ce dictionnaire est de 24 pour les caractères traditionnels, contre 15 pour les caractères simplifiés.

Dans l'index trié par nombre de traits, « politique démographique » (人口政策) précède « croissance démographique » (人口增長), car bien que les deux expressions soient regroupées dans la catégorie des caractères à deux traits en raison du premier caractère 人, les troisième et quatrième caractères 政 et 增 comportent respectivement 9 et 15 traits. Il convient également de noter que le caractère traditionnel 長, qui comportait 8 traits, a été simplifié en 长, qui n'en compte plus que 4.

Avec 15 traits, les locuteurs chinois peuvent se tromper dans le comptage ; c'est pourquoi nous espérons que l'index *zhuyin* sera utile. Enfin, et comme pour les deux éditions précédentes (1982 et 1992) du volume chinois, nous avons ajouté un index trié en anglais. Grâce à la correspondance entre les numéros et les termes d'une édition unifiée dans notre base de données, il est possible de proposer l'index de l'édition anglaise unifiée et d'y joindre la traduction chinoise de chaque terme. Nous pourrions tout aussi bien créer un index trié en français, en espagnol ou en italien pour ce volume chinois unifié, puisque ces trois volumes sont également unifiés.

L'index anglais du dictionnaire traditionnel contient les transcriptions phonétiques *zhuyin* et *pinyin*, tandis que l'index du dictionnaire simplifié ne contient que la transcription *pinyin*. Bien que le *pinyin* et l'écriture simplifiée soient largement utilisés en Chine continentale, l'écriture traditionnelle et l'alphabet phonétique *zhuyin* sont enseignés dès l'école primaire à Taïwan. Bien entendu, des cours de *pinyin* sont proposés aux étudiants des grandes universités, comme l'Academia Sinica, pour ceux qui ont besoin de communiquer avec la Chine continentale. En revanche, les jeunes Chinois de Chine continentale lisent rarement les caractères traditionnels et ignorent généralement tout du système phonétique *zhuyin* et de ses origines, bien qu'il ait été inventé en Chine continentale et non à Taïwan.

Une brève histoire de l'écriture chinoise moderne

Nous souhaitons savoir si les progrès technologiques rapprocheraient les deux systèmes d'écriture et si notre dictionnaire, publié dans les deux variantes, resterait pertinent à l'avenir. Revenons sur l'histoire de deux réformes fondamentales : la phonologie chinoise moderne (1910-1920) et la simplification des caractères (1950-1955).

La première étape consista à démontrer qu'une langue vernaculaire chinoise pouvait remplacer le chinois classique, à l'instar de Dante Alighieri (1304-1321) qui, dans la **Divine Comédie**, avait démontré que l'italien toscan pouvait supplanter le latin, langue savante et religieuse. Les langues vernaculaires chinoises commencèrent à remplacer le chinois classique au début du XXe. Le mouvement Baihua (voir Gao 2018, Weng 2020) encouragea l'usage du chinois écrit vernaculaire, qui se développa et se diffusa grâce aux revues et aux livres entre 1910 et 1920, principalement à partir d'un dialecte du nord, le mandarin. Une seconde étape suivit : la standardisation de la prononciation. S'appuyant sur la tradition phonologique Wade-Giles (Giles 1892), le système phonétique *zhuyin* visait à clarifier et à simplifier la structure des syllabes mandarin en marquant explicitement les composantes initiale, médiane et finale.

En 1919, le Mouvement du 4 mai devint une force nationaliste et anti-impérialiste majeure de l'histoire chinoise moderne. Il fut déclenché par l'indignation populaire suscitée par les termes du traité de Versailles, qui accordait au Japon le contrôle des concessions allemandes dans la province du Shandong au lieu de les restituer à la Chine. Ce faisant, les pays de l'Entente ne reconnurent pas le rôle de la Chine dans la Première Guerre mondiale. La Chine avait fourni plus

de 140 000 travailleurs pour soutenir l'effort de guerre de l'Entente. Les Chinois considérant leur mise à l'écart comme une atteinte à leur souveraineté ne signèrent pas le traité. À Pékin, des manifestations étudiantes éclatèrent le 4 mai 1919, centrées sur la place Tiananmen, donnant lieu à des appels à un changement politique et social. Le Mouvement du 4 mai est considéré comme un moment charnière de l'histoire chinoise, jetant les bases de la transition de la Chine vers le socialisme (Boissoneault 2017).

Dans le cadre de ce mouvement, *le baihua* et *le zhuyin* ont transformé l'éducation et unifié la langue chinoise, jetant ainsi les bases du mandarin moderne. Cependant, le système phonétique du *zhuyin* est aujourd'hui quasiment inconnu des jeunes Chinois de Chine continentale. Bien que son influence ait été limitée par la multitude de dialectes, *le zhuyin* et *le pinyin* ont tous deux démontré leur efficacité à Taïwan et en Chine continentale pour promouvoir le mandarin dans le système éducatif.

L'héritage du *zhuyin* est reconnu dans de nombreux ouvrages et entretiens, notamment à partir de 1999 par Zhou Youguang (décédé en 2017 à l'âge de 100 ans), principal auteur de la transcription *pinyin* en 1955. Ce dernier semble avoir voulu rappeler à la Chine continentale l'importance du *zhuyin* et reconnaît même son adoption à Taïwan : « Le *pinyin* n'est pas né de rien, le *zhuyin* en est l'un des prédécesseurs » (Zhou Youguang, *entretien du 10 février 2005*). « Sans les symboles du *zhuyin*, la Chine n'aurait pas pu s'appuyer sur des décennies d'expérience dans l'enseignement de la phonétique ; le *zhuyin* a établi le principe d'une séparation claire entre initiales et finales ; son succès démontre que l'annotation phonétique peut coexister avec les caractères chinois et ne menace pas leur existence » (Zhou 2003). « Taïwan a préservé le *zhuyin*, où il constitue une tradition culturelle très précieuse. Il y est utilisé avec une grande efficacité. » (Entretien avec Zhou Youguang, CCTV, 2003).

Cependant, le *pinyin* se heurta à la difficulté de compléter l'alphabet latin, car il fallait éviter de confondre trois consonnes absentes des pays latins (à l'exception du son allemand « ich ») avec les sons déjà présents en *pinyin* : zh, ch et sh. Plutôt que d'inventer de nouveaux caractères étrangers à l'alphabet latin ou des signes diacritiques comme č, š et ž, Zhou Youguang préféra s'inspirer de l'expérience soviétique des années 1920-1930, où le ***Latinxua Sin Wenz*** (*Nouveau système de latinisation de l'écriture*), inventé par des linguistes russes comme N.A. Semenov et des linguistes chinois comme Qu Qiubai, avait été testé avec succès auprès de centaines de milliers de travailleurs chinois en Union soviétique et dans les régions sous influence communiste. À cette époque, les caractères latins étaient considérés comme *l'écriture internationale du prolétariat*, et non le cyrillique.

À première vue, réutiliser des lettres consonantiques latines pour représenter des sons spécifiques au mandarin peut sembler audacieux. Pourtant, c'est précisément ce qu'a fait LatinXua Sin Wenz — avec succès et de manière systématique — en attribuant de nouvelles valeurs phonétiques à trois lettres : J, Q et X.

En LatinXua, **J** correspond à l'initiale du Zhuyin ㄐ, une consonne « [affriquée alvéolo-palatale sourde](#) », comparable au son initial de *Dieu* en français, ou au *dj* de *adjacent* et au *j* de *jump* en anglais. **Q** correspond à ㄑ, sa forme aspirée, semblable au *ti* de *tien* ou *question* en français, et proche du son *ch* de *cheese* en anglais. **X** représente ㄒ, une « fricative alvéolo-palatale sourde », comparable au *ch* de *chat* en français, au *ssi* de *mission* ou au *sh* de *sheep* en anglais.

Contrairement aux systèmes de romanisation antérieurs, le LatinXua a conservé les quatre signes diacritiques tonaux hérités du Zhuyin, permettant ainsi de marquer explicitement les tons sans modifier l'orthographe de la syllabe. Ceci contraste avec le Gwoyeu Romatzy, adopté officiellement par la République de Chine en 1928 comme système de romanisation officiel du chinois standard, qui encodait les tons par des alternances orthographiques – souvent par redoublement des voyelles – ce qui offrait une plus grande précision phonologique, mais une

transparence moindre. Par exemple, les syllabes *mā, má, mǎ, mà* (妈,麻,马,骂) étaient transcrites *mha, ma, maa, mah*.

À cet égard, LatinXua anticipait l'un des principaux atouts adoptés plus tard par le pinyin (*Hanyu Pinyin*) : la réutilisation économique de l'alphabet latin pour représenter la phonologie du mandarin de manière claire, cohérente et avec une complexité visuelle minimale, démontrant ainsi qu'un système basé sur le latin pouvait être à la fois pédagogiquement efficace et linguistiquement robuste.

Après avoir décrit la genèse de ces deux ouvrages de l'édition unifiée du Dictionnaire démographique multilingue et les raisons qui nous ont poussés à publier un dictionnaire court en *caractères simplifiés* et un dictionnaire plus long en *caractères traditionnels*, nous pouvons nous interroger sur l'avenir de ces deux systèmes d'écriture dans un monde de plus en plus numérisé. Il semble probable que, du fait de leur proximité structurelle (langue, phonétique), les deux systèmes perdureront.

L'écriture chinoise simplifiée remonte à soixante ans, tandis que l'écriture traditionnelle existe depuis plusieurs milliers d'années. Cependant, les taux d'analphabétisme sont aujourd'hui très faibles dans les deux régions, ce qui prouve l'efficacité des systèmes phonétiques. Parmi les personnes analphabètes, la possibilité de transcrire vocalement la plupart des dialectes parlés en mandarin sur un smartphone pour échanger des SMS ou sur des réseaux sociaux comme WeChat rend inutile la connaissance du *pinyin* ou du *zhuyin* !

Des barrières politiques persistent au niveau étatique, mais aussi au niveau des fabricants de smartphones. En effet, il est possible de converser en chinois simplifié sur de nombreux smartphones Android en saisissant du texte sur un clavier *zhuyin* (comme celui présenté ci-dessus), mais cela est étrangement impossible sur un iPhone !

Les dictionnaires Demopædia, disponibles en ligne et en version imprimée, sont publiés sous licence Creative Commons ShareAlike.

Les versions imprimées sont accessibles via Lulu.com, un service d'impression à la demande, en recherchant « Demopaedia ». Vous pouvez également télécharger les versions PDF ou EPUB des dictionnaires en chinois et dans d'autres langues en cliquant sur le [lien de téléchargement](#) de demopaedia.org. Si votre navigateur modifie automatiquement l'URL de http à https, veuillez la modifier manuellement pour rétablir http afin d'y accéder.

Références

Boissoneault, L. (2017). *The surprisingly important role China played in WWI*. *Smithsonian Magazine*. Special report: *World War I: 100 years later*.
<https://www.smithsonianmag.com/history/surprisingly-important-role-china-played-wwi-180964268/>

Gao, Y. (2018). *The Baihua Movement and ideological revolution* (pp. 73–90). In *The birth of twentieth-century Chinese literature*. Palgrave Macmillan.
https://doi.org/10.1057/978-1-137-55936-4_4

Giles, H. A. (1892/1982). *A Chinese–English dictionary*. Kelly & Walsh.

Weng, J. (2020). *Vernacular language movement*. In *Oxford bibliographies in Chinese studies*. Oxford University Press.
<https://doi.org/10.1093/OBO/9780199920082-0180>

Zhou, Y. (1999). *Hanyu pinyin fang'an de lishi* [《汉语拼音方案的历史》; *History of the Hanyu Pinyin scheme*]. Commercial Press.

Zhou, Y. (2002). *Wenzi gaige liushi nian* [《文字改革六十年》; *Sixty years of script reform*]. Beijing Language and Culture University Press.

Zhou, Y. (2003). *Zhou Youguang wenji* (Vol. 3) [《周有光文集》第三卷; *Collected works of Zhou Youguang*]. Zhonghua Book Company.

Fewer and fewer opportunities to write—Do you still care if your words look good? (2017, March 14). *People's Daily Online*.

<http://edu.people.com.cn/n1/2017/0314/c1053-29144147.html>