

L'écriture chinoise, 3000 ans d'analyses et rationalisations

Abstract

Since their early creation, each Chinese characters -which is basically a visual non-phonetic representation of a concept- have both evolve and been duplicated in multiple variants. Across history, governments have continuously try to standardize their shapes and forbidden variants in order to keep the Chinese plain united by those thousands characters. On their side, individuals have done their best to analyses those characters, and create databases such as etymological ones, semantic ones (dictionaries), phonetic ones (zhuyin, pinyin, hiragana, hangeul), graphic ones (calligraphic dictionaries), which are now join by digital encodage tables, geometric descriptions, and foreign languages fields. It's actually the full set of meta-data relate to Chinese characters which is now expand, and link together by computing possibilities.

This article exposes quickly the most noticeable of those changes and achievements made the last 3.000 years in Chinese characters description, with a special focus on recent developments (20th century).

Table of Contents

L'écriture chinoise,.....	1
3000 ans d'analyses et rationalisations.....	1
Abstract.....	1
Ère archaïque et impériale.....	5
L'époque archaïque et la multitude des variantes.....	5
Dynastie Qin et Han : les premières grandes rationalisations.....	6
Réforme de Li Si.....	6
Styles Lishu et Kaishu : vers une écriture de traits.....	7
Lexicographies impériales : une approche graphique et combinatoire.....	8
Xu Shen et le Shuowen Jiezi	8
Dictionnaire Kangxi.....	8
Ères moderne.....	10
Premières propositions de phonétisations.....	10
Phonétisation chinoise par associations.....	10
Premiers contacts avec systèmes phonétiques.....	10
Propositions modernes.....	10
Premiers usages à grande échelle.....	11
Réformes post-1945.....	12
La question phonétique en RPC.....	12
Enjeux.....	12
Phonétisation de 1956 et 1958 : vers le Hanyu Pinyin.....	12
La question de la simplification.....	13
Enjeux.....	13
Bases historiques.....	13
Réforme japonaise.....	13
Réforme chinoise.....	14
La question de l'ordre des traits.....	15
Enjeux.....	15
Historique des publications.....	15
La question graphique.....	17
Enjeux: Développement de l'édition et besoin de standardisation graphique.....	17
Le cas taïwanais : l'établissement de graphies officielles.....	17
Le cas chinois : l'établissement de graphies officielles, les documents de références.....	19
La question des Encodages.....	19
Premières tentatives et tentatives nationales.....	19
Unicode.....	20
Les nouvelles descriptions géométriques et informatiques.....	21
Conclusion.....	23
Bilan.....	23
Analyse.....	25
Sources:	26
Annexes (Pièces jointes).....	28

Licence.....30

L'écriture chinoise, 3.000 ans d'analyses et rationalisations

Ère archaïque et impériale

L'époque archaïque et la multitude des variantes

Dans la Chine ancienne, les Jiaguwen, gravés sur des carapaces de tortues et omoplates de bovins avaient des allures très variables d'un support à l'autre, d'un site archéologique à un autre. Le rôle rituel, la faiblesse des échanges écrits, la barrière de la distance, mais aussi la main humaine permet l'apparition de très nombreuses variantes pour un même caractère en style Jiaguwen, Jinwen, ou Dazhuan. Il existe ainsi des catalogues entiers de variantes Jiaguwen des caractères 龍, 魚, 鳥, etc. R. Sears, spécialiste des sinogrammes anciens, publie par exemple 80 variantes Jiaguwen, 62 Jinwen et 25 Dazhuan du caractère 馬 ma3, cheval.¹ Dans une même 'phrases', les caractères varient dans leurs tailles, et les représentations d'un même caractère peuvent également varier. C'est un concept que l'on dessine, aussi, la graphie de chaque caractère n'est pas vraiment fixé.²

Les Zhou occidentaux auraient tenté une première codification vers -800, afin de mettre de l'ordre dans les usages régionaux et individuels. Une liste de 1000 caractères en 15 sections aurait été constituée et diffusée sur lamelles de bambou, avec plus ou moins d'effet.³ Suite à l'affaiblissement du pouvoir central Zhou, des principautés stabilisent peu à peu leurs graphies Dazhuan autour d'une graphie approximative lors du processus d'affirmation de l'autorité, mais la situation reste encore largement chaotique à l'échelle de la plaine chinoise.

Rappel des styles majeurs (pinyin et chinois) :

Jiǎgǔwén 甲骨文 Jīnwén 金文 Dàzhuàn 大篆 Xiǎozhuàn 小篆 Lìshū 隶書

Kǎishū (t) 楷書(繁体) Kǎishū (s) 楷书(简体) Xíngshū 行书 Cǎoshū 草书



1 Richard Sears, Chineseetymology.org :

<http://www.chineseetymology.org/CharacterASP/CharacterEtymology.aspx?characterInput=马>

2 Alleton, 2005. p.77

Dynastie Qin et Han : les premières grandes rationalisations

Réforme de Li Si

Avec l'établissement de l'empire Qin, c'est la victoire du légalisme et de la rigueur sur l'ordre ancien des Zhou et sur la multitude des principautés. L'empereur Qin, assisté de son chancelier Lǐ Sī (李斯) lance une vague de réformes et standardisations afin d'unifier l'empire, casser les particularismes, faciliter l'administration, les communications, le commerce et la circulation des troupes Qin. Cette réforme Qin inclue notamment la standardisation :

- des poids et mesures ;
- de la largeur des ;
- les sinogrammes ;

Le Shuowen Jiezi de Xu Shen l'explique ainsi :

[...] Lors que l'empereur Qin unifia le monde, le chancelier Li Si proposa la standardisation des écrits, et la cessation de la circulation des scripts qui n'étaient pas en accord avec le script [officiel] Qin.⁴

Li Si confia la création d'une liste de caractères officiels à un groupe d'érudits. Une liste d'environ 3.000 caractères est produite. Très largement inspirés de ceux déjà en vigueur dans le royaume Qin, cette liste interdit la multitude des variantes, simplifie les tracés en faveur de courbes régulières et de droites, régularise le cadre d'écriture des caractères (un rectangle vertical)⁵ et crée aussi de nouveaux caractères.⁶ De la multitude des Dazhuan, Li Si fait émerger 3.000 graphies ensuite approuvées par un édit impérial. Ce nouveau standard est connu sous le nom de "Xiaozhuan".

Ces Xiaozhuan se diffusent à travers l'empire Qin et permettent une communication effective et sûre entre contrées lointaines. C'est un élément décisif de la construction étatiques et d'une identité Chinoise/Han. Cette standardisation accélérant les processus de renseignement, d'administration, de fiscalité, de gestion du nouvel espace militaire : derrière ces Xiaozhuan, c'est aussi les pas des armées

3 Alleton, 2005, p.76

4 Gong & Qi, 2004, p.17

5 Alleton, 2005, p.79

6 Gong & Qi, 2004, p.17

et de l'État Qin qui s'entendent.

Cette standardisation pourrait être à l'origine de la « Proscription des Cent Écoles » et de l'autodafé de -213. La destruction des ouvrages anciens ayant pu être considéré comme la seule véritable solution pour l'anéantissement des scripts régionaux, des variantes désormais proscrites, et d'affirmation du nouvel ordre.⁷

Styles Lishu et Kaishu : vers une écriture de traits

En dehors de l'harmonisation par standardisation, c'est à dire par réduction de la multitude des variantes à une seule, une autre rationalisation s'opère. Le style Lishu, qui remplaçant les courbes Xiaozhuan par des traits, est autorisé dès la dynastie Qin,⁸ le Xiaozhuan s'utilisant dans les pratiques officielles ou rituelles, tandis que le Lishu est utilisé pour les documents d'usage quotidien.⁹

Vers le II^e siècle de notre ère (dynastie Han), c'est le style Kaishu qui devient la norme, plus rectiligne, plus franc, et faisant clairement apparaître les changements de direction. Ce style est encore aujourd'hui le plus utilisé, du fait sa double clarté -supérieure au Lishu- et esthétique -supérieure au Songti-.¹⁰

La révolution suivante est le développement de la xylographie/imprimerie, permettant des diffusions massives et relativement 'centralisées', ne dépendant plus de copistes isolés. Le style Kaishu, trop complexe à graver dans le bois, est adapté pour le besoin en style Songti/Mingti, plus rapide à graver, plus anguleux, et renforcé dans ses angles. Ce style Songti est l'apogée de la rationalisation des styles graphique. Les traits et angles droits sont la règle, tandis que les courbes sont presque absentes. Des empereurs soutiennent d'ailleurs l'utilisation double du Songti et du Kaishu par la proto-industrie de l'édition. C'est également ce mécanisme d'impression qui aurait permis la stabilité graphique de ces styles à travers l'ère impériale.¹¹

7 Alleton, 2005. p.79

8 Blakney, 2007, p6.

9 Alleton, 2005, p80.

10 Alleton, 2005. p84-85.

11 Blakney, 2007. p6.

Lexicographies impériales : une approche graphique et combinatoire

Xu Shen et le Shuowen Jiezi

L'autodafé de -213 peut être considérée comme un succès, puisqu'à la fin de la dynastie Han, les caractères Dazhuan et Xiaozhuan n'était plus familiers qu'à quelques spécialistes.¹² C'est pour cette raison que le lexicographe Xǔ Shèn (許慎, 58?-147?) compile vers +100 le *Shuōwén Jiězì* (說文解字) : un dictionnaire étymologique contenant 9.353 caractères en Xiaozhuan.^{13,14}

Jusqu'alors, les systèmes de classification des sinogrammes étaient soit sémantiques, par champs lexical, soit selon la division en « six écritures » : 1. images, 2. symboles, 3. composés sémantiques, 4. composés par éléments sémantique et phonétique, 5. composé par transformation graphique d'un caractère, 6. emprunts-rébus.¹⁵

Xǔ Shèn invente lui la classification graphique par clef (*bùshǒu* 部首), établissant 540 'radicaux' avec lesquels il classe l'ensemble des sinogrammes. C'est une révolution, puisque cette méthode permet la recherche d'un caractère par sa graphie, afin de trouver sa signification. Mais aussi par le fait qu'il affirme l'aspect combinatoire des sinogrammes.

Dictionnaire Kangxi

D'autres dictionnaires, de plus en plus complets, sont publiés. En 1008, le Guangyun affiche 26,194 caractères, tandis qu'une multitude de classements divers sont utilisés.

Une nouvelle révolution vient avec le dictionnaire Kangxi (1716: 康熙字典 *Kāngxī zìdiǎn*), contenant 40.545 caractères, et présentant un nouveau système d'organisation par 214 radicaux associés à leurs réductions et variantes. (voir annexes) Ce dictionnaire produit sous la direction de l'empereur Kangxi et présenté en pleine période de prospérité fait dès lors



Scan du dictionnaire Kangxi, avec sa police d'écriture.

autorité. Œuvre produite par l'autorité impériale, les pièces d'impression et la police d'écriture associée -produite spécialement pour l'œuvre- réaffirme aussi la graphie officielle des 40.545 caractères

12 Alleton, 2005. p81.

13 Britannica, 2009.

14 Alleton, 2005, p81.

15 Alleton, 2005. p44.

exposés en servant de modèle. Cette œuvre sera ensuite largement réutilisée lors des affirmations étatiques nationaliste, communiste, mais aussi japonaise et coréenne, ces gouvernements souhaitant également, périodiquement, réaffirmer une graphie officielle pour les caractères.

Ères moderne

Premières propositions de phonétisations

Phonétisation chinoise par associations

Le premier système employé semble être le Fangqie, sur le système de « X se lit comme A(initiale)-B(finale) », A et B étant des sinogrammes plus communs. Mais le choix des briques phonétiques n'est pas arrêté, et cela semble à chaque fois une improvisation.¹⁶

Premiers contacts avec systèmes phonétiques

Le premier contact franc avec un système d'écriture fondé sur la phonétique (et non la sémantique) à lieu au début de l'ère chrétienne, avec l'utilisation des premiers textes sanscrits. Ceci encourage les premières analyses syllabiques du chinois, vers le VIIème siècle. En 1260, durant la période mongole, afin de faciliter le contrôle sur l'espace chinois, Kubilai Khan demande à Pa'sse-pa une extension de l'alphabet mongole afin de permettre la transcription des sons chinois. Mais les fonctionnaires font un usage du mongole seul, excluant les sons chinois, tandis que le commun rejette cet alphabet phonétique et continue d'écrire par sinogrammes. Au XVIIe, les élites chinoises doivent apprendre le mandchou, une écriture alphabétique inspirée du mongole, mais n'appliquent pas la méthode aux caractères chinois.¹⁷

Propositions modernes

Matteo Ricci (1605), Nicolas Trigault (1626) seraient les premiers auteurs de systèmes alphabétiques de transcription des sinogrammes. Plusieurs dizaines de transcriptions alphabétiques, en fonction des besoins de la nation européenne, apparaissent ainsi.¹⁸

1648 voit la première tentative chinoise, mais c'est avec le chaos politique du XIXe, la montée du courant réformiste-nationaliste, ayant à la fois le souhait de renforcer la Chine, et de renforcer son

16 Alleton, p118

17 Alleton, p116-117

18 Alleton, p119-120

unité, qu'apparaît un mouvement favorable à la création de systèmes phonétiques pour décrire la prononciation correcte des caractères. En 1892, Lu Ganzhang (盧贛章) affirme la nécessité politique d'un alphabet.¹⁹ Aussi, de 1892 à 1918, c'est plus de 50 systèmes qui seront développés. Chen en cite 46 notables, classés en 5 catégories : 1. alphabétiques (8) ; 2. basés sur les traits chinois (17) ; 3. sténographique (10); 4. numériques (3); 5. autres (8).²⁰

Premiers usages à grande échelle

Avec la « Conférence sur L'Unification de la Prononciation » se tenant à Beijing en 1913, les réformateurs analysent et discutent longuement d'un dialecte standard, de la phonétique précise de ce dialecte, des systèmes de transcription existants, et proposent enfin le *Zhùyīn Zìmǔ* (注音字母), composée de 40 caractères adoptés en 1918 comme transcription phonétique officielle. (voir annexes)

En 1922, cinq tons y sont ajoutés tons : (neutre) ˊ ˋ ˊˋ ˋˊ ˙ 21

L'emploi du Zhuyin est celui d'une « annotation à droite » (voir image), précisant la phonétique officielle de caractères nouveaux ou difficiles. Le but est double :



1. diminuer l'analphabétisme,
2. soutenir la langue nationale (國語) que l'État républicain tente d'imposer pour renforcer l'unité de la Chine.

Mais la situation reste complexe, puisque d'autres transcriptions persistent, tels le Gwoyeu Romatzyh (1926) et le Latinxua (1929), en Chine,²² ou le Wade-Giles, à l'étranger.

19 Chen, p62

20 Chen, pp64-77

21 Chen, pp78-82

22 Alleton, p121-122

Réformes post-1945

La question phonétique en RPC

Enjeux

Avec les agitations des années 30, l'activité et les concurrences politiques discréditent les réformes récentes et encouragent la production de nouveaux modèles. Dès l'arrivée au pouvoir, en 1949, les communistes affirment la nécessité d'une écriture totalement alphabétique afin de lutter contre l'analphabétisme. Les caractères chinois sont perçus comme inutilement complexes et une relique du système impérialiste limitant l'accès à l'éducation.²³

Phonétisation de 1956 et 1958 : vers le Hanyu Pinyin

Quelques principes de bases sont d'abord désignés:²⁴

1. prononciation: selon le mandarin tel que prononcé à Beijing,
2. structure: syllabes sans limites dans le nombre de lettres ;
3. tons: 5 tons
4. séparateur : pas de séparateurs.

Une première proposition est diffusée à échelle nationale en 1956. Suivant les commentaires, les symboles sont modifiés, supprimant les caractères spéciaux afin de faciliter l'utilisation des machines à écrire étrangères.

Proposition (1956)	ž ċ š	z/ㄗ ç/q š/x	...ŋ j...	(rien)
Hanyu Pinyin (1958)	zh ch sh	ji qi xi	...ng y...	' (séparateur)

L'objectif est toujours d'unifier massivement la prononciation sur une langue commune, le Putonghua (普通话), mais également d'habituer au pinyin, tandis qu'une première simplification d'urgence des

²³ Alleton, p122.

²⁴ Hsia, pp122-126

sinogrammes est prévue. L'objectif long terme est réaffirmé comme étant l'utilisation de ce système Pinyin 100% phonétique.²⁵

La question de la simplification

Enjeux

La suprématie des puissances européennes poussent les intellectuelles est-asiatiques à souligner l'importance de l'alphabétisation des masses, et par conséquent, de la nécessité de simplification des sinogrammes.

Bases historiques

La complexité des sinogrammes à toujours encouragé les hommes à les simplifier. Chen cite quelques exemples anciens :

篆 → 𠄎+豕 (Zhou dynasty) ; 無 → 无 (Han dynasty) ;

與 → 与 (Sui dynasty) ; 淚 → 泪 (Ming dynasty)²⁶

Les utilisateurs de ces simplifications utilisant des formes ultra-cursives, un/des composants, ou des concepts liées au sens pour créer un caractères de même prononciation et sens, mais d'écriture plus simple. Les États ont cependant toujours veillés à réduire la prolifération de ces variantes, réaffirmant périodiquement les graphies correctes et excluant les innovations individuelles et régionales.²⁷

Réforme japonaise

Les japonais, défaits par les États-Unis en 1945, annoncent dès novembre 1946 un nouveau jeu de 1850 caractères officiels (1850 Tōyō kanji 当用漢字), dont une partie est simplifiée, ce sont les Shinjitai (新字體).²⁸

25 Alleton, p122.

26 Chen, p47.

27 Alleton, p85

28 Alleton, p113

Réforme chinoise

En RPC, la « Commission de Réforme de l'Écriture » (文字委员会) est favorable à une « pinyinisation » de l'écriture mais conscient de la nécessité d'une période de cohabitation pinyin-sinogrammes et de l'urgence de l'alphabétisation des masses. En plus de l'enseignement du Pinyin, elle propose donc aussi une simplification des caractères chinois. En 1954, la commission analyse les variantes collectées puis sélectionne les plus commodes et les plus populaires, en excluant les autres variantes. Le 28 janvier 1958, une liste de 515 mots esimplifiés, accompagnée de 54 radicaux dont les graphies sera systématiquement simplifié, et promulguée.²⁹ Pour ces 515 sinogrammes, le nombre de trait moyen est abaissé de 16 traits à 8,16 traits, permettant une écriture notablement plus rapide.³⁰

La simplification se fait selon 6 axes³¹ :

1. 雲 → 云 : emprunt des formes anciennes,
2. 言 → 讠 : emprunt des formes cursives,
3. 體 → 体 : emplois populaires,
4. 燈 → 灯 : simplification du phonétique,
5. 幾 → 几 : création/remplacement complet
6. 飛 → 飞 : suppression partielle.

Les tentatives postérieures de simplifications échoues. Le contexte politique des années fin-1970/début80 -plus libéral- ne permet pas une mobilisation en faveur des 853 sinogrammes simplifiés de la seconde vague (1977), et cette seconde réforme est finalement abandonnée.³²

29 Alleton, pp85-89

30 Alleton, p88

31 Alleton, pp87-88

32 Alleton, p88

La question de l'ordre des traits

Enjeux

La description des sinogrammes à aussi conduit à la description de leur exécution. Les traits des caractères chinois s'« empilent » de haut en bas, et de gauche à droite, selon un autre globalement dicté par la graphie (voir image). L'écriture est également soumise à des priorités : horizontale prioritaire verticales, etc. L'ordre d'exécution, s'il a peut de conséquence sur les style Kaishu et Lishu, à ensuite une influence croissante sur les styles Xinshu et Caoshu. Aussi, les chinois ont tenté de définir certaines règles pour l'exécution des sinogrammes.

Le but est double : faciliter l'apprentissage de l'écriture aux enfants (par l'apport d'un ordre d'exécution clairement défini), et uniformiser les écritures manuscrites (Xinshu).

Enseigner un tracé strict permet en effet d'obtenir des caractères manuscrits (Xinshu) d'allures similaires pour tout individu utilisant le même ordre d'exécution.



Image illustrant un détail standardisé par l'équipe de standardisation, et l'association quand à la graphie et l'ordre d'écriture.

Historique des publications

Dès 1615, le dictionnaire écrit par Mei Yingzuo (梅膺祚) contient dans son premier chapitre des exemples décrivant l'ordre de certains caractères, et dont le lecteur déduit des règles générales d'écriture. Sous la dynastie Qing, le 『父師善誘法』 (*Fushi shanyou fa*) de Tang Biao (唐彪) contient 50 exemples. La période républicaine voit également la production de plusieurs ouvrages sur le sujet, mais cela reste la vue des auteurs seuls.³³

En 1958, au Japon, le Ministère japonais de l'Éducation publie le 『筆順指導の手びき』 (*Hitsujun shidō no tebiki*), qui reste le standard japonais d'exécution des Kanji de 1958 à 1977, année où le Japon

³³ Li Xian > online version > 標準字體筆順手冊 > Introduction.

laisse aux établissements éducatifs le libre choix de leurs supports pédagogiques.

En 1995, avec le 『常用國字標準字體筆順手冊』 (*Changyong Guozhi Biaozhun Ziti Shouce*) de Li Xian (李鑾), présenté à Taipei, l'ordre des traits est effectivement standardisé par l'État nationaliste pour les caractères traditionnels.³⁴

En 1997, c'est au tour de la RPC de publier le 『現代漢語通用字筆順規範』 (*Xiandai Hanyu Tongyong Zi Bishun Guifan*), exposant l'ordre officiel, en RPC, de 7.000 caractères.³⁵

Ainsi, les 3 principaux pays utilisant les sinogrammes se sont dotés de descriptions standardisées de l'exécution des sinogrammes, dans le double but d'aider l'apprentissage des sinogrammes, et d'unifier les écritures manuscrites.

34 Li Xian > online version > 標準字體筆順手冊 > Introduction.

35 現代漢語通用字筆順規範

La question graphique

Enjeux: Développement de l'édition et besoin de standardisation graphique

Les années 50 ont largement entamée l'illettrisme des population chinoise. L'éducation progresse, les supports écrits se diffusent, la presse se développe. Les années 70 est 80, avec la révolution information informatique, le décollage économique, et l'élévation du niveau de vie accentue la production-consommation d'écrits. La problématique est a nouveau celle de Kangxi, ou de Li Si : afin de faciliter la collaboration des forces humaines, administratives, économiques du territoire (PRC ou ROC), et ici face à un nouvel outil de diffusion, il devient nécessaire de diffuser des conventions d'écriture, graphiques et numériques, pour les sinogrammes.

Mais les polices d'écritures utilisées par l'industrie de l'édition (on se sert des polices d'écritures pour construire ou graver les tambours d'imprimantes), construites à partir de sources variées, commencent à poser problème. Deux polices d'écritures peuvent en effet produire un même caractère chinois, dans le même style Songti, mais avec de notable variation de traits.

Aussi, dans la continuité du travail fait sous Kangxi, des initiatives de stabilisation de la forme des sinogrammes sont à nouveau relancées, avec comme cibles majeures l'édition et ses polices d'écritures, dont l'influence et l'ubiquité est acquise. L'édition et l'administration passant à l'ère numérique, la standardisation graphique s'associe aux entreprises numériques. L'exemple de Taïwan est le plus documenté.³⁶

Le cas taïwanais : l'établissement de graphies officielles

A Taïwan, en 1973,³⁷ la National Taiwan Normal University (NTNU、國立師範大學) se voit confiée le travail d'analyse et d'affirmation d'une graphie standard attachée à chaque caractère. De nombreuses sources lexicologiques traditionnelles (Shuowen Jiezi, Kangxi dictionary), calligraphiques (中國書法大

36 Le Ministère taïwanais de l'Éducation mettant ses ouvrages de référence en accès libre et gratuit sur internet, il est plus facile de se renseigner sur le cas taïwanais. — Pour ce qui est du cas continental de 'standardisation' et 'numérisation' des caractères, des ouvrages et process similaires ont eu lieu (cf plus bas), mais leurs documents ne sont pas accessibles en ligne.

37 國六十二年二月一日

字典), statistiques (常用字彙初稿), sont utilisées.³⁸ Le 常用國字標準字體表 ("Graphie standardisée des caractères fréquents"), fixant la graphie de 4.808 caractères traditionnels fréquents est ainsi publié en 1979 par le Ministère de l'Éducation de Taïwan.

Le travail consistait en une fine opération de tries, d'analyses, et de choix graphiques, l'œuvre présente enfin la graphie officielle affirmée de chaque caractères listé. Une sélection de 517 caractères représentatifs ou difficiles est particulièrement renseignés avec : 1. la graphie officielle affirmée ; 2. le radical ; 3. le nombre de traits hors-radical et total ; 4. la graphie Xiaozhuan ;

5. la phonétique locale officielle ; 6. une explication textuelle afin de comprendre et d'éviter les erreurs de graphie ; 7. l'ordre des traits. (voir image) Pour les autres caractères, moins détaillés, leurs explications complètes peuvent être obtenues par cascade, en se basant sur les 517 exemples.

Le jeu de cette documentation officielle est ainsi constitué et complété :³⁹

- "常用國字標準字體表" (1978) contenant 4.808 caractères, avec une révision (1979: 訂正本), avec 137 modifications.
- "次常用國字標準字體表" (1981 稿本, 7894 字+2845 possibles; 1982 正本: 6332 字+4399 rares possibles)

Le total s'élève ainsi à 4808+6332=11.140 caractères standardisés, expliqués, fixés.

Le jeu est encore étendue à destination des compagnies informatiques :

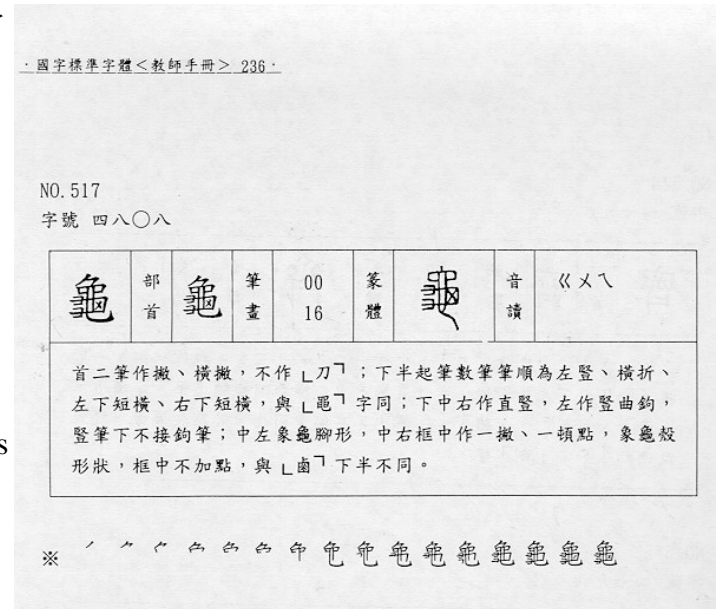
- "罕用國字標準字體表" (1983), avec 18.388 caractères rares.
- "異體字表" (1984), avec 18.588 caractères rares.

L'ordre des traits est également standardisé :

- "常用國字標準字體筆順手冊" (1995)

38 Ceng Rongfen > online version > 標準字體的研訂原則與實例 > 常用字的選擇

39 Ceng Rongfen > online version > 標準字體的研訂簡史



L'exemple de 龜, extrait du livre de Ceng Rongfen, 1979.

L'initiative est d'abord lancée par l'État, mais dès 1982, l'équipe en charge de l'établissement de standards pour la graphie des caractères, au service du ministère de l'éducation, entre en collaboration avec plusieurs acteurs de l'industrie informatique, et la création d'une première police d'écriture Songti de 4.808 caractères standardisés est créée (1982), le jeu complet de 11.151 caractères standardisés est annoncé en 1993.⁴⁰

Le cas chinois : l'établissement de graphies officielles, les documents de références

Le même mouvement de quantification, clarification, standardisation a eu lieu du côté chinois. Dès 1965, le 印刷通用漢字字形表 (1965: 6196 字) définit la graphie correcte et officielle à employer suite aux réformes communistes. Ce document est remplacé en 1988 par le 現代漢語常用字表 (3500 字), puis sa mise à jour, le 現代漢語通用字表 (7000 字 : 3500+3500). La définition de ces graphies servira de base aux polices d'écritures en encodage GB.

La question des Encodages

Premières tentatives et tentatives nationales

Les premiers ordinateurs sur le sol asiatique ne fonctionnent qu'avec les caractères occidentaux, puis, une multitude d'encodages spécifique à une compagnie (Adobe, Sony) font leur apparition. Un fichier ne peut alors être transmis d'un groupe à l'autre.

Les premières tentatives de quantification et numérisation des sinogrammes à s'être largement répandues sont respectivement la série des encodages GB, et la série Big5.

La série des **encodages GB** (国标 Gúobiāo : Standard national) est lancée en 1980 par la Chine populaire, afin de standardiser l'encodage des caractères simplifiés, et permettre l'échange de fichiers.

Pareillement, la **série Big5** est lancée par la République de Chine (Taïwan) en 1984, afin de standardiser l'encodage des sinogrammes traditionnels.

La standardisation de ces encodages, c'est à dire la diffusion sur toutes les machines d'un encodage unique déterminé : GB pour la Chine, Big5 pour Taïwan, Hongkong et Macao, permet enfin à des

40 Ceng Rongfen > online version > 標準字體與中文電腦的發展關係 : point 2 et 5.

fichiers de pouvoir être transmis d'une machine à une autre, en restant compréhensible. Des tables de conversions permettent également la conversion de Big5 à GB, et inversement, tandis que de larges tables additives sont produites par certains groupes spécifiques : le gouvernement d'HongKong (~3.000 字), celui de Singapour, des ministères aux besoins pointus. La situation a resté un temps là du fait de la rivalité politique entre ROC et PRC.

S'ajoute à ces encodages des "polices d'écritures", qui sont la couche visuelle et 'human friendly', permettant d'afficher des sinogrammes lisibles par l'homme.

Unicode

L'étape suivante dans la définition des sinogrammes est d'abord franchie par le dépassement de ces encodages locaux (GB et Big5), par l'encodage "Unicode". L'Unicode est une tentative internationale d'encodage universel de tous les caractères écrits humains. Sous impulsion américaine, le groupe UniHan (Groupe pour l'Unification des caractères Han) est créé et intégré au projet Unicode. Une nouvelle table est construite pour accueillir les sinogrammes. Japonais, taïwanais, chinois, coréens, s'accordent sur une organisation selon les radicaux de Kangxi, puis selon le nombre de traits, avec inclusion de tous les caractères traditionnels et simplifiés. En partant ainsi des caractères eux-mêmes, les tensions nationalistes liées aux divergences d'ordre alphabétique selon une phonétique (Pinyin ? Zhuyin ? Katakana ? Hangeul ?) s'effacent. On obtient pour la première fois un encodage international.

L'Unicode, dans sa volonté d'unification graphique et informatique, s'emploie à collecter les analyses, et à les fusionner. Les recherches taiwanaises, chinoises, hongkongaises, coréennes précédemment citées et leurs groupes de recherches sont intégrés au projet, et lui servent de base graphique.

Clairement tournée vers le support numérique, le groupe de recherche UniHan confirme plusieurs innovations.

La liste des 36 traits CJK(V),⁴¹ ainsi que la liste des 214 radicaux traditionnels⁴² et des 115 composants graphiques complémentaires⁴³ sont des tentatives de création d'un jeu complet de 'briques', permettant

41 *CJK Strokes*, The Unicode Standard, V.5.1, p238. <http://unicode.org/charts/PDF/U31C0.pdf>

42 *Kangxi Radicals*, The Unicode Standard, V.5.1, p277-281 : <http://www.unicode.org/charts/PDF/U2F00.pdf>

43 *CJK Radicals Supplement*, The Unicode Standard, V.5.1, p273-276 : <http://www.unicode.org/charts/PDF/U2E80.pdf>

de recréer l'ensemble des 100.000 sinogrammes historiques par inclusion successives, positionnements, et transformations géométriques (modification de la largeur et/ou hauteur).

Les nouvelles descriptions géométriques et informatiques

Le front de bataille dans la complétion des méta-données liées aux sinogrammes et à l'informatique est aujourd'hui du côté des 'Chinese characters Description Languages', basés sur les langages de programmations tels que le XML, qui permettent la description de graphiques par courbes mathématiques et coordonnées, et l'inclusion en cascade.

L'image ci-à droite expose le système : le caractère 敖 ào, est constitué des traits Heng-Heng-Shu, du caractère 万 wan et du composant 攴 pu (réduction de 攴). Les points de contrôles permettent d'appliquer des transformations.⁴⁴

Figure 2-4. 敖 CDL for BMP graph 敖 [U+6556], 1



敖 ào: H-H-S-万-攴, plus des transformations.

Du côté code, dans le cas du système dit CDL,⁴⁵ un trait Dian est dessiné en XML par des points associés à des courbes de bezier tel que :

```
<path
  id="heng"
  style="fill:#black;fill-opacity:1"
  d="M 328.4,4.9 C 330.3,433.0 332.3,440.8 335.7,444.4 C
    339.5,448.5 344.7,4.0 345.7,440.0 C 347.9,430.6 341.4
    2396,422.70 335.4,415.8 C 329.3,408.9 327.3,415.6 328.4,421.9 z"
  sodipodi:nodetypes="csss" />
```

⁴⁴ Richard Cook, Chinese Character Description Languages, 2003. p.110. <http://www.wenlin.com/cdl/03387-cook-cdl.pdf>

⁴⁵ T. Bishop & R. Cook, 2003.

Chaque radical et composant est une somme de traits. Les caractères complexes sont des sommes de traits, radicaux, composants complémentaires, ou de caractères plus simple, tel que :

```
<cdl char="太 ">
  <comp char="大" points="0,0 128,128" />
  <stroke type="dian" points="45,104 66,128" />
</cdl>
```

Ainsi, un sinogrammes est définit graphiquement, mais aussi par ses composants constituants. Cette description XML permet de stocker des méta données comme :

- Les éléments constitutifs : 太 = 大 , 丶
- Les traits et ordre des traits : 太 = Heng-Pie-Na-Dian
- La direction des traits : la direction d'un Heng étant connue : de gauche a droite.
- La position relative des composants : le 大 est au dessus du Dian ; le Pie croise le Heng.

Du coté des utilisateurs, ces données peuvent être utilisées dans la reconnaissance des sinogrammes sur écran tactiles, comme c'est le cas dans l'IPhone.

Du coté des compagnies informatiques et de la standardisation des graphies, ce système permet la génération de polices d'écritures rationnelles, puisque toute la police sera basée sur un lot de ~40 traits. La mise à jour des 20.000 sinogrammes de cette police se faisant par la correction du seul élément à modifier. Les projets CDL, SCML, et HanGlyph sont fondés sur de telles approches graphiques et géométriques. Le projet CDL décrit 73.000 caractères selon de telles méthodes.⁴⁶

46 <http://www.wenlin.com/cdl/>

Conclusion

Bilan

La période de l'invention des caractères chinois, puis celles des premières royautés sont des périodes de multiplication des variantes graphiques des caractères. Les autorités ont bien tenté des uniformisations, comme la liste des Zhou, vers -800, mais la faiblesse du pouvoir central, et la division de la plaine chinoise ne permet pas d'uniformiser les pratiques.

De l'ère Qin-Han, qui voit plusieurs pas majeurs dans la rationalisation des caractères, l'on peut retenir 3 faits :

- **La codification Qin** : Lǐ Sī (李斯) institue « une seule graphie pour chaque caractère, et un style homogène » de courbes régulières, et de droites, pour un lot d'au moins 3.000 caractères,
- **L'apparition du trait** : on passe de la prédominance des courbes sinueuses, à la domination du trait. D'abord par le style Lìshū (隸書), puis plus encore par le style Kǎishū (楷書),
- **Le système de radicaux** : Xǔ Shèn (許慎) rationalise le listage des caractères, par une approche graphique.

Le reste de l'ère impériale expose d'autres étapes majeures :

- **L'imprimerie** : le développement de l'imprimerie permet une production 'centralisée', rationnelle, ne dépendant plus de copistes. Des styles (Kaishu, Songti) sont soutenus par l'administration impériale.
- **La police Songti** : développé par l'imprimerie, cette police est l'apogée de la rationalisation des styles graphiques. Les traits et angles droits, les courbes simples, sont ici la règle.
- **Le système de radicaux** : il est amélioré, et depuis Kangxi basé sur 214 clefs, accompagnées de variantes ou réduction.

Il est également à noter les premiers systèmes phonétiques : Fanqie, Ricci, Wade-Giles.

L'ère contemporaine, des grandes réformes pour l'alphabétisation de la population chinoise, peut être

daté des années 1890. Ici, il y a volonté de standardiser la phonétique associée à chaque caractère chinois, mais aussi la volonté de simplifier la complexité des caractères. Au vu de la multitude des tentatives, il convient de se consacrer aux tentatives réussies :

- **Le système Zhuyin Fuhao (1913)** : un alphabet de d'initiales, médianes, basé sur les sinogrammes, associé à 5 tons ;
- **Le système Hanyu Pinyin (1958)** : utilisant alphabet latin, associé à 5 tons.
- **La simplification communiste (1956)** : avec le remplacement par des variantes ou nouvelles graphies plus simples.

Ces systèmes phonétiques ont permis l'émergence de dictionnaires où les caractères sont classés par phonétiques, mais ont aussi facilité le passage à l'ère informatique. La simplification de 1956 de fait facilite l'apprentissage des sinogrammes, et l'alphabétisation des masses. Le projet de romanisation complète, initialement prévu dans de nombreuses zones, n'a abouti que dans le cas particulier de la Corée.

Avec les affirmations étatiques du 20^{ème} siècle, d'autres standardisations ont également eu lieu, je pense en particulier au travail de :

- **La redéfinition et standardisation des graphies correctes** : effectués en Chine, à Taïwan, au Japon, d'abord à destination des éducations nationales respectives, mais aussi à destination de l'industrie de l'édition et de l'informatique, afin de conserver l'universalité de graphies dans le monde numérique ;
- **La définition d'un 'ordre des traits standard'** : principalement afin que les écritures manuscrites Xingshu d'un individu 'A' restent lisible par un individu 'B'.

Enfin, il est à signaler le travail d'unification informatique, mais aussi de collecte des méta-données, avec :

- **le renfort de l'interopérabilité** : la création d'un encodage international facilitant l'échange de documents numériques ;
- **La collecte des méta-données associées** : par la création de bases de données associant à chaque caractères toutes les données précédemment créés (composants, phonétiques locales,

traits, définitions chinoise, étymologie, définition anglaise, etc.)

- **La numérisation des sinogrammes** : la définition géo-mathématique des sinogrammes.

Analyse

Aussi, que ce soit pour les cas anciens (Li Si, Xushen) ou pour les cas récents chinois ou taïwanais (encodages, standards), c'est en fait une sélection, une clarification, une stabilisation, et une numérisation de tout une masse de méta-données qui tournaient depuis longtemps autour des sinogrammes. Les variations et variantes anciennes, graphiques, phonétiques, sémantiques, d'ordre d'écriture, sont examinées, triées, dans le but d'être validés ou effacés. Cette première longue étape de nettoyage de l'héritage lié aux sinogrammes prend ensuite la direction et forme de la diffusion d'une norme affirmée commune. Dans la masse des variantes et pratiques, ces standardisations sont largement des sélections artificielles soudaines, coupant les branches de l'arbre pour n'en garder qu'un tronc droit. Mais d'un point de vue pragmatique, cette diffusion de conventions facilite la communication, les échanges, les contrôles, et est donc un mouvement vers l'intégration et l'unification de l'ensemble chinois, mais aussi vers l'intégration à l'espace monde et à ses échanges.

Sources:

Histoire des sinogrammes et réformes

- ALLETON V., L'écriture chinoise, 1970, PUF, (coll. Que sais-je ?, n°1374, 6ème édition 2002), ISBN 2 13 052921 6.
- XUE Wu , 1994, *Writing reform at a crossroads in contemporary China*, Simon Fraser University.
- GONG Qi, QI Huang, Jerry NORMAN, Helen WANG, 2004, *Chinese characters then and now*, Volume 1 of Ginkgo series, Edition Voldemeer. Springer. 268 pages. ISBN: 9783211227954.
Site: <http://books.google.com/books?id=YdTAugdUQDQC&lpg=PA7&ots=mINvvVH2wt&lr=&pg=PA17>
- **Chinese languages, Qin dynasty standardisation.** (2009). In *Encyclopædia Britannica online*:
Site: <http://www.britannica.com/EBchecked/topic/112557/Chinese-languages>

Phonétique et standardisation :

- CHEN, J.T.S., Les réformes de l'écriture chinoise, Mémoire de l'institut de hautes études chinoises, vol.12, collège de France, PUF, 240 pages, 1966. ISBN: 2-85757-017-1
- HSIA, Tao-Tai, China's language reforms, Far Eastern Publications, New haven, Connecticut, Minor serie A, n°21, 1956.

Standardisation des graphies et de l'ordre des traits:

- CENG Rongfen (曾榮汾) & al., 國字標準字體(教師手冊), Taiwan Minister of the Education. ISBN: 957-00-4240-0. Site: http://www.edu.tw/files/site_content/M0001/std/c4.htm?open
1979 (甲) : 常用國字標準字體表,
1982 (乙) : 次常用國字標準字體表
- LI Xian (李鑾) & al., 1995, 常用國字標準字體筆順手冊, Taiwan Minister of the Education. ISBN 957-00-7082-X Site : http://www.edu.tw/files/site_content/M0001/bishuen/c8.htm?open
- 現代漢語通用字筆順規範, 453pages, 1997, editeur: 语文出版社, ISBN:7801262018
<http://www.amazon.cn/detail/product.asp/168-5091939-8361003?prodid=zjbk002902>

Numérisations :

- ZHOU Jing, *Combat over Chinese character unification*, 2008, China.org.cn.
Site: <http://beijing2008.german.china.org.cn/english/news/242268.htm>
- BISHOP T. & COOK R., *A Specification for CDL : Character Description Language*,
Site: http://www.wenlin.com/cdl/cdl_spec_2003_10_31.pdf
- COOK, Richard : *The Extreme of Typographic Complexity: Character Set Issues Relating to Computerization of The Eastern Han Chinese Lexicon Shuowenjiezi*, STEDT Project, Linguistic Department, University of California, Berkeley, 2001.
Site: <http://linguistics.berkeley.edu/~rscook/pdf/IUC18-SWJZZ.wp3-600.pdf>

Liste des 40 caractères et 5 tons du Zhuyin Fuhao:

ㄅ	ㄆ	ㄇ	ㄏ	一	ㄨ	ㄌ	ㄩ	ㄛ	ㄝ
ㄉ	ㄊ	ㄋ	ㄌ	ㄝ	ㄨ	ㄨ	ㄨ	ㄨ	ㄨ
ㄍ	ㄎ	ㄎ		ㄨ	ㄨ	ㄨ	ㄨ	ㄨ	ㄨ
ㄐ	ㄑ	ㄒ							
ㄓ	ㄔ	ㄕ							
ㄖ	ㄗ	ㄘ		ㄨ	ㄨ	ㄨ			
ㄙ	+5tones:			ˊ	ˋ	ˊ	ˋ	ˊ	ˋ

List of 'Zhuyin Fuhao' characters (AR PL UKai TW)

Licence

Afin de favoriser la diffusion des savoirs, ce texte est réalisé sous **licence Creative Commons : CC-by-3.0**.

Si vous souhaitez réutiliser tout ou partie de ce texte, alors vous devez mentionner l'auteur original des passages empruntés ou réécrits et la licence :

Hugo Lopez, CC-by-3.0