

Considérations concernant l'utilisation de l'écriture latine dans les variantes de domaines de premier niveau internationalisés

Rapport final du groupe d'étude VIP (*Variant Issues Project* - Projet sur les problèmes liés aux variantes) de l'ICANN pour l'écriture latine

Résumé

Ce groupe d'étude a examiné tous les caractères de la version 6.1.0 de la table de code des caractères Unicode associés à l'écriture latine et valides en vertu du protocole IDNA2008. Il a identifié différents « risques de confusion » demandant un examen attentif dans le classement d'un sous-ensemble d'un répertoire plus large pour une utilisation locale. La résolution de ces questions dépend cependant principalement des conventions orthographiques locales. Celles-ci tendent à traiter les mêmes caractères de différentes manières. Des chaînes qui sont très similaires dans le contexte d'une langue peuvent ne pas avoir les mêmes connotations dans une autre langue.

Etant donné que l'alphabet latin est celui qui est utilisé par le plus grand nombre de communautés linguistiques, essayer de donner un aperçu global des besoins de chacune d'entre elles est une entreprise presque impossible. Essayer de faire une synthèse de ces besoins malgré tout serait faire preuve d'un manque de compréhension envers les communautés qui n'ont pas encore rejoint la discussion portant sur les IDN (Noms de domaine internationalisés).

Le groupe d'étude conclut donc qu'il n'existe aucune base pour le traitement catégorique d'un point de code quelconque assigné à un élément de l'alphabet latin comme étant équivalent à tout autre point de code. Il ne considère pas non plus que cette base existe au-delà de ce qui est déjà incorporé au protocole IDNA2008. Le processus de candidature aux TLD de l'ICANN ne devrait pas donner suite à des demandes de chaînes latines multiples en partant du principe qu'elles sont des variantes les unes des autres. Si un examen rigoureux s'avère nécessaire pour évaluer le risque de confusion des étiquettes TLD proposées, cela n'en fait pas pour autant des variantes au sens de l'étude VIP.

Deux tables de références sont jointes au présent rapport dans lequel sont énumérés et présentés les points de code valides associés à l'écriture latine. La table primaire énumère les éléments qui peuvent être considérés comme appartenant à des systèmes d'écriture courants et qui possèdent donc une utilité potentielle pour les répertoires locaux d'IDN. La table secondaire comprend des caractères qui ne possèdent pas cet attribut ou qui sont moins bien adaptés pour un usage sans restrictions. Bien que la première puisse être considérée comme l'univers de points de code disponibles pour les futures étiquettes TLD basées sur l'alphabet latin, le groupe d'étude ne recommande pas d'accepter de manière inconditionnelle un point de code sans avoir au préalable une garantie spécifique de son utilisation dans le contexte linguistique prévu. La table d'écriture latine de la zone racine peut par conséquent être un

classement descriptif des points de code qui apparaissent réellement dans cette zone plutôt qu'une tabulation normative de points de code pouvant y être inclus sans autre considération.

1. Introduction

Dans le cadre des efforts visant à augmenter la disponibilité de noms de domaine de premier niveau internationalisés (IDN), l'ICANN a lancé un programme limité destiné à introduire des domaines de premier niveau de code pays en écritures non latines¹ et s'apprête à lancer un programme qui permet la création de domaines de premier niveau générique dans n'importe quelle langue². Un des problèmes de ce processus est qu'il remet en question l'hypothèse selon laquelle un domaine unique peut être représenté par une étiquette unique dans le système de noms de domaine. Avec les IDN, et avec la norme Unicode sur laquelle ils sont construits, certaines étiquettes peuvent être présentées sous différentes formes. Cela arrive lorsqu'un caractère abstrait (voir Section 6.1) dans l'étiquette est trouvé dans plus d'un point de code de la table de code des caractères Unicode ou peut être représenté par des séquences alternatives de points de code.

Pour faciliter le développement de politiques d'assignation et d'allocation de ces multiples étiquettes dans le nouveau processus de domaine de premier niveau, l'ICANN a établi le « Projet de variantes de TLD IDN ». Son rôle est d'identifier et d'analyser les problèmes à aborder et de fournir des recommandations sur lesquelles les politiques devraient se baser. Dans sa première étape, cette étude vise à analyser en détail et à parvenir à une approche commune de ce que les parties prenantes considèrent comme étant des variantes et des aspects importants les concernant. Ces énoncés de problèmes seront réunis en un seul document contenant les problèmes pour lesquels des solutions doivent être trouvées et évaluées.

Pour obtenir un large échantillon de données, six équipes d'étude de cas ont été constituées — chacune pour un alphabet particulier utilisé dans les systèmes d'écriture d'un segment important de la population mondiale. Ces alphabets sont l'arabe, le chinois, le cyrillique, le devanagari, le grec et le latin. Le cas du latin, que nous analysons ici, diffère des autres à plusieurs égards. En premier lieu, parce que c'est l'alphabet sur lequel le système de noms de domaine (DNS) est basé et que son déploiement date d'une trentaine d'années ce qui a permis d'étayer de nombreuses observations et conclusions. L'inclusion de l'alphabet latin de base « A à Z » dans le jeu de caractères ASCII répond en grande mesure aux besoins de plusieurs communautés linguistiques. Les motifs qui poussent à utiliser l'écriture latine pour les IDN peuvent donc différer de ceux qui justifient l'utilisation des langues non-latines.

En deuxième lieu, en tant que langues ayant des liens historiques avec le cyrillique et le grec, il existe des points communs entre ces trois alphabets. Le déploiement pratique de plusieurs écritures implique l'évaluation préalable des risques de confusion et des problèmes de sécurité communs existant pour l'ensemble de ces écritures.

Comme base de travail, les équipes d'étude de cas ont reçu une liste de questions écrites³. Bien qu'il ne soit pas obligatoire d'y répondre, elles ont été conçues pour aborder certains des domaines clés que les équipes devaient examiner pour avoir une idée d'ensemble de l'énoncé du problème.

1 <http://www.icann.org/en/topics/idn/fast-track/>

2 <http://newgtlds.icann.org/>

3 <https://community.icann.org/download/attachments/16842778/Draft+Questions.pdf>

En substance, les équipes d'étude de cas ont été invitées à évaluer le concept de variantes du point de vue de l'expérience de l'utilisateur. Les utilisateurs du DNS — non seulement les utilisateurs d'ordinateurs mais aussi les utilisateurs de téléphones portables qui utilisent les domaines, les systèmes techniques de maintenance des administrateurs de système, les éditeurs de logiciels, les registres, les bureaux d'enregistrement, etc. — ont certaines attentes concernant la façon dont les noms de domaines devraient fonctionner. L'introduction de variantes devrait être fondamentalement destinée à combler les lacunes définies pour lesquelles ces attentes doivent être satisfaites afin d'assurer des opérations prévisibles et fiables. En abordant les attentes à satisfaire comme un énoncé de problème, des solutions potentielles peuvent être évaluées en fonction de leur capacité à combler ces lacunes, en relation avec leurs implications techniques et politiques.

2. Définitions

Pour être sûr que les groupes de travail utilisent tous le même vocabulaire, il leur a été recommandé de se mettre d'accord sur une terminologie⁴ spécifique lorsque cela était nécessaire. Les définitions données dans le RFC 6365⁵ doivent être considérées comme la base de la terminologie d'internationalisation pour améliorer les définitions fournies. Lorsque cette documentation ne propose pas une définition claire nécessaire pour cette étude de cas, des détails supplémentaires sont ajoutés, le cas échéant, dans le présent document.

3. L'écriture latine

Le répertoire de base des caractères pour le DNS est donné dans la norme ASCII. Bien que tout octet puisse être utilisé dans un nom de DNS, une règle distincte pour les noms d'hôte limite le répertoire aux lettres latines de « a » à « z » en majuscule et en minuscule, aux chiffres de « 0 » à « 9 » et aux traits d'union « - » (« ensemble LDH »). Bien qu'il existe des différences dans la manière dont le DNS traite les formes majuscules et minuscules des lettres, les utilisateurs les considèrent en général comme transparentes. Ces caractères ont été utilisés pour tous les enregistrements de noms dans le DNS jusqu'à l'apparition des IDN en 2003.

Le protocole utilisé pour la mise en œuvre des IDN est formellement connu comme « Noms de domaine internationalisés dans les applications » ou IDNA. Il utilise un « codage compatible avec ASCII » pour des caractères donnés du répertoire Unicode plus exhaustif. Le codage et le décodage et l'affichage de caractères Unicode dépendent du soutien existant au niveau de l'application. Les enregistrements correspondants dans le DNS proprement dit restent basés sur l'ASCII. La version la plus récente de ce protocole est dénommée IDNA2008⁶, (pour sa date de lancement; mais elle a été publiée en 2010). Cette version et la version initiale IDNA2003 se basent toutes deux exclusivement sur la norme Unicode⁷ pour dériver les points de code autorisés.

⁴ <https://community.icann.org/download/attachments/16842778/Draft+Definitions.pdf>

⁵ « Terminologie utilisée dans l'internationalisation dans l'IETF » (RFC 6365), <http://tools.ietf.org/html/rfc6365>

⁶ « Internationalised Domain Names in Applications (IDNA) : Définitions et cadre du document » (RFC 5890) <http://tools.ietf.org/html/rfc5890>

⁷ <http://unicode.org/standard/standard.html>

La norme Unicode assigne plusieurs propriétés de caractères à chaque point de code, dont un qui indique l'écriture. Cette propriété est utilisée pour subdiviser la table de caractères Unicode en blocs avec en en-tête l'élément désignant la langue, tel que cela est publié sur le site Internet d'Unicode⁸.

L'alphabet latin est présenté dans neuf de ces blocs qui sont indiqués dans l'Annexe B1. Le bloc « latin de base » reprend le répertoire ASCII et inclut l'ensemble LDH auquel les noms d'hôtes ont été historiquement limités. Les lettres latines au-delà de l'ensemble LDH, ainsi que les formes marquées diacritiquement et autrement décorées sont présentées dans les blocs latins étendus et supplémentaires avec d'autres lettres latines dans la rubrique « Symboles phonétiques ».

Le protocole IDNA2008 permet d'utiliser un grand nombre de points de code provenant de ces blocs (n'interdisant que celui qui est intitulé « Lettres latines plein chasse »). Il impose également certaines limitations catégoriques, par exemple, sur le point de code qui représente des signes de ponctuation ou des symboles. Les points de code autorisés sont qualifiés de « valide du point de vue du protocole » ou PVALID.

Le sous-groupe des points de code validés par le protocole assigné à  l'écriture latine que le groupe a considéré comme pertinent pour l'étude des variantes latines est présenté dans l'Annexe B2. Une deuxième liste de points de code validés par le protocole qui ne sont pas utilisés dans le système d'écriture quotidien mais qui pourraient néanmoins figurer dans des IDN dans certaines circonstances est présentée dans l'Annexe B3. Ces listes offrent une énumération exhaustive de ce que le groupe a considéré comme les points de code de l'écriture latine.

4. Attentes des utilisateurs

4.1. Sensibilité à la casse

Le fait de se baser sur l'écriture latine pour les noms de domaines non internationalisés a donné lieu à une attente de l'utilisateur clairement établie concernant le fonctionnement des noms de domaine en langue latine. En particulier, l'absence de distinction entre les majuscules et les minuscules dans les chaînes LDH permet leur représentation précise selon une méthodologie qui ne s'applique pas aux IDN. Cependant, étant donné que les utilisateurs écrivent souvent les étiquettes LDH en majuscules, en minuscules ou dans un mélange des deux, il est raisonnable de supposer qu'ils s'attendent à être en mesure de faire la même chose avec d'autres caractères latins.

4.2. Les navigateurs Web

Le logiciel de navigation du Web conserve généralement la casse d'un nom de domaine LDH comme il a été saisi, en laissant les effets de toute normalisation de casse visibles pour l'utilisateur. Cette possibilité ne s'applique pas nécessairement à l'environnement IDN et un certain nombre de caractéristiques d'autres navigateurs supportant les IDN ne possèdent pas d'équivalents dans l'environnement de travail traditionnel. Les difficultés que les utilisateurs peuvent rencontrer pour s'adapter à ce système sont aggravées par l'absence d'une approche uniforme à la mise en œuvre de l'IDNA par tous les développeurs de navigateurs. Certains navigateurs affichent la forme codée ASCII d'un IDN (l'« étiquette A ») plutôt que la forme localisée attendue (l'« étiquette U ») si le TLD ne figure pas sur une liste blanche dans un répertoire central maintenu par le développeur. D'autres

navigateurs peuvent exposer l'étiquette A sur la base de la logique dérivée de paramètres de langage du système opératif sous-jacent, ainsi que de ceux du navigateur.

4.3. Claviers

A-Z et 0-9 sont présents sur la plupart des claviers d'ordinateur. Les caractères en latins étendus qui comprennent des signes diacritiques ou qui sont représentés dans d'autres formes spéciales se trouvent sur les claviers dans les régions où ces caractères sont utilisés. Bien qu'il existe des différences mineures dans la disposition des claviers latins, telles que QWERTY ou AZERTY, l'ensemble LDH est disponible sur tous les claviers.

5. Pratiques pertinentes concernant les domaines non IDN

Les registres de domaine qui permettent l'enregistrement d'IDN rassemblent généralement des répertoires plus petits de points de code autorisés de l'ensemble valide selon le protocole. (De fait, l'IDNA2008 attend le classement de répertoires locaux plus petits). Ils sont en général documentés sous forme de tableaux, regroupés par écriture ou par langue, selon ce qui correspond le mieux à la communauté ciblée. Beaucoup sont publiés dans un répertoire central maintenu par l'ICANN sur le site de l'IANA.⁹

Dans la pratique préalable, il n'y avait aucun équivalent à cette sélectivité. L'ensemble LDH offre vingt-six lettres qui peuvent être complétées par dix chiffres et un trait d'union. Les lettres peuvent être utilisées sans distinction entre majuscules et minuscules. Le protocole DNS traite les deux formes comme étant équivalentes dans la mesure où « oldnews » et « OldNews » correspondent à la même étiquette. Cela permet à des étiquettes visiblement similaires comme « OldNews » et « 01dNews » de cohabiter, tel que mentionné ci-dessus. L'utilisation répandue de polices qui compliquent leur différenciation visuelle peut entraîner une confusion importante (par ex., « OldNews » versus « OldNews »). Le groupe n'est pas au courant des tentatives visant à éviter ces problèmes en limitant les caractères autorisés à un sous-ensemble LDH, à l'exception de la zone racine où les caractères numériques sont exclus des noms de domaine pour éviter toute confusion avec les adresses IP.

Table : Ambiguïtés visuelles LDH dans le système existant¹⁰

G	6	
I	1	1
O	0	
S	5	
U	V	
Z	2	

⁹ <http://www.iana.org/domains/idn-tables>

¹⁰ Table issue de « Punycode: A Bootstring encoding of Unicode for Internationalised Domain Names in Applications » (RFC 3492), Section 5. <http://tools.ietf.org/html/rfc3492>

5.2 Considérations orthographiques

Les différences entre l'orthographe américaine et britannique, telles que « organization » et « organisation » sont fréquemment soulignées mais le groupe ignore si cela a été inclus dans des politiques de registre. L'enregistrement d'une étiquette sous une forme ne donne aucun droit sur l'autre forme et le titulaire du nom qui souhaite éviter le risque de confusion chez l'utilisateur a toujours eu la liberté d'enregistrer les deux formes. Une des raisons pour lesquelles cela n'était pas perçu comme une source de préoccupation était le fait que les étiquettes étaient considérées comme des systèmes mnémoniques et que toutes les propriétés semblables au mot qu'elles affichaient possédaient un intérêt secondaire.

Conformément à RFC 3454 :

« Les utilisateurs s'attendent probablement à ce que toutes les équivalences orthographiques soient acceptées comme étant valides ou qu'aucune ne le soit.

Des exemples d'orthographe équivalents sont 'theater' vs. 'theatre', et 'hemoglobin' vs. 'hæmoglobin' en anglais américain vs. anglais britannique ».¹¹

« Des orthographe spécifiques à certaines langues telles que 'Aepfel' vs. 'Äpfel', qui sont parfois considérées comme équivalentes en allemand, peuvent ne pas être considérées comme telles dans d'autres langues »¹².

Dans le document de la charte initiale du projet concernant les questions relatives aux variantes des IDN, un exemple latin a été donné d'un hypothétique domaine de premier niveau pour la ville de Cologne (Köln), Allemagne. Les règles orthographiques de la langue allemande permettent d'écrire « .koeln » ou « .köln ».

Pour l'étude VIP, la mesure dans laquelle un internaute allemand pourrait considérer ces deux étiquettes comme étant similaires ou les deux orthographe « .strasse » et « .straße » pour utiliser un autre exemple, illustre une situation dont il faut tenir compte mais ne se prête pas à la résolution algorithmique. Dans tous les cas, cela met en évidence à quel point ces équivalences mathématiques ne peuvent pas être imposées à l'alphabet latin.

6. Taxonomie des variantes

Différents termes ont été utilisés pour désigner le concept d'équivalence mais aucun n'a encore été accompagné d'une définition qui soit adéquate dans tous les contextes ou cela est nécessaire. Cette difficulté terminologique est probablement due à des perceptions et à des besoins différents d'une communauté linguistique à l'autre ou découle en partie de ces derniers. Cela est même vrai parmi les communautés de langue qui partagent les mêmes écritures. Cela est particulièrement vrai pour l'écriture latine qui est utilisée pour écrire un nombre de langues supérieur à toutes les écritures combinées.

La relation de variante la plus évidente dans l'alphabet latin est entre les majuscules et les minuscules d'une lettre donnée. Comme indiqué, le DNS permet de ne pas se soucier de cet aspect tant que le

¹¹ « Préparation des chaînes internationalisées ('stringprep') » (RFC 3454).

<http://tools.ietf.org/html/rfc3454>

¹² Ibid.

répertoire est restreint à l'ensemble LDH. La forme de l'égalité qu'il transmet à ces caractères est souvent considérée comme étant applicable à d'autres relations de variantes, en reconnaissant le fait qu'un cadre descriptif correspondant plus large serait nécessaire mais sans en proposer aucun.

Il est en outre communément admis que toute complexité supérieure pouvant être liée à la mise en œuvre de ce cadre spéculatif ne représenterait pas une tâche démesurée pour la communauté technique. De fait, l'insensibilité à la casse ASCII est soutenue par une opération mathématique banale. Concevoir des moyens pour l'exécution algorithmique de toute autre forme de relation de variante — de la simple convertibilité des majuscules et des minuscules des lettres de l'alphabet latin au-delà de l'ensemble LDH, à l'équivalence entre les caractères qui sont considérés comme semblables à certains autres égards, n'est en aucun cas une entreprise facile. Dans tous les cas, les caractères utilisés dans les systèmes d'écriture de plusieurs langues ont souvent des relations de variantes qui diffèrent d'une langue à l'autre, et là aussi une quantification universelle serait également nécessaire.

Le protocole IDNA aborde les préoccupations centrales en matière de relations de variantes dans l'écriture latine, et traite, par exemple, directement de la sensibilité à la casse (quoique d'une manière inattendue). D'autres problèmes de variantes sont si profondément dépendants des considérations locales qu'une solution au niveau du protocole général semble difficile à trouver.

6.1. Formes d'affichage

Un caractère abstrait, mettons la LETTRE MINUSCULE LATINE M (selon la présentation Unicode formelle des noms de caractères) est instancié par un ou par plusieurs glyphes, chacun ayant une forme d'affichage spécifique. C'est à ce niveau que des attributs tels que les empattements, les italiques et la graisse sont quantifiés.

m *m* **m** *m* **m** *m*

Une police est une collection nommée de glyphes (par ex., Helvetica ou Times Toman), que vont supporter un ou plusieurs scripts et qui peut utiliser le même glyphe dans des scripts différents. Un exemple classique est celui de la LETTRE MINUSCULE LATINE A et de LA LETTRE MINUSCULE CYRILLIQUE A :

a *a*

Certaines polices sont délibérément conçues pour rendre de telles distinctions aussi visibles que possible mais l'utilisateur type est équipé de polices spécifiques aux paramètres locaux qui ne supportent normalement pas les sous-ensembles extensifs du répertoire Unicode. La représentation des différents caractères abstraits par des glyphes indiscernables, en utilisant des polices couramment disponibles dans des résolutions d'affichage normal, a donc un impact important sur les considérations spécifiques aux autres communautés quand il s'agit d'évaluer la relation de variante potentielle entre deux points de code.

6.2. Les glyphes identiques

Les tables d'écriture latine fournies dans les Annexes B2 et B3 (pour les explications voir la section 6.4) ne contiennent qu'une instance de deux caractères abstraits ayant le même glyphe. Il s'agit de la

LETTRE MINUSCULE LATINE E CULBUTÉE (U+01DD), et de la LETTRE MINUSCULE LATINE SCHWA (U+0259) :

ə ə

Dans tout système d'écriture dans lequel un des deux apparaît, l'autre n'est pas inclus comme substitut. Cela donne lieu au débat portant sur l'utilité d'indiquer la relation de variante entre eux dans tout répertoire IDN dans lequel ils apparaissent. L'absence d'attributs spécifiques dans la zone racine indique cependant un besoin de réglementation contextuelle explicite concernant l'utilisation des ces caractères dans les étiquettes TLD.

6.3. Casse

La différence la plus immédiate entre la présentation d'une chaîne de lettres latine limitée à l'ensemble LDH et une chaîne dans l'ensemble plus large valide selon l'IDNA, peut être illustré en revenant aux étiquettes de casse mélangées. L'IDNA2008 ne fonctionne pas sur les étiquettes LDH et n'accepte que les lettres minuscules dans les étiquettes qu'il traite. Cela signifie que les étiquettes « clubamerica » et « ClubAmerica » correspondent dans le DNS, mais « ClubAmérica » est invalide et ne peut être utilisé que comme « clubamérica »¹³. Dans ce cas, la présence ou l'absence de l'accent ne devrait pas changer la signification que l'étiquette est destinée à transmettre et le titulaire du nom peut accorder davantage de valeur à la flexibilité dans la représentation de la casse plutôt qu'à la marque diacritique.

Il y a moins de différences dans d'autres situations n'impliquant pas plus qu'une lettre non ASCII. Une personne portant le nom suédois de « Östen » et souhaitant l'utiliser comme étiquette dans un nom de domaine acceptera probablement une restriction en minuscule plutôt que de renoncer au tréma et d'utiliser l'étiquette « osten » qui signifie fromage en suédois. En allemand, l'expansion d'un « ö » à un « oe » est une alternative établie mais avec une applicabilité limitée aux noms propres. Cependant, cela ne s'applique pas au cas du suédois dans lequel le « ö » ne se décompose pas en deux lettres et n'est pas considéré comme un « o » diacritiquement marqué. C'est une lettre à part entière (la dernière dans un alphabet de 29 lettres) et les deux lettres ne sont pas des variantes l'une de l'autre. Dans des contextes de saisie de texte, si un « ö » n'est pas disponible la solution est d'utiliser un « o ».

6.4. Variantes décoratives et contrastives

Dans les cas où les signes diacritiques clarifient la prononciation mais ne changent pas la signification d'un mot — « naïvete » versus « naïveté » — leur utilisation est dénommée « décorative ». Dans les cas où l'utilisation de signes diacritiques modifie ou clarifie le sens — « resume » versus « résumé » — leur utilisation est dénommée « contrastive ». Indépendamment de la distinction entre l'utilisation décorative et contrastive, les lettres comportant des signes diacritiques sont généralement appelées « décorées ». Ce même terme est aussi appliqué aux lettres qui apparaissent dans des formes modifiées, qu'elles portent un signe ou pas :

n ŋ ƚ

Une illustration plus large des lettres latines décorées disponible dans les IDN, avec les lettres de base qui ne sont pas utilisées en anglais et ne figurent donc pas dans l'ensemble LDH, est présentée dans les

¹³

IDNA2003 permet les majuscules mais normalise les minuscules. IDNA2008 donne la liberté aux responsables du logiciel pour inclure la normalisation de casse dans le traitement des IDN.

tables de script incluses dans les Annexes B2 et B3. Celles-ci offrent une liste exhaustive des points de code étiquetés comme lettres dans les blocs de l'alphabet latin dans la version 6.1.0 de l'Unicode (qui est essentiellement identique à la version 6.0.0 actuellement en vigueur et dont le lancement est prévu pour février 2012) et qui sont valides en vertu du protocole IDNA2008. Les blocs phonétiques incluent également des éléments latins qui sont utilisés dans des systèmes d'écriture qui apparaissent dans l'espace d'IDN ou qui pourraient vraisemblablement être utilisés dans l'avenir. Actuellement les tables comprennent les points de code des blocs phonétiques qui sont spécialement nécessaires pour contenir l'« alphabet de référence en Afrique », ou ARA. L'ARA est utilisé à de nombreuses reprises dans les illustrations présentées ici parce qu'il est bien adapté à cette fin et pour souligner le fait que les préoccupations concernant les variantes de caractères dans l'alphabet latin ne se limitent pas aux langues européennes. Beaucoup de points de code inclus dans l'ARA apparaissent également dans d'autres alphabets basés sur le latin.

Les blocs latins ont été considérés jusqu'au milieu du latin étendu C, en s'arrêtant à U+2C74. Les points de code présentés à partir de là sont utilisés dans le discours analytique sur la langue mais ils n'apparaissent dans aucun des systèmes d'écriture quotidiens connus par le groupe, à la seule exception de U+A78C mentionné ci-dessous. Cependant, ces points de code peuvent être inclus dans des répertoires locaux par des registres où il est spécifié qu'ils sont nécessaires, qui les intègrent dans des chaînes qui clarifient leur fonction (en utilisant de moyens basés sur l'étiquette pour résoudre les problèmes concernant les caractères) et qui sont capables de préciser les limitations contextuelles requises pour éviter leur utilisation inadéquate dans d'autres contextes. Il semble raisonnable d'affirmer que ces points de code ne possèdent actuellement aucune utilité qui justifie leur inclusion dans la zone racine, et le groupe recommande qu'ils soient entièrement évités dans ce contexte.

6.5. Caractères pré-composés

Les blocs en latin d'Unicode comprennent un nombre immense de points de code isolés qui représentent un caractère de base combiné avec un ou plusieurs signes diacritiques ou qui apparaissent dans d'autres formes décorées. D'autres signes peuvent néanmoins être appliqués à ces caractères pré-composés. Aucun signe combinatoire séparé n'est inclus dans la table mais beaucoup de signes qui apparaissent dans les caractères pré-composés sont aussi valides séparément sous le protocole IDNA dans leurs formes combinatoires. Toute personne qui collationne une table de caractères locale sera donc libre d'inclure une des formes suivantes du même caractère :

le caractère recomposé,

LETTRE LATINE MINUSCULE C AVEC CÉDILLE ET ACCENT AIGU ç

qui peut aussi être représenté en combinant des caractères à deux points de code séparés,

LETTRE LATINE MINUSCULE C AVEC CÉDILLE +ACCENT AIGU COMBINATOIRE ç
LETTRE LATINE MINUSCULE C AVEC ACCENT AIGU + CÉDILLE COMBINATOIRE
ç

ou à trois points de codes séparés,

LETTRE LATINE MINUSCULE C +ACCENT AIGU COMBINATOIRE + CÉDILLE
COMBINATOIRE ç

LETTRE LATINE MINUSCULE C+ CÉDILLE COMBINATOIRE +ACCENT AIGU COMBINATOIRE ć

Les représentations de point de code un, deux et trois sont équivalentes à tous points de vue et par conséquent une instanciation non ambiguë d'une série de caractères de variantes. Les caractères pré-composés dans le deuxième et dans le troisième exemples peuvent être représentés dans les deux formes pré et post combinées, ce qui donne deux autres séries. Le nombre de séries de variantes pouvant être sélectionnées récursivement dans la table de cette manière serait vraiment extraordinaire si de telles relations devaient être indiquées séparément. En termes pratiques ce problème peut être largement ou entièrement évité par la normalisation effectuée dans l'IDNA2008.

6.6. Signes combinatoires

Dans la mesure où les signes combinatoires séparés sont exclus de la table il n'est pas nécessaire d'indiquer l'ensemble de formes variantes qu'ils peuvent générer. Il y a cependant des systèmes d'écriture qui utilisent des caractères qui ne sont pas inclus dans la table Unicode dans leur forme pré-composée. L'ARA comprend plusieurs consonnes soulignées dont certaines sont disponibles sous formes pré-composées, comme la LETTRE MINUSCULE LATINE T LIGNE SOUSCRITE (U+1E6F) :

t

tandis que d'autres requièrent des signes combinatoires séparés, comme la LETTRE C LATINE MINUSCULE + DIACRITIQUE TRAIT SOUSCRIT COMBINATOIRE (U+0063 U+0332) :

c

Le glossaire Unicode se réfère à un caractère composite qui n'existe pas dans la forme pré-composée en tant que « groupe de graphèmes ».¹⁴

Il y a aussi des situations où des signes semblables, telles que le TIRET BAS COMBINATOIRE et le MACRON SOUSCRIT COMBINATOIRE peuvent être utilisés de façon interchangeable ; d'autres signes plus distinctifs du point de vue du graphème, telle que le POINT SOUSCRIT COMBINATOIRE; sont également des éléments de ce qui pourrait être traité comme une série de variantes fonctionnelle. Cependant, il est rare qu'il existe une relation de variantes intrinsèque entre les signes eux-mêmes. Cette caractéristique est modulée par le caractère de base avec lequel elles sont combinées et par le système d'écriture dans lequel les caractères combinés sont utilisés. Le traitement des signes comme des variantes les uns des autres dans un système d'écriture n'exclut en aucune manière la possibilité de les utiliser de façon indépendante dans un autre système.

Les caractères combinés peuvent facilement être placés dans une table sans ajouter l'énumération séparée des signes combinatoires en indiquant les relations de variantes, le cas échéant. Toutefois, il est nécessaire qu'il y ait une certaine forme de contrainte contextuelle sur les caractères auxquels un signe donné peut être appliqué. Ceci représente un bon exemple d'une situation dans laquelle un grave problème de gestion des variantes peut être évité au niveau du protocole. L'initiation d'une telle action n'est cependant pas une préoccupation actuelle. Le protocole IDNA2008 reconnaît la nécessité de réduire les formes variantes de ce type à une seule représentation canonique et normalise algorithmiquement tous les points multicode alternatifs dans la forme pré-composée, s'il en existe une. Dans le cas contraire il est toujours nécessaire d'appliquer une politique de contrainte.

14

« Une unité de texte horizontalement segmentable, consistant en une base de graphèmes ...avec un nombre quelconque de marques de non espacement qui lui est appliqué ».

Lorsqu'elles sont incluses dans une étiquette Unicode candidate, les cinq formes dotées de signé de la lettre 'c' dans l'exemple précédent, apparaissent indistinctement comme :

U+0061 U+1E09 U+0065	; açe
U+0061 U+00E7 U+0301 U+0065	; açe
U+0061 U+0107 U+0327 U+0065	; açe
U+0061 U+0063 U+0301 U+0327 U+0065	; açe
U+0061 U+0063 U+0327 U+0301 U+0065	; açe

Bien que seule la première forme soit une étiquette U valide, les autres formes sont normalisées avant de générer ce qui sera alors la même étiquette A pour toutes — « xn--ae-ess ». Il est important de noter que cette normalisation n'est pas effectuée par l'algorithme d'encodage lui-même. Chacune des cinq étiquettes U candidates, autant les valides que les non valides, possède sa propre forme Punycode.

U+0061 U+1E09 U+0065	; açe	# ae-ess
U+0061 U+00E7 U+0301 U+0065	; açe	# ae-4ia66s
U+0061 U+0107 U+0327 U+0065	; açe	# ae-vla09s
U+0061 U+0063 U+0301 U+0327 U+0065	; açe	# ace-ldc1qd
U+0061 U+0063 U+0327 U+0301 U+0065	; açe	# ace-ldc3q

[Digression : Il existe des applications de « conversion d'IDN » en ligne qui effectuent l'encodage Punycode et préfixent le résultat avec 'xn--'. Il faut donc s'attendre à ce que certains titulaires de nom prospectifs soumettent des demandes avec l'étiquette A présentées sous une forme non valide. En outre, les lettres pré-composées sont souvent plus typographiquement stables que leurs points multicode équivalents. Les erreurs classiques dans l'affichage de ces derniers sont le placement erroné des signes combinatoires ; par exemple, dans le cas açe, avec le passage d'un des signes combinatoires de 'c' à 'e' sans changer les points de code. On peut donc s'attendre à ce que les demandes d'étiquettes U soient présentées de manière incorrecte. Il est concevable que l'apparition simultanée d'étiquettes A et U incorrectes ne permette pas d'identifier l'étiquette qui est réellement demandée.

6.7. Ponctuation

Le seul signe de ponctuation permis dans le répertoire LDH est le TRAIT D'UNION SIGNE MOINS (U+002D). Le répertoire IDNA2008 comprend un certain nombre de caractères qui ne sont pas des signes de ponctuation mais des glyphes qui sont difficiles à différencier les uns des autres (et possèdent des noms Unicode qui sont des hybrides lettres/signes de ponctuation). L'utilisation de ces caractères exige évidemment beaucoup de prudence. Un exemple est représenté par LETTRE MODIFICATIVE VIRGULE CULBUTÉE (U+02BB) qui est utilisé dans de nombreux contextes, dont un est le « okina », qui indique le coup de glotte dans 'Hawai'i', par exemple :

Un autre exemple de caractère utilisé pour indiquer un coup de glotte mais qui est une lettre latine et non pas un signe de ponctuation est la LETTRE MINUSCULE LATINE SALTILLO (U+A78C) :

Son annotation Unicode signale que « elle est largement utilisée dans de nombreuses langues au Mexique et dans d'autres régions, y compris en Izere, au Nigeria » et elle peut donc apparaître dans les

tables d'écriture latine locales. Des considérations similaires sont valables pour l'exemple 'okina' dans un contexte de langue polynésienne. Le seul dénominateur commun pour réunir ces caractères dans une série de variantes de caractères est leur similitude graphémique. Autrement, ils sont utilisés dans des systèmes d'écriture différents qui ne les reconnaissent pas mutuellement. En tout cas, la variante APOSTROPHE (U+0027), qui de toute évidence serait la préférée, ne peut pas apparaître dans une étiquette U. Le problème est que cela constitue une représentation de repli partagée par les autres formes qui rarement sont indiquées directement sur les claviers et qui généralement sont remplacées par une apostrophe. Cela s'associe avec un taux d'erreur élevé quand une étiquette qui contient un caractère similaire à une apostrophe est transcrit sur papier. La probabilité de substitution erronée dans des situations créées numériquement est tout aussi élevée.

Le problème avec des caractères qui ressemblent à des apostrophes est aggravé par l'apparition alternée du caron sur certaines consonnes, soit comme le signe nominale attendu observé dans la première des illustrations ci-dessous soit comme signe similaire à une apostrophe dans le deuxième exemple. Ces deux interprétations différentes de la LETTRE MINUSCULE LATINE T CARON (U+0165) fournissent une base potentielle pour une autre forme de collection de variantes :

ť ť

On trouve une autre similitude lettre/signe de ponctuation dans l'ARA. La LETTRE LATINE COUP DE GLOTTE (U+0294) ressemble au POINT D'INTERROGATION (U+003F), qui ne peut pas apparaître dans une étiquette U.

? ?

En prenant un des quatre caractères utilisés pour représenter les consonnes à clic dans de nombreuses langues parlées en Afrique du Sud (mais qui ne sont pas incluses dans l'ARA), la LETTRE LATINE CLIC RÉTROFLÈXE (U+01C3) ressemble, et dans beaucoup de polices, partage le même glyphe que le POINT D'EXCLAMATION (U+0021), qui est aussi absent des étiquettes U :

! !

pour des raisons évidentes, ces caractères, avec tous les autres caractères valides qui ressemblent à des signes de ponctuation non valides ne doivent pas pouvoir être utilisés comme substituts.

6.8. Points de code des systèmes d'écriture non reliés entre eux

Une préoccupation similaire est causée par les lettres qui résident dans des points de code différents, qui sont utilisés dans des systèmes d'écritures non reliés et qui sont généralement représentés par des glyphes similaires, voire identiques. Il s'agit de la LETTRE MINUSCULE LATINE E CULBUTÉE (U+01DD), et de la LETTRE MINUSCULE LATINE SCHWA (U+0259) :

ə ə

L'utilisation du même glyphe pour les deux caractères dans de nombreuses polices couramment déployées n'implique pas que les caractères soient des formes variantes les uns des autres, et aucun système d'écriture n'est supposé inclure plus d'un d'entre eux non plus. Cependant, des problèmes imprévus pourraient surgir si un répertoire incluant un de ces points de code devait être compilé sans tenir compte de l'autre. Par conséquent, outre toute colonne dans une table de scripts généralisée qui indique des relations de variantes entre points de code, une colonne séparée pourrait indiquer les points de code qui ne devraient pas apparaître dans le même répertoire localisé sans un examen détaillé. Si les

deux sont en effet nécessaires, les règles contextuelles pour leur utilisation doivent être clairement précisées et appliquées.

L'inverse s'applique également aux caractères dans un système d'écriture isolée. L'ARA comprend à la fois LA LETTRE MINUSCULE LATINE A (U+0061) et la LETTRE MINUSCULE LATINE ALPHA (U+0251), qui ne peuvent pas être considérées comme des variantes l'une de l'autre.

a

ɑ

Toutefois, il est clair que la deuxième d'entre elles a besoin d'être restreinte pour être utilisée dans des étiquettes écrites avec un alphabet qui les comprenne toutes les deux. Là encore, l'inclusion de cette dernière dans une étiquette TLD requiert une réglementation contextuelle explicite.

Il y a un défi de taille à relever concernant l'équilibre nécessaire entre le besoin de minimiser les possibilités d'abus générées par la disponibilité simultanée de ces deux caractères et les attentes légitimes concernant l'orthographe — voire la nécessité — dans des contextes linguistiques où les différences entre les deux lettres sont reconnues et comprises. Ceci s'applique à beaucoup d'autres situations présentées ici.

L'expertise nécessaire pour affronter ce type de situations sans risque de discrimination culturelle va au-delà des connaissances courantes du système d'écriture de la communauté respective et doit suivre l'évolution de ses règles orthographiques. Cela exige également une connaissance approfondie des environnements de travail locaux, notamment en ce qui concerne la capacité des moteurs de traduction couramment utilisés à afficher correctement les caractères décorés dans des polices largement déployées et une prise de conscience de toutes les divergences entre le point de code d'un caractère donné généré par un clavier courant et le point de code Unicode de ce caractère. Il existe encore la nécessité de se rapprocher de la communauté pour être en mesure de suivre le développement dans l'environnement de l'utilisateur et d'assurer l'ajustement rapide du système IDN afin de refléter les changements qui ont eu lieu.

Si au sein d'une communauté linguistique il existe des variations au niveau des points de code produits pour un caractère donné dans une saisie de texte par clavier, la préparation d'un répertoire unique de caractères IDN pour cette communauté devrait aussi prendre en compte ces différences. (C'est une question importante pour la gestion des variantes, mais ne pose pas de difficultés particulières pour l'écriture latine). Une connaissance similaire du profil local est nécessaire pour évaluer si un traitement spécial doit éventuellement être accordé aux caractères correspondant à différents points de code qui, au moins dans certains contextes, sont généralement affichés au moyen de glyphes impossibles à distinguer y compris par les membres les plus érudits des communautés linguistiques qui s'en servent.

Outre une certaine familiarité avec le profil local, l'évaluation et la gestion de ces risques exige de bien comprendre leurs effets sur des utilisateurs extérieurs à la communauté qui, d'une part sont capables de lire aisément l'écriture de base utilisée pour les étiquettes IDN, mais d'autre part sont incapables de reconnaître des distinctions clé entre des points de code attribués à des graphèmes familiers.

7. Similitude visuelle avec d'autres écritures

Le groupe signale que certains points de code attribués aux écritures cyrillique et grecque sont généralement représentés au moyen de glyphes, aussi utilisés pour afficher des points de code attribués à

l'écriture latine. D'autres écritures non incluses dans l'étude VIP, telles que la cherokee, présentent le même type de chevauchement avec l'écriture latine.

Un exemple de cette propriété est la chaîne « paypal », écrite intégralement en écriture latine, versus « paypal » écrite avec le « а » cyrillique. Ces deux chaînes sont visuellement identiques, même pas similaires, mais sont encodées différemment.¹⁵

Les équipes d'étude des cas latin et cyrillique se sont réunies pour discuter le risque potentiel de confusion et de préjudice associé à l'utilisation mixte de points de code attribués à deux écritures différentes. Les équipes n'avaient connaissance d'aucune langue écrite ni en latin ni en cyrillique nécessitant la combinaison de ces écritures pour exprimer une idée ou un terme, ou possédant essentiellement une valeur mnémonique. Compte tenu de la confusion que cette combinaison pourrait entraîner alors qu'elle n'est pas vraiment nécessaire, les groupes ont convenu qu'elle devrait être évitée. Le problème plus épineux du « risque de confusion associé à une écriture » - c'est-à-dire la juxtaposition d'étiquettes d'écritures uniques ne pouvant pas être différenciées visuellement - n'a pas été abordée, car elle dépassait le cadre des études consacrées spécifiquement à une écriture. Elle devrait par conséquent être considérée en détail dans l'examen transversal.

8. Évaluation de candidatures aux TLD avec variantes

L'ICANN met à disposition un outil destiné à tester des chaînes à l'aide de l'« algorithme SWORD ».¹⁶ Celui-ci a fait l'objet d'essais de pré-production en ligne pendant la période précédant le lancement du programme des nouveaux gTLD. D'après sa notice, cet outil est capable de prendre en charge un ensemble de langues : « *Cette version pré-production de l'algorithme prend en charge la plupart des caractères courants des écritures arabe, chinoise, cyrillique, devanagari, grecque, japonaise, coréenne et latine. ... Les écritures latine et grecque appartiennent à la famille des écritures européennes et pourraient être comparées. ... Cette version de l'algorithme n'inclut pas la validation des chaînes du candidat en vue de confirmer leur respect des protocoles IDNA* ». [italique ajoutée]

L'intérêt d'un examen des candidatures TLD au moyen de cette méthode ou de tout autre mécanisme automatisé d'évaluation de chaînes tient à ce qu'un tel mécanisme devrait permettre de signaler et d'identifier, par exemple, deux candidatures différentes rivalisant pour .koeln et .köln respectivement. Celles-ci pourraient être considérées comme un ensemble conflictuel, mais une telle suggestion va au-delà du mandat ou de la mission du groupe de rédaction initial.¹⁷

9. Impact des variantes sur le fonctionnement des opérateurs de registre/ bureaux d'enregistrement.

¹⁵ « Version préliminaire de la norme technique Unicode 39 : Mécanismes de sécurité Unicode », Section 4. <http://unicode.org/reports/tr39/tr39-1.html>

¹⁶ <http://icann.sword-group.com/icann-algorithm/Default.aspx>

¹⁷ Il faut signaler que le site SWORD affichait une erreur quand ces deux chaînes étaient comparées ou que köln était testé — « *köln : L'étiquette ne doit comporter que des lettres minuscules en pur ASCII ; 'ö' (U+00F6) a été trouvé en position 1.* »

La gestion de variantes de TLD – par exemple, la délégation d'un domaine sous une variante de TLD- peut entraîner des difficultés au niveau des opérateurs de registre et des bureaux d'enregistrement. Du point de vue de l'opérateur de registre, les éléments suivants devraient être mis au point pour l'introduction d'une variante : des politiques et des procédures, y compris en matière de lancement et de toute possible synchronisation des domaines sous la variante concernée ; des redevances pour les bureaux d'enregistrement, des procédures de transfert y compris des options de verrouillage de domaine.

Du point de vue du bureau d'enregistrement, des consultations approfondies avec les bureaux d'enregistrement agréés de l'opérateur de registre qui envisage d'introduire une variante devraient être prévues pour tout développement associé à des variantes. Cela tient à plusieurs raisons, dont la plus pressante concerne le fait que pour la plupart des TLD, les domaines sont enregistrés via des bureaux d'enregistrement et non pas directement par les opérateurs de registre. Le lancement d'une variante sans l'implication des bureaux d'enregistrement peut conduire à ce qu'un tout petit groupe d'utilisateurs privilégiés en soient bénéficiés. Lors de l'introduction des IDN au deuxième niveau, de nombreux bureaux d'enregistrement n'ont pas pris immédiatement en charge les écritures au fur et à mesure de leur introduction par l'opérateur de registre.

C'est pourquoi l'impact des variantes sur les opérateurs de registre/bureaux d'enregistrement devrait être soigneusement exploré, non seulement au niveau de leur fonctionnement –qui devrait être dans tous les cas examiné- mais plus particulièrement au niveau de leur introduction, avec l'accent mis sur les attentes des utilisateurs finaux.

L'impact des variantes de TLD sur les opérateurs de registre et les bureaux d'enregistrement peut dépendre largement des différentes méthodes de mise en œuvre. Les parties prenantes devront s'investir largement dans toute proposition de mise en œuvre afin de vérifier que les services fournis par les opérateurs de registre et les bureaux d'enregistrement assurent un fonctionnement du DNS stable, sûr, cohérent et non ambigu. Dans ce cadre, il faudra veiller à assurer une communication claire, ainsi qu'une bonne connaissance des variantes de TLD, de façon à limiter toute possible confusion des IDN chez l'utilisateur final, le registrant, le bureaux d'enregistrement et l'opérateur de registre. Des aspects tels que le comportement des applications, la résolution et l'enregistrement de services, le service WHOIS et la logique commerciale doivent être examinés pour déterminer si ces objectifs sont atteignables.

10. Impact des variantes sur les procédures de résolution de litiges

10.1. Règles existantes en matière de résolution de litiges

Un grand nombre d'opérateurs de registre TLD suivent la Procédure uniforme de résolution de litiges en matière de noms de domaine (UDRP). En vertu de cette procédure, les noms de domaine liés à des marques commerciales faisant l'objet d'un litige ne peuvent être annulés, suspendus ou transférés par un bureau d'enregistrement que si ledit litige a été résolu par un accord entre les parties, une action en justice, ou un arbitrage. Les litiges qui découleraient de l'enregistrement abusif de noms de domaine (par exemple, le cybersquattage) peuvent être traités par des procédures administratives accélérées que le titulaire des droits de marque initie par le dépôt d'une plainte auprès d'un fournisseur de services de résolution des litiges approuvé.

Une commission d'arbitrage vérifie si le domaine en litige (l'ensemble de la chaîne ou une composante de celle-ci) est identique ou similaire à s'y méprendre à une marque commerciale ou à une marque de service dont le plaignant détient les droits, et s'il comporte un risque de confusion avec une marque de

commerce ou de service en ce qui concerne la source, le parrainage, l'affiliation ou l'approbation du site Web du registrant ou d'un produit ou service du site Web ou de l'emplacement du registrant.

Aspects à prendre en considération pour évaluer si la similarité d'un domaine avec une marque commerciale ou une marque de service peut porter à confusion :

- a) la partie du domaine constituant l'étiquette enregistrée (normalement la partie de deuxième niveau à gauche du domaine de premier niveau) ; ou
- b) dans des cas exceptionnels, autant les composantes à gauche que les composantes à droite du point (les composantes de premier niveau et de deuxième niveau du domaine).

10.2. Impact des variantes latines

Si des variantes d'un seul TLD possèdent une similitude propice à confusion et que l'opérateur de registre permet à plus d'une partie d'enregistrer des domaines dans l'ensemble conflictuel, le nombre de litiges risque de s'accroître. En effet, le nombre de litiges risque d'augmenter suite à des politiques permettant par exemple, qu'un opérateur de registre ayant deux variantes de TLD – « .koeln » et « .köln »- délègue un domaine sous « .koeln » à un registrant et la même étiquette à un autre registrant sous « köln » -par exemple, « hotel-airport.koeln » enregistré par une partie et « hotel-airport.köln » par une autre partie.

11. Conclusions

Vous trouverez ci-dessous les conclusions du groupe d'étude qui s'est penché sur le cas latin :

1. Le répertoire de caractères latins pouvant être utilisés dans les IDN comporte plusieurs formes qui présentent des similitudes visuelles. Une évaluation minutieuse s'avère nécessaire lorsqu'il s'agit de classer un sous-ensemble du répertoire complet en vue d'une utilisation locale. La résolution des problèmes que cela peut soulever dépend toutefois en grande mesure des conventions orthographiques locales. Celles-ci traitent souvent les mêmes caractères de différentes manières. Les chaînes ayant une similitude propice à confusion dans le cadre d'une langue peuvent ne pas avoir cette connotation dans une autre langue.
2. Etant donné que l'alphabet latin est celui qui est utilisé par le plus grand nombre de communautés linguistiques, essayer de donner un aperçu global des besoins de chacune d'entre elles est une entreprise presque impossible. Essayer de faire une synthèse de ces besoins malgré tout serait faire preuve d'un manque de compréhension envers les communautés qui n'ont pas encore rejoint la discussion portant sur l'IDN.
3. Compte tenu des deux points précédents, le groupe d'étude ne trouve pas de justification pour traiter catégoriquement tout point de code attribué à un élément de l'écriture latine comme étant équivalent à toute autre point de code. Le groupe ne pense pas d'ailleurs qu'une telle justification existe au-delà de ce qui a été déjà incorporé dans le protocole IDNA2008.
4. Le processus de candidature aux TLD de l'ICANN ne devrait pas donner suite à des demandes de chaînes latines multiples en partant du principe qu'elles sont des variantes les unes des autres. Si un examen rigoureux s'avère nécessaire pour évaluer le risque de confusion des étiquettes TLD proposées, cela n'en fait pas pour autant des variantes au sens de l'étude VIP.
5. Même si le protocole IDNA permet l'affichage d'un large nombre de points de code dans les étiquettes TLD basées sur l'écriture latine, le groupe d'étude ne préconise pas l'inclusion de points de code dans la zone racine du DNS sans une démonstration spécifique de son bien fondé dans le contexte linguistique concerné. Contrairement aux zones TLD, qui peuvent posséder des attributs liés à des spécificités locales à partir desquels il est possible d'évaluer ce type de situations, la racine est une ressource globalement partagée où doivent trouver leur place toutes les communautés linguistiques de la façon la plus équitable possible.
6. La table d'écriture latine pour la zone racine peut donc être un classement descriptif des points de code apparaissant réellement dans cette zone plutôt qu'une tabulation prescriptive de points de code pouvant y être inclus sans aucun examen ultérieur.

ANNEXES

A. Équipe d'étude du cas latin

L'équipe d'étude du cas latin a été constituée et s'est réunie pour la première fois à l'occasion de la 41^e réunion de l'ICANN tenue à Singapour en juin 2011. Par la suite, elle a tenu des réunions téléphoniques bimensuelles afin de considérer des questions clé, de discuter et d'analyser leur pertinence et leurs implications vis à vis de l'écriture latine, et d'élaborer le présent rapport.

A1. Membres de l'équipe

Les membres de l'équipe d'étude de cas sont :

Nom	Rôle.
Jothan Frakes	Coordinateur de l'étude de cas
Harald Alvestrand	Membre de l'équipe
Andrzej Bartosiewicz	Membre de l'équipe
Eric Brown	Membre de l'équipe
Cary Karp	Membre de l'équipe (représentant l'organisation hôte)
Nadya Morozova	Membre de l'équipe
Francisco Obispo	Membre de l'équipe
Giovanni Seppia	Membre de l'équipe
Will Shorter	Membre de l'équipe
Wil Tan	Membre de l'équipe
Avri Doria,	Observateur
Leo Vegoda	Liaison étude de cas
Francisco Arias	Expert en Opérations de registre
Kim Davies	Expert en Sécurité
Nicholas Ostler	Expert en Linguistique
Steve Sheng	Expert en Politiques
Andrew Sullivan	Expert en Protocoles

A2. Déclarations d'intérêt

Dans un souci de transparence, on a demandé aux membres de l'équipe de fournir des déclarations écrites attestant de l'intérêt qu'ils portent aux différents sujets d'étude et contenant des informations pertinentes à cet égard.

Ces déclarations sont publiées en ligne sur le site Web de l'équipe d'étude de cas.¹⁸

A3. Remerciements

L'équipe remercie l'Internet Infrastructure Foundation, l'organisation hôte de l'étude de cas latin.

L'équipe reconnaît aussi que certains éléments clé du présent rapport sont issus d'une note informative d'introduction apportée par Cary Karp au nom de l'organisation hôte, avec le soutien du Musée Suédois d'Histoire Naturelle.¹⁹

B. Répertoire de caractères latins

Une définition de ce que l'on entend par caractères dans le cadre de l'étude de cas s'avérait nécessaire pour focaliser le travail de l'équipe.

Un sous-ensemble limité de points de code Unicode est considéré un IDN valable conformément aux spécifications du protocole. Ces codes « validés par le protocole » ou « PVALID » déterminent quels points de code peuvent être représentés dans le cadre du protocole IDNA.

Cet ensemble peut par la suite être limité aux points de code qui reflètent ceux appartenant aux différents blocs Unicode de points de code attribués à l'écriture latine.

B1. Attribution de bloc Unicode de points de code à l'écriture latine

Les blocs latins dans la Table de codes de caractères Unicode version 6.1.0 sont :

Nom de la Table Unicode	Série point de code
Latin de base	U+0020 U+007F
Supplément Latin-1	U+0080 U+00FF
Latin étendu A	U+0100 U+017F
Latin étendu B	U+0180 U+024F
Latin étendu C	U+2C60 U+2C7F
Latin étendu D	U+A720 U+A7FF
Latin étendu supplémentaire	U+1E00 U+1EFF
Ligatures latines	U+FB00 U+FB0F
Lettre latines pleine chasse	U+FF00 U+FF5E

B2. Points de code validés par le protocole, attribués à l'écriture latine

Le classement des tables de cette section et de la section suivante est conforme aux principes décrits dans la section 6.4 du rapport. Ces tables sont destinées à soutenir et à illustrer l'action de l'étude et ne sauraient en aucun cas être considérées comme officielles. La déclaration normative des points de code validés par le protocole se trouve dans la norme RFC 5892²⁰, avec la table de référence correspondante dans le Registre de protocole IANA.²¹

Vous trouverez ci-dessous les points de code attribués à l'écriture latine censés apparaître dans les répertoires de caractères IDN classés localement, par ordre alphabétique :

¹⁹ <http://mm.icann.org/pipermail/latin-vip/attachments/20111007/ed0ab5a4/latin-support-brief-final-0001.pdf>

²⁰ <http://tools.ietf.org/html/rfc1591>

²¹ <http://www.iana.org/assignments/idnabis-tables/idnabis-tables.xml>

U+0061 ; a # LETTRE MINUSCULE LATINE A
 U+00E0 ; à # LETTRE MINUSCULE LATINE A ACCENT GRAVE
 U+00E1 ; á # LETTRE MINUSCULE LATINE A ACCENT AIGU
 U+00E2 ; â # LETTRE MINUSCULE LATINE A ACCENT CIRCONFLEXE
 U+00E3 ; ã # LETTRE MINUSCULE LATINE A TILDE
 U+00E4 ; ä # LETTRE MINUSCULE LATINE A TRÉMA
 U+00E5 ; â # LETTRE MINUSCULE LATINE A ROND EN CHEF
 U+0101 ; ā # LETTRE MINUSCULE LATINE A MACRON
 U+0103 ; ă # LETTRE MINUSCULE LATINE A BRÈVE
 U+0105 ; ą # LETTRE MINUSCULE LATINE A OGONEK
 U+01CE ; ǎ # LETTRE MINUSCULE LATINE A CARON
 U+01DF ; ā̄ # LETTRE MINUSCULE LATINE A TRÉMA ET MACRON
 U+01E1 ; ā̇ # LETTRE MINUSCULE LATINE A POINT EN CHEF ET MACRON
 U+01FB ; ǎ̇ # LETTRE MINUSCULE LATINE A ROND EN CHEF ET ACCENT AIGU
 U+0201 ; ̀̀ # LETTRE MINUSCULE LATINE A DOUBLE ACCENT GRAVE
 U+0203 ; ă̂ # LETTRE MINUSCULE LATINE A BRÈVE RENVERSÉE
 U+1E01 ; ȁ # LETTRE MINUSCULE LATINE A ROND SOUSCRIT
 U+1EA1 ; ȁ̇ # LETTRE MINUSCULE LATINE A POINT SOUSCRIT
 U+1EA3 ; ȁ̂ # LETTRE MINUSCULE LATINE A CROCHET EN CHEF
 U+1EA5 ; ȁ̂̇ # LETTRE MINUSCULE LATINE A ACCENT CIRCONFLEXE ET AIGU
 U+1EA7 ; ȁ̂̀ # LETTRE MINUSCULE LATINE A ACCENT CIRCONFLEXE ET GRAVE
 U+1EA9 ; ȁ̂̃ # LETTRE MINUSCULE LATINE A ACCENT CIRCONFLEXE ET CROCHET
 U+1EAB ; ȁ̂̃̇ # LETTRE MINUSCULE LATINE A ACCENT CIRCONFLEXE ET TILDE
 U+1EAD ; ȁ̂̃̂ # LETTRE MINUSCULE LATINE A ACCENT CIRCONFLEXE ET POINT SOUSCRIT
 U+1EAF ; ȁ̂̃̇̂ # LETTRE MINUSCULE LATINE A BREVE ET ACCENT AIGU
 U+1EB1 ; ȁ̂̃̇̀ # LETTRE MINUSCULE LATINE A BREVE ET ACCENT GRAVE
 U+1EB3 ; ȁ̂̃̇̃ # LETTRE MINUSCULE LATINE A BREVE ET CROCHET EN CHEF
 U+1EB5 ; ȁ̂̃̇̃̇ # LETTRE MINUSCULE LATINE A BREVE ET TILDE
 U+1EB7 ; ȁ̂̃̇̃̂ # LETTRE MINUSCULE LATINE A BREVE ET POINT SOUSCRIT
 U+0227 ; ȁ̂̃̇̂̂ # LETTRE MINUSCULE LATINE A POINT EN CHEF
 U+2C65 ; ȁ̂̃̇̂̂̂ # LETTRE MINUSCULE LATINE A BARRÉ
 U+0251 ; α # LETTRE MINUSCULE LATINE ALPHA
 U+00E6 ; æ # LETTRE MINUSCULE LATINE AE
 U+0062 ; b # LETTRE MINUSCULE LATINE B
 U+0180 ; ƀ # LETTRE MINUSCULE LATINE B BARRÉ
 U+0183 ; Ɓ # LETTRE MINUSCULE LATINE B POTENCE
 U+1E03 ; ȃ # LETTRE MINUSCULE LATINE B POINT EN CHEF
 U+1E05 ; ȃ̇ # LETTRE MINUSCULE LATINE B POINT SOUSCRIT
 U+1E07 ; ȃ̂ # LETTRE MINUSCULE LATINE B LIGNE SOUSCRITE
 U+0253 ; Ɓ̂ # LETTRE MINUSCULE LATINE B CROSSE
 U+0063 ; c # LETTRE MINUSCULE LATINE C
 U+00E7 ; ç # LETTRE MINUSCULE LATINE C CÉDILLE
 U+0107 ; ć # LETTRE MINUSCULE LATINE C ACCENT AIGU
 U+0109 ; ȉ # LETTRE MINUSCULE LATINE C ACCENT CIRCONFLEXE
 U+010B ; ȋ # LETTRE MINUSCULE LATINE C POINT EN CHEF
 U+010D ; ȋ̃ # LETTRE MINUSCULE LATINE C CARON
 U+0188 ; ȋ̂ # LETTRE MINUSCULE LATINE C CROSSE
 U+023C ; ȋ̂̂ # LETTRE MINUSCULE LATINE C BARRÉ
 U+0255 ; ȋ̂̂̂ # LETTRE MINUSCULE LATINE C BOUCLÉ
 U+A793 ; ☐ # LETTRE MINUSCULE LATINE C RAYÉ
 U+1E09 ; Ć # LETTRE MINUSCULE LATINE C CÉDILLE ET ACCENT AIGU

U+2184 ; ƿ # LETTRE MINUSCULE LATINE C RÉFLÉCHI
 U+0064 ; d # LETTRE MINUSCULE LATINE D
 U+010F ; d̂ # LETTRE MINUSCULE LATINE D CARON
 U+0111 ; đ # LETTRE MINUSCULE LATINE D BARRÉ
 U+018C ; đ̄ # LETTRE MINUSCULE LATINE D POTENCE
 U+0221 ; đ̸ # LETTRE MINUSCULE LATINE D BOUCLÉ
 U+1E0D ; đ̇ # LETTRE MINUSCULE LATINE D POINT SOUSCRIT
 U+1E0F ; đ̈ # LETTRE MINUSCULE LATINE D LIGNE SOUSCRITE
 U+1E11 ; đ̊ # LETTRE MINUSCULE LATINE D CÉDILLE
 U+1E13 ; đ̋ # LETTRE MINUSCULE LATINE D ACCENT CIRCONFLEXE SOUSCRIT
 U+1E0B ; đ̌ # LETTRE MINUSCULE LATINE D POINT EN CHEF
 U+0256 ; đ̍ # LETTRE MINUSCULE LATINE D HAMEÇON RÉTROFLEXE
 U+0257 ; đ̎ # LETTRE MINUSCULE LATINE D CROSSE
 U+1E9F ; δ # LETTRE MINUSCULE LATINE DELTA
 U+00F0 ; ð # LETTRE MINUSCULE LATINE ED
 U+0065 ; e # LETTRE MINUSCULE LATINE E
 U+00E8 ; è # LETTRE MINUSCULE LATINE E ACCENT GRAVE
 U+00E9 ; é # LETTRE MINUSCULE LATINE E ACCENT AIGU
 U+00EA ; ê # LETTRE MINUSCULE LATINE E ACCENT CIRCONFLEXE
 U+00EB ; ë # LETTRE MINUSCULE LATINE E TRÉMA
 U+0113 ; ē # LETTRE MINUSCULE LATINE E MACRON
 U+0115 ; ě # LETTRE MINUSCULE LATINE E BREVE
 U+0117 ; è̇ # LETTRE MINUSCULE LATINE E POINT EN CHEF
 U+0119 ; ę # LETTRE MINUSCULE LATINE E OGONEK
 U+011B ; ě̇ # LETTRE MINUSCULE LATINE E CARON
 U+0205 ; è̈ # LETTRE MINUSCULE LATINE E DOUBLE ACCENT GRAVE
 U+0207 ; ê̈ # LETTRE MINUSCULE LATINE E BREVE RENVERSÉE
 U+1E15 ; è̊ # LETTRE MINUSCULE LATINE E MACRON ET ACCENT GRAVE
 U+1E17 ; é̊ # LETTRE MINUSCULE LATINE E MACRON ET ACCENT AIGU
 U+1E19 ; ę̇ # LETTRE MINUSCULE LATINE E ACCENT CIRCONFLEXE SOUSCRIT
 U+1E1B ; ę̈ # LETTRE MINUSCULE LATINE E TILDE SOUSCRIT
 U+1E1D ; ě̉ # LETTRE MINUSCULE LATINE E CÉDILLE ET BREVE
 U+1EB9 ; ę̊ # LETTRE MINUSCULE LATINE E POINT SOUSCRIT
 U+1EBB ; ě̋ # LETTRE MINUSCULE LATINE E CROCHET EN CHEF
 U+1EBD ; ě̌ # LETTRE MINUSCULE LATINE E TILDE
 U+1EBF ; é̋ # LETTRE MINUSCULE LATINE E ACCENT CIRCONFLEXE ET AIGU
 U+1EC1 ; è̋ # LETTRE MINUSCULE LATINE E ACCENT CIRCONFLEXE ET GRAVE
 U+1EC3 ; ě̌ # LETTRE MINUSCULE LATINE E ACCENT CIRCONFLEXE ET CROCHET EN CHEF
 U+1EC5 ; ě̍ # LETTRE MINUSCULE LATINE E ACCENT CIRCONFLEXE ET TILDE
 U+1EC7 ; ê̍ # LETTRE MINUSCULE LATINE E ACCENT CIRCONFLEXE ET POINT SOUSCRIT
 U+0229 ; ę̍ # LETTRE MINUSCULE LATINE E CÉDILLE
 U+0247 ; ę̎ # LETTRE MINUSCULE LATINE E BARRÉ
 U+025B ; ε # LETTRE MINUSCULE LATINE EPSILON
 U+01DD ; ə # LETTRE MINUSCULE LATINE E CULBUTÉ
 U+0259 ; ə̇ # LETTRE MINUSCULE LATINE SCHWA
 U+0066 ; f # LETTRE MINUSCULE LATINE F
 U+0192 ; ḟ # LETTRE MINUSCULE LATINE F CROSSE
 U+1E1F ; f̈ # LETTRE MINUSCULE LATINE F POINT EN CHEF
 U+0067 ; g # LETTRE MINUSCULE LATINE G
 U+011D ; ĝ # LETTRE MINUSCULE LATINE G ACCENT CIRCONFLEXE
 U+011F ; ğ # LETTRE MINUSCULE LATINE G BREVE
 U+0121 ; ğ̇ # LETTRE MINUSCULE LATINE G POINT EN CHEF

U+0123 ; ġ # LETTRE MINUSCULE LATINE G CÉDILLE
 U+01E5 ; ĝ # LETTRE MINUSCULE LATINE G BARRÉ
 U+01E7 ; ğ # LETTRE MINUSCULE LATINE G CARON
 U+01F5 ; ħ # LETTRE MINUSCULE LATINE G ACCENT AIGU
 U+0260 ; ģ # LETTRE MINUSCULE LATINE G CROSSE
 U+1E21 ; ģ̄ # LETTRE MINUSCULE LATINE G MACRON
 U+1D77 ; ģ̆ # LETTRE MINUSCULE LATINE G CULBUTÉ
 U+0263 ; γ # LETTRE MINUSCULE LATINE GAMMA
 U+0068 ; h # LETTRE MINUSCULE LATINE H
 U+0125 ; ĥ # LETTRE MINUSCULE LATINE H ACCENT CIRCONFLEXE
 U+0127 ; ħ # LETTRE MINUSCULE LATINE H BARRÉ
 U+021F ; ħ̄ # LETTRE MINUSCULE LATINE H CARON
 U+1E23 ; ħ̆ # LETTRE MINUSCULE LATINE H POINT EN CHEF
 U+1E25 ; ḣ # LETTRE MINUSCULE LATINE H POINT SOUSCRIT
 U+1E27 ; ĥ̄ # LETTRE MINUSCULE LATINE H TRÉMA
 U+1E29 ; ĥ̆ # LETTRE MINUSCULE LATINE H CÉDILLE
 U+1E2B ; ĥ̇ # LETTRE MINUSCULE LATINE H BREVE SOUSCRIT
 U+2C68 ; ḧ # LETTRE MINUSCULE LATINE H CRAMPONNÉ
 U+1E96 ; ĥ̆ # LETTRE MINUSCULE LATINE H LIGNE SOUSCRITE
 U+0266 ; ĥ̇ # LETTRE MINUSCULE LATINE H CROSSE
 U+02AE ; ĥ̆̄ # LETTRE MINUSCULE LATINE H CULBUTÉ CROSSE RÉFLÉCHIE
 U+02AF ; ĥ̆̄ # LETTRE MINUSCULE LATINE H CULBUTÉ CROSSE RÉFLÉCHIE ET HAMEÇON
 U+0069 ; i # LETTRE MINUSCULE LATINE I
 U+00EC ; ì # LETTRE MINUSCULE LATINE I ACCENT GRAVE
 U+00ED ; í # LETTRE MINUSCULE LATINE I ACCENT AIGU
 U+00EE ; î # LETTRE MINUSCULE LATINE I ACCENT CIRCONFLEXE
 U+00EF ; ï # LETTRE MINUSCULE LATINE I TRÉMA
 U+0129 ; ï̄ # LETTRE MINUSCULE LATINE I TILDE
 U+012B ; ī # LETTRE MINUSCULE LATINE I MACRON
 U+012D ; ĩ # LETTRE MINUSCULE LATINE I BRÈVE
 U+012F ; ĭ # LETTRE MINUSCULE LATINE I OGONEK
 U+0131 ; i̇ # LETTRE MINUSCULE LATINE I SANS POINT
 U+01D0 ; ĭ̄ # LETTRE MINUSCULE LATINE I CARON
 U+0209 ; ï̄ # LETTRE MINUSCULE LATINE I DOUBLE ACCENT GRAVE
 U+020B ; î̇ # LETTRE MINUSCULE LATINE I BRÈVE RENVERSÉE
 U+1E2D ; ĭ̇ # LETTRE MINUSCULE LATINE I TILDE SOUSCRIT
 U+1E2F ; ĭ̇̄ # LETTRE MINUSCULE LATINE I TRÉMA ET ACCENT AIGU
 U+1EC9 ; ĭ̇̄ # LETTRE MINUSCULE LATINE I CROCHET EN CHEF
 U+1ECB ; ĭ̇̄ # LETTRE MINUSCULE LATINE I POINT SOUSCRIT
 U+0269 ; ι # LETTRE MINUSCULE LATINE IOTA
 U+006A ; j # LETTRE MINUSCULE LATINE J
 U+0135 ; ĵ # LETTRE MINUSCULE LATINE J ACCENT CIRCONFLEXE
 U+01F0 ; ĵ̄ # LETTRE MINUSCULE LATINE J CARON
 U+0237 ; j̇ # LETTRE MINUSCULE LATINE J SANS POINT
 U+0249 ; ĵ̇ # LETTRE MINUSCULE LATINE J BARRÉ
 U+025F ; J̇ # LETTRE MINUSCULE LATINE J SANS POINT BARRÉ
 U+006B ; k # LETTRE MINUSCULE LATINE K
 U+0137 ; ħ̆ # LETTRE MINUSCULE LATINE K CÉDILLE
 U+0199 ; ħ̆̄ # LETTRE MINUSCULE LATINE K CROSSE
 U+01E9 ; ħ̆̄ # LETTRE MINUSCULE LATINE K CARON
 U+1E31 ; ħ̆̄ # LETTRE MINUSCULE LATINE K ACCENT AIGU
 U+1E33 ; ħ̆̄ # LETTRE MINUSCULE LATINE K POINT SOUSCRIT

U+1E35 ; k̡ # LETTRE MINUSCULE LATINE K LIGNE SOUSCRITE
 U+2C6A ; ḳ # LETTRE MINUSCULE LATINE K CRAMPONNÉ
 U+0138 ; κ # LETTRE MINUSCULE LATINE KRA
 U+006C ; l # LETTRE MINUSCULE LATINE L
 U+013A ; l̇ # LETTRE MINUSCULE LATINE L ACCENT AIGU
 U+013C ; l̈ # LETTRE MINUSCULE LATINE L CÉDILLE
 U+013E ; l̋ # LETTRE MINUSCULE LATINE L CARON
 U+0142 ; ľ # LETTRE MINUSCULE LATINE L BARRÉ

 U+019A ; ł # LETTRE MINUSCULE LATINE L RAYÉ
 U+0234 ; l̥ # LETTRE MINUSCULE LATINE L BOUCLÉ
 U+1E37 ; l̦ # LETTRE MINUSCULE LATINE L POINT SOUSCRIT
 U+1E39 ; ļ # LETTRE MINUSCULE LATINE L POINT SOUSCRIT ET MACRON
 U+1E3B ; l̨ # LETTRE MINUSCULE LATINE L LIGNE SOUSCRITE
 U+1E3D ; l̩ # LETTRE MINUSCULE LATINE L ACCENT CIRCONFLEXE SOUSCRIT
 U+2C61 ; l̪ # LETTRE MINUSCULE LATINE L À DOUBLE BARRE
 U+006D ; m # LETTRE MINUSCULE LATINE M
 U+1E3F ; ṁ # LETTRE MINUSCULE LATINE M ACCENT AIGU
 U+1E41 ; m̈ # LETTRE MINUSCULE LATINE M POINT EN CHEF
 U+1E43 ; m̋ # LETTRE MINUSCULE LATINE M POINT SOUSCRIT
 U+006E ; n # LETTRE MINUSCULE LATINE N
 U+00F1 ; ñ # LETTRE MINUSCULE LATINE N TILDE
 U+0144 ; ṅ # LETTRE MINUSCULE LATINE N ACCENT AIGU
 U+0146 ; n̈ # LETTRE MINUSCULE LATINE N CÉDILLE
 U+0148 ; n̋ # LETTRE MINUSCULE LATINE N CARON

 U+019E ; ň # LETTRE MINUSCULE LATINE N À LONG JAMBAGE DE DROITE
 U+01F9 ; ñ̇ # LETTRE MINUSCULE LATINE N ACCENT GRAVE
 U+0235 ; n̥ # LETTRE MINUSCULE LATINE N BOUCLÉ
 U+1E45 ; n̦ # LETTRE MINUSCULE LATINE N POINT EN CHEF
 U+1E47 ; ņ # LETTRE MINUSCULE LATINE N POINT SOUSCRIT
 U+1E49 ; n̨ # LETTRE MINUSCULE LATINE N LIGNE SOUSCRITE
 U+1E4B ; n̩ # LETTRE MINUSCULE LATINE N ACCENT CIRCONFLEXE SOUSCRIT
 U+0272 ; n̪ # LETTRE MINUSCULE LATINE N HAMEÇON À GAUCHE
 U+014B ; ŋ # LETTRE MINUSCULE LATINE ENG
 U+006F ; o # LETTRE MINUSCULE LATINE O
 U+00F2 ; ò # LETTRE MINUSCULE LATINE O ACCENT GRAVE
 U+00F3 ; ó # LETTRE MINUSCULE LATINE O ACCENT AIGU
 U+00F4 ; ô # LETTRE MINUSCULE LATINE O ACCENT CIRCONFLEXE
 U+00F5 ; õ # LETTRE MINUSCULE LATINE O TILDE
 U+00F6 ; ö # LETTRE MINUSCULE LATINE O TRÉMA
 U+00F8 ; ø # LETTRE MINUSCULE LATINE O BARRÉ
 U+014D ; õ̇ # LETTRE MINUSCULE LATINE O MACRON
 U+014F ; ṏ # LETTRE MINUSCULE LATINE O BRÈVE
 U+0151 ; õ̋ # LETTRE MINUSCULE LATINE O DOUBLE ACCENT AIGU
 U+01A1 ; ø̇ # LETTRE MINUSCULE LATINE O CORNU
 U+01D2 ; ṏ # LETTRE MINUSCULE LATINE O CARON
 U+01EB ; q̇ # LETTRE MINUSCULE LATINE O OGONEK

 U+01ED ; q̈ # LETTRE MINUSCULE LATINE O OGONEK ET MACRON
 U+01FF ; ó̇ # LETTRE MINUSCULE LATINE O BARRÉ ET ACCENT AIGU
 U+020D ; ò̇ # LETTRE MINUSCULE LATINE O DOUBLE ACCENT GRAVE
 U+020F ; ô̇ # LETTRE MINUSCULE LATINE O BRÈVE RENVERSÉE
 U+022B ; ö̇ # LETTRE MINUSCULE LATINE O TRÉMA ET MACRON

U+022D ; ð # LETTRE MINUSCULE LATINE O TILDE ET MACRON
 U+022F ; ô # LETTRE MINUSCULE LATINE O POINT EN CHEF
 U+0231 ; õ # LETTRE MINUSCULE LATINE O POINT EN CHEF ET MACRON
 U+1E4D ; ö # LETTRE MINUSCULE LATINE O TILDE ET ACCENT AIGU
 U+1E4F ; õ # LETTRE MINUSCULE LATINE O TILDE ET TRÉMA
 U+1E51 ; ò # LETTRE MINUSCULE LATINE O MACRON ET ACCENT GRAVE
 U+1E53 ; ó # LETTRE MINUSCULE LATINE O MACRON ET ACCENT AIGU
 U+1ECD ; ȝ # LETTRE MINUSCULE LATINE O POINT SOUSCRIT
 U+1ECF ; ȝ # LETTRE MINUSCULE LATINE O CROCHET EN CHEF
 U+1ED1 ; ô # LETTRE MINUSCULE LATINE O ACCENT CIRCONFLEXE ET AIGU
 U+1ED3 ; ò # LETTRE MINUSCULE LATINE O ACCENT CIRCONFLEXE ET GRAVE
 U+1ED5 ; õ # LETTRE MINUSCULE LATINE O ACCENT CIRCONFLEXE ET CROCHET EN CHEF
 U+1ED7 ; ð # LETTRE MINUSCULE LATINE O ACCENT CIRCONFLEXE ET TILDE
 U+1ED9 ; ô # LETTRE MINUSCULE LATINE O ACCENT CIRCONFLEXE ET POINT SOUSCRIT
 U+1EDB ; ó # LETTRE MINUSCULE LATINE O CORNU ET ACCENT AIGU
 U+1EDD ; ò # LETTRE MINUSCULE LATINE O CORNU ET ACCENT GRAVE
 U+1EDF ; ô # LETTRE MINUSCULE LATINE O CORNU ET CROCHET EN CHEF
 U+1EE1 ; õ # LETTRE MINUSCULE LATINE O CORNU ET TILDE
 U+1EE3 ; ȝ # LETTRE MINUSCULE LATINE O CORNU ET POINT SOUSCRIT
 U+0254 ; ɔ # LETTRE MINUSCULE LATINE O OUVERT
 U+0275 ; ø # LETTRE MINUSCULE LATINE O BARRÉ
 U+0153 ; œ # DIAGRAMME SOUDÉ MINUSCULE LATIN OE
 U+0070 ; p # LETTRE MINUSCULE LATINE P
 U+01A5 ; ꝑ # LETTRE MINUSCULE LATINE P CROSSE
 U+1E55 ; ꝑ # LETTRE MINUSCULE LATINE P ACCENT AIGU
 U+1E57 ; ꝑ # LETTRE MINUSCULE LATINE P POINT EN CHEF
 U+0071 ; q # LETTRE MINUSCULE LATINE Q
 U+024B ; q # LETTRE MINUSCULE LATINE Q AVEC HAMEÇON
 U+0072 ; r # LETTRE MINUSCULE LATINE R
 U+0155 ; ř # LETTRE MINUSCULE LATINE R ACCENT AIGU
 U+0157 ; ɽ # LETTRE MINUSCULE LATINE R CÉDILLE
 U+0159 ; ɽ # LETTRE MINUSCULE LATINE R CARON
 U+0211 ; ɽ # LETTRE MINUSCULE LATINE R DOUBLE ACCENT GRAVE
 U+0213 ; ɽ # LETTRE MINUSCULE LATINE R BREVE RENVERSÉE
 U+1E59 ; ɽ # LETTRE MINUSCULE LATINE R POINT EN CHEF
 U+1E5B ; ɽ # LETTRE MINUSCULE LATINE R POINT SOUSCRIT
 U+1E5D ; ɽ # LETTRE MINUSCULE LATINE R POINT SOUSCRIT ET MACRON
 U+1E5F ; ɽ # LETTRE MINUSCULE LATINE R LIGNE SOUSCRITE
 U+024D ; ɽ # LETTRE MINUSCULE LATINE R BARRÉ
 U+027D ; ɽ # LETTRE MINUSCULE LATINE R HAMEÇON RÉTROFLEXE
 U+027F ; ɽ # LETTRE MINUSCULE LATINE R SANS ÔBIT RÉFLÉCHI
 U+0073 ; s # LETTRE MINUSCULE LATINE S
 U+015B ; ś # LETTRE MINUSCULE LATINE S ACCENT AIGU
 U+015D ; ŝ # LETTRE MINUSCULE LATINE S ACCENT CIRCONFLEXE
 U+015F ; ś # LETTRE MINUSCULE LATINE S CÉDILLE
 U+0161 ; š # LETTRE MINUSCULE LATINE S CARON
 U+0219 ; ŝ # LETTRE MINUSCULE LATINE S VIRGULE SOUSCRITE
 U+1E61 ; Ś # LETTRE MINUSCULE LATINE S POINT EN CHEF
 U+1E63 ; ŝ # LETTRE MINUSCULE LATINE S POINT SOUSCRIT
 U+1E65 ; ś # LETTRE MINUSCULE LATINE S ACCENT AIGU ET POINT EN CHEF
 U+1E67 ; š # LETTRE MINUSCULE LATINE S CARON ET POINT EN CHEF
 U+1E69 ; ŝ # LETTRE MINUSCULE LATINE S POINT SOUSCRIT ET POINT EN CHEF

U+023F ; § # LETTRE MINUSCULE LATINE S À PARAPHE
 U+00DF ; ß # LETTRE MINUSCULE LATINE S DUR
 U+1E9C ; □ # LETTRE MINUSCULE LATINE S LONG BARRÉ OBLIQUEMENT
 U+1E9D ; □ # LETTRE MINUSCULE LATINE S LONG BARRE HAUTE
 U+0283 ; ſ # LETTRE MINUSCULE LATINE ECH
 U+0074 ; t # LETTRE MINUSCULE LATINE T
 U+0163 ; ʈ # LETTRE MINUSCULE LATINE T CÉDILLE
 U+0165 ; ʈ̃ # LETTRE MINUSCULE LATINE T CARON
 U+0167 ; ʈ̄ # LETTRE MINUSCULE LATINE T BARRÉ
 U+01AB ; ʈ̥ # LETTRE MINUSCULE LATINE T HAMEÇON PALATAL
 U+01AD ; ʈ̦ # LETTRE MINUSCULE LATINE T CROSSE
 U+1E6B ; ʈ̧ # LETTRE MINUSCULE LATINE S POINT EN CHEF
 U+1E6D ; ʈ̨ # LETTRE MINUSCULE LATINE T POINT SOUSCRIT
 U+1E6F ; ʈ̩ # LETTRE MINUSCULE LATINE T LIGNE SOUSCRITE
 U+1E71 ; ʈ̪ # LETTRE MINUSCULE LATINE T ACCENT CIRCONFLEXE SOUSCRIT
 U+0236 ; ʈ̫ # LETTRE MINUSCULE LATINE T BOUCLÉ
 U+021B ; ʈ̬ # LETTRE MINUSCULE LATINE T VIRGULE SOUSCRITE
 U+1E97 ; ʈ̭ # LETTRE MINUSCULE LATINE T TRÉMA
 U+2C66 ; ʈ̮ # LETTRE MINUSCULE LATINE T BARRÉ OBLIQUEMENT
 U+0288 ; ʈ̯ # LETTRE MINUSCULE LATINE T HAMEÇON RÉTROFLEXE
 U+00FE ; þ # LETTRE MINUSCULE LATINE THORN
 U+0075 ; u # LETTRE MINUSCULE LATINE U
 U+00F9 ; ù # LETTRE MINUSCULE LATINE U ACCENT GRAVE
 U+00FA ; ú # LETTRE MINUSCULE LATINE U ACCENT AIGU
 U+00FB ; û # LETTRE MINUSCULE LATINE U ACCENT CIRCONFLEXE
 U+00FC ; ü # LETTRE MINUSCULE LATINE U TRÉMA
 U+0169 ; ũ # LETTRE MINUSCULE LATINE U TILDE
 U+016B ; ū # LETTRE MINUSCULE LATINE U MACRON
 U+016D ; ŭ # LETTRE MINUSCULE LATINE U BRÈVE
 U+016F ; ŭ̂ # LETTRE MINUSCULE LATINE U ROND EN CHEF
 U+0171 ; ú̂ # LETTRE MINUSCULE LATINE U DOUBLE ACCENT AIGU
 U+0173 ; ȳ # LETTRE MINUSCULE LATINE U OGONEK
 U+01B0 ; ȳ̄ # LETTRE MINUSCULE LATINE U CORNU
 U+01D4 ; ů # LETTRE MINUSCULE LATINE U CARON
 U+01D6 ; ů̄ # LETTRE MINUSCULE LATINE U TRÉMA ET MACRON
 U+01D8 ; ů̂ # LETTRE MINUSCULE LATINE U TRÉMA ET ACCENT AIGU
 U+01DA ; ů̃ # LETTRE MINUSCULE LATINE U TRÉMA ET CARON
 U+01DC ; ů̄ # LETTRE MINUSCULE LATINE U TRÉMA ET ACCENT GRAVE
 U+0215 ; ů̆ # LETTRE MINUSCULE LATINE U DOUBLE ACCENT GRAVE
 U+0217 ; ů̇ # LETTRE MINUSCULE LATINE U BRÈVE RENVERSÉE
 U+1E73 ; ȳ̈ # LETTRE MINUSCULE LATINE U TRÉMA SOUSCRIT
 U+1E75 ; ȳ̉ # LETTRE MINUSCULE LATINE U TILDE SOUSCRIT
 U+1E77 ; ȳ̊ # LETTRE MINUSCULE LATINE U ACCENT CIRCONFLEXE SOUSCRIT
 U+1E79 ; ů̋ # LETTRE MINUSCULE LATINE U TILDE ET ACCENT AIGU
 U+1E7B ; ů̌ # LETTRE MINUSCULE LATINE U MACRON ET TRÉMA
 U+1EE5 ; ȳ̍ # LETTRE MINUSCULE LATINE U POINT SOUSCRIT
 U+1EE7 ; ů̎ # LETTRE MINUSCULE LATINE U CROCHET EN CHEF
 U+1EE9 ; ů̏ # LETTRE MINUSCULE LATINE U CORNU ET ACCENT AIGU
 U+1EEB ; ů̐ # LETTRE MINUSCULE LATINE U CORNU ET ACCENT GRAVE
 U+1EED ; ů̑ # LETTRE MINUSCULE LATINE U CORNU ET CROCHET EN CHEF
 U+1EEF ; ů̒ # LETTRE MINUSCULE LATINE U CORNU ET TILDE
 U+1EF1 ; ȳ̓ # LETTRE MINUSCULE LATINE U CORNU ET POINT SOUSCRIT

U+0076 ; v # LETTRE MINUSCULE LATINE V
 U+1E7D ; ṽ # LETTRE MINUSCULE LATINE V TILDE
 U+1E7F ; ṿ # LETTRE MINUSCULE LATINE V POINT SOUSCRIT
 U+1EFD ; □ # LETTRE MINUSCULE LATINE V MIDDLE-WELSH
 U+2C71 ; v̂ # LETTRE MINUSCULE LATINE V HAMEÇON A DROITE
 U+2C74 ; v̄ # LETTRE MINUSCULE LATINE V BOUCLÉ
 U+028B ; v̥ # LETTRE MINUSCULE LATINE V CROSSE
 U+028C ; v̧ # LETTRE MINUSCULE LATINE V CULBUTÉ
 U+0077 ; w # LETTRE MINUSCULE LATINE W
 U+0175 ; ŵ # LETTRE MINUSCULE LATINE W ACCENT CIRCONFLEXE
 U+1E81 ; ẉ # LETTRE MINUSCULE LATINE W ACCENT GRAVE
 U+1E83 ; w̤ # LETTRE MINUSCULE LATINE W ACCENT AIGU
 U+1E85 ; w̥ # LETTRE MINUSCULE LATINE W TRÉMA
 U+1E87 ; w̧ # LETTRE MINUSCULE LATINE W POINT EN CHEF
 U+1E89 ; w̨ # LETTRE MINUSCULE LATINE W POINT SOUSCRIT
 U+1E98 ; w̩ # LETTRE MINUSCULE LATINE W ROND EN CHEF
 U+2C73 ; w̪ # LETTRE MINUSCULE LATINE W CROSSE

 U+01BF ; p̥ # LETTRE LATINE WYNN
 U+0078 ; x # LETTRE MINUSCULE LATINE X
 U+1E8B ; x̣ # LETTRE MINUSCULE LATINE S POINT EN CHEF
 U+1E8D ; x̤ # LETTRE MINUSCULE LATINE X TRÉMA
 U+0079 ; y # LETTRE MINUSCULE LATINE Y
 U+00FD ; ý # LETTRE MINUSCULE LATINE Y ACCENT AIGU
 U+00FF ; ÿ # LETTRE MINUSCULE LATINE Y TRÉMA
 U+0177 ; ŷ # LETTRE MINUSCULE LATINE Y ACCENT CIRCONFLEXE

 U+01B4 ; ỵ̂ # LETTRE MINUSCULE LATINE Y CROSSE
 U+0233 ; ŷ̤ # LETTRE MINUSCULE LATINE Y MACRON
 U+024F ; ŷ̥ # LETTRE MINUSCULE LATINE Y BARRÉ
 U+1E99 ; ŷ̧ # LETTRE MINUSCULE LATINE Y ROND EN CHEF
 U+1EF3 ; ŷ̨ # LETTRE MINUSCULE LATINE Y ACCENT GRAVE
 U+1EF5 ; ŷ̩ # LETTRE MINUSCULE LATINE Y POINT SOUSCRIT
 U+1EF7 ; ŷ̪ # LETTRE MINUSCULE LATINE Y CROCHET EN CHEF
 U+1EF9 ; ŷ̫ # LETTRE MINUSCULE LATINE Y TILDE
 U+1EFF ; □ # LETTRE MINUSCULE LATINE Y BOUCLÉ
 U+1E8F ; ŷ̬ # LETTRE MINUSCULE LATINE Y POINT EN CHEF
 U+021D ; Ʒ # LETTRE MINUSCULE LATINE YOGH
 U+007A ; z # LETTRE MINUSCULE LATINE Z
 U+017A ; ẑ # LETTRE MINUSCULE LATINE Z ACCENT AIGU
 U+017C ; ẓ # LETTRE MINUSCULE LATINE Z POINT EN CHEF
 U+017E ; z̤ # LETTRE MINUSCULE LATINE Z CARON

 U+01B6 ; z̥ # LETTRE MINUSCULE LATINE Z BARRÉ
 U+0225 ; z̧ # LETTRE MINUSCULE LATINE Z CROSSE
 U+0240 ; z̨ # LETTRE MINUSCULE LATINE Z À PARAPHE
 U+1E91 ; z̩ # LETTRE MINUSCULE LATINE Z ACCENT CIRCONFLEXE
 U+1E93 ; z̪ # LETTRE MINUSCULE LATINE Z POINT SOUSCRIT
 U+1E95 ; z̫ # LETTRE MINUSCULE LATINE Z LIGNE SOUSCRITE
 U+2C6C ; z̬ # LETTRE MINUSCULE LATINE Z CRAMPONNÉ
 U+0292 ; Ʒ̣ # LETTRE MINUSCULE LATINE EJ
 U+01B9 ; Ʒ̤ # LETTRE MINUSCULE LATINE EJ RÉFLÉCHI
 U+01BA ; Ʒ̥ # LETTRE MINUSCULE LATINE EJ HAMEÇON
 U+01EF ; Ʒ̧ # LETTRE MINUSCULE LATINE EJ CARON

U+01C0	; # LETTRE LATINE CLIC DENTAL
U+01C1	; # LETTRE LATINE CLIC LATÉRAL
U+01C2	; † # LETTRE LATINE CLIC ALVÉOLAIRE
U+01C3	; ! # LETTRE LATINE CLIC RÉTROFLEXE
U+0294	; ? # LETTRE LATINE COUP DE GLOTTE

B3. Points de code de réserve

Les points de code suivants sont également disponibles dans les IDN en écriture latine, mais le groupe d'étude ne s'attend à les voir apparaître que dans des circonstances spéciales qui ne comprennent pas leur utilisation dans la zone racine du DNS et les a donc placés dans une table différente. Ces points de code sont répertoriés par ordre numérique dans leurs en-têtes ou leurs sous en-têtes de bloc Unicode

EXTENSIONS IPA

U+0250	; v # LETTRE MINUSCULE LATINE a CULBUTÉE
U+0252	; v # LETTRE MINUSCULE LATINE ALFA CULBUTÉE
U+0258	; e # LETTRE MINUSCULE LATINE e RÉFLÉCHIE
U+025A	; œ # LETTRE MINUSCULE LATINE SCHWA CROCHET
U+025C	; ɜ # LETTRE MINUSCULE LATINE EPSILON REFLECHI OUVERTE
U+025D	; ɝ # LETTRE MINUSCULE LATINE EPSILON REFLECHI CROCHET
U+025E	; e̞ # LETTRE MINUSCULE LATINE EPSILON REFLECHI FERMÉE
U+0264	; ɣ # LETTRE MINUSCULE LATINE PETIT GAMMA
U+0265	; ɥ # LETTRE MINUSCULE LATINE HACHE LFA CULBUTÉE
U+0267	; ɧ # LETTRE MINUSCULE LATINE ENG CROSSE
U+0268	; i̇ # LETTRE MINUSCULE LATINE I BARRÉE
U+026A	; I # LETTRE LATINE PETITE CAPITALE I
U+026B	; l̥ # LETTRE MINUSCULE LATINE L TILDE MEDIAN
U+026C	; l̊ # LETTRE MINUSCULE LATINE L SANGLÉE
U+026D	; l̝ # LETTRE MINUSCULE LATINE L HAMEÇON RÉTROFLEXE
U+026E	; l̞ # LETTRE MINUSCULE LATINE LEJ
U+026F	; w̥ # LETTRE MINUSCULE LATINE M CULBUTÉE
U+0270	; ẘ # LETTRE MINUSCULE LATINE M HAMPÉ CULBUTÉ
U+0271	; m̝ # LETTRE MINUSCULE LATINE M HAMEÇON
U+0273	; n̝ # LETTRE MINUSCULE LATINE N HAMEÇON RÉTROFLEXE
U+0274	; N # LETTRE LATINE PETITE CAPITALE
U+0276	; œ̥ # LETTRE LATINE PETITE CAPITALE OE
U+0277	; ω̥ # LETTRE MINUSCULE LATINE OMÉGA FERMÉ
U+0278	; φ̥ # LETTRE MINUSCULE LATINE PHI
U+0279	; r̥ # LETTRE MINUSCULE LATINE R CULBUTÉ
U+027A	; r̊ # LETTRE MINUSCULE LATINE R CROSSE CULBUTÉ
U+027B	; r̝ # LETTRE MINUSCULE LATINE R CROSSE CULBUTÉ
U+027C	; r̞ # LETTRE MINUSCULE LATINE R PROLONGÉ
U+027E	; r̥ # LETTRE MINUSCULE LATINE R SANS OBIT
U+0280	; R # LETTRE LATINE PETITE CAPITALE
U+0281	; R̥ # LETTRE LATINE PETITE CAPITALE R RENVERSÉ

U+0282 ; ſ LETTRE MINUSCULE LATINE S HAMEÇON RÉTROFLEXE
 U+0284 ; ƒ# LETTRE MINUSCULE LATINE J SANS POINT BARRÉ CROSSE
 U+0285 ; ʀ# LETTRE MINUSCULE LATINE R SANS OBIT RÉFLÉCHI HAMEÇON
 RÉTROFLEXE
 U+0286 ; ʃ LETTRE MINUSCULE LATINE ECH BOUCLÉ
 U+0287 ; ʀ# LETTRE MINUSCULE LATINE T CULBUTÉ
 U+0289 ; ʁ# LETTRE MINUSCULE LATINE U BARRÉ
 U+028A ; ʊ# LETTRE MINUSCULE LATINE UPSILON
 U+028D ; ʌ LETTRE MINUSCULE LATINE W CULBUTÉE
 U+028E ; ʎ# LETTRE MINUSCULE LATINE Y CULBUTÉ
 U+028F ; ʏ# LETTRE LATINE PETITE CAPITALE Y
 U+0290 ; ʐ# LETTRE MINUSCULE LATINE Z HAMEÇON RÉTROFLEXE
 U+0291 ; ʑ# LETTRE MINUSCULE LATINE Z BOUCLÉ
 U+0293 ; ʒ# LETTRE MINUSCULE LATINE EJ BOUCLÉ
 U+0295 ; ʔ# LETTRE LATINE COUP DE GLOTTE RÉFLÉCHI
 U+0296 ; ʕ# LETTRE LATINE COUP DE GLOTTE CULBUTÉ
 U+0297 ; ʞ# LETTRE LATINE C ÉTIRÉ
 U+0298 ; ʟ # LETTRE LATINE CLIC BILABILA
 U+0299 ; ʙ# LETTRE LATINE PETITE CAPITALE B
 U+029A ; ɘ# LETTRE MINUSCULE LATINE EPSILON FERMÉE OUVERTE
 U+029B ; ɠ# LETTRE MINUSCULE LATINE PETITE CAPITALE G CROSSE
 U+029C ; ɨ# LETTRE LATINE PETITE CAPITALE H
 U+029D ; ʝ# LETTRE MINUSCULE LATINE J QUEUE CROISÉE
 U+029E ; ɣ# LETTRE MINUSCULE LATINE K CULBUTÉ
 U+029F ; ʌ# LETTRE LATINE PETITE CAPITALE L
 U+02A0 ; ɔ# LETTRE MINUSCULE LATINE Q CROSSE
 U+02A1 ; ʔ# LETTRE LATINE COUP DE GLOTTE BARRÉ
 U+02A2 ; ʕ# LETTRE LATINE COUP DE GLOTTE BARRÉ RÉFLÉCHI
 U+02A3 ; ɀ# LETTRE MINUSCULE LATINE DIGRAMME DZ
 U+02A4 ; Ɂ# LETTRE MINUSCULE LATINE DIGRAMME DEZ
 U+02A5 ; ɂ# LETTRE MINUSCULE LATINE DIGRAMME DZ BOUCLÉ
 U+02A6 ; Ƀ# LETTRE MINUSCULE LATINE DIGRAMME TS
 U+02A7 ; Ʉ# LETTRE MINUSCULE LATINE DIGRAMME TECH
 U+02A8 ; Ʌ# LETTRE MINUSCULE LATINE DIGRAMME TC BOUCLÉ
 U+02A9 ; Ɇ# LETTRE MINUSCULE LATINE DIGRAMME FENG
 U+02AA ; ɇ# LETTRE MINUSCULE LATINE DIGRAMME LS
 U+02AB ; Ɉ# LETTRE MINUSCULE LATINE DIGRAMME LZ
 U+02AC ; ɉ# LETTRE LATINE PERCUSSION BILABIALE
 U+02AD ; Ɋ# LETTRE LATINE PERCUSSION BIDENTALE

EXTENSIONS PHONÉTIQUES

U+1D6B ; ɥ # LETTRE MINUSCULE LATINE UE
 U+1D6C ; ɇ# LETTRE MINUSCULE LATINE B TILDE MEDIAN
 U+1D6D ; Ɉ# LETTRE MINUSCULE LATINE D TILDE MÉDIAN
 U+1D6E ; ɉ# LETTRE MINUSCULE LATINE f TILDE MÉDIAN
 U+1D6F ; Ɋ# LETTRE MINUSCULE LATINE M TILDE MÉDIAN
 U+1D70 ; ɋ# LETTRE MINUSCULE LATINE N TILDE MEDIAN
 U+1D71 ; Ɍ# LETTRE MINUSCULE LATINE P TILDE MEDIAN

U+1D72	; ɣ # LETTRE MINUSCULE LATINE R TILDE MÉDIAN
U+1D73	; ɣ̄ # LETTRE MINUSCULE LATINE R SANS OBIT TILDE MÉDIAN
U+1D74	; s̄ # LETTRE MINUSCULE LATINE S TILDE MEDIAN
U+1D75	; t̄ # LETTRE MINUSCULE LATINE T TILDE MEDIAN
U+1D76	; z̄ # LETTRE MINUSCULE LATINE Z TILDE MEDIAN
U+1D79	; ȝ # LETTRE MINUSCULE LATINE G INSULAIRE
U+1D7A	; th̄ LETTRE MINUSCULE LATINE TH BARRÉ DIAGONALEMENT
U+1D7B	; ī LETTRE LATINE PETITE CAPITALE I BARRÉE
U+1D7C	; i̇ LETTRE MINUSCULE LATINE IOTA BARRÉ
U+1D7D	; p̄ LETTRE MINUSCULE LATINE P BARRÉ
U+1D7E	; ū LETTRE LATINE PETITE CAPITALE U BARRÉE
U+1D7F	; u̇ LETTRE MINUSCULE LATINE UPSILON BARRÉ
U+1D80	; b̄ LETTRE MINUSCULE LATINE B HAMEÇON PALATAL
U+1D81	; d̄ LETTRE MINUSCULE LATINE D HAMEÇON PALATAL
U+1D82	; f̄ LETTRE MINUSCULE LATINE F HAMEÇON PALATAL
U+1D83	; ḡ LETTRE MINUSCULE LATINE G HAMEÇON PALATAL
U+1D84	; k̄ LETTRE MINUSCULE LATINE K HAMEÇON PALATAL
U+1D85	; l̄ LETTRE MINUSCULE LATINE L HAMEÇON PALATAL
U+1D86	; m̄ LETTRE MINUSCULE LATINE M HAMEÇON PALATAL
U+1D87	; n̄ LETTRE MINUSCULE LATINE N HAMEÇON PALATAL
U+1D88	; p̄ LETTRE MINUSCULE LATINE P HAMEÇON PALATAL
U+1D89	; r̄ LETTRE MINUSCULE LATINE R HAMEÇON PALATAL
U+1D8A	; s̄ LETTRE MINUSCULE LATINE S HAMEÇON PALATAL
U+1D8B	; ʃ̄ LETTRE MINUSCULE LATINE ECH HAMEÇON PALATAL
U+1D8C	; v̄ LETTRE MINUSCULE LATINE V HAMEÇON PALATAL
U+1D8D	; x̄ LETTRE MINUSCULE LATINE X HAMEÇON PALATAL
U+1D8E	; z̄ LETTRE MINUSCULE LATINE Z HAMEÇON PALATAL
U+1D8F	; ā LETTRE MINUSCULE LATINE A HAMEÇON RÉTROFLEXE
U+1D90	; ᾱ LETTRE MINUSCULE LATINE ALPHA HAMEÇON RÉTROFLEXE
U+1D91	; c̄ # LETTRE MINUSCULE LATINE CROSSE ET HAMEÇON RÉTROFLEXE
U+1D92	; ē LETTRE MINUSCULE LATINE E HAMEÇON RÉTROFLEXE
U+1D93	; ε̄ LETTRE MINUSCULE LATINE EPSILON HAMEÇON RÉTROFLEXE
U+1D94	; ɜ̄ LETTRE MINUSCULE LATINE EPSILON RÉFLÉCHI HAMEÇON RÉTROFLEXE
U+1D95	; ə̄ LETTRE MINUSCULE LATINE SCHWA HAMEÇON RÉTROFLEXE
U+1D96	; ī LETTRE MINUSCULE LATINE I HAMEÇON RÉTROFLEXE
U+1D97	; ō LETTRE MINUSCULE LATINE O OUVERT HAMEÇON RÉTROFLEXE
U+1D98	; ʃ̄ LETTRE MINUSCULE LATINE ECH HAMEÇON RÉTROFLEXE
U+1D99	; ū LETTRE MINUSCULE LATINE U HAMEÇON RÉTROFLEXE
U+1D9A	; ɹ̄ LETTRE MINUSCULE LATINE ECH HAMEÇON RÉTROFLEXE

LATIN ÉTENDU C

Lettres claudiennes

U+2C76 ; h̄ # LETTRE MINUSCULE LATINE MOITIÉ DE H

Ajouts pour UPA

U+2C77 ; ø̄ # LETTRE MINUSCULE LATINE PHI SANS QUEUE

U+2C78	; □	# LETTRE MINUSCULE LATINE E ENCOCHE
U+2C79	; □	# LETTRE MINUSCULE LATINE R HAMEÇON RÉTROFLEXE
U+2C7A	; □	# LETTRE LATINE MINUSCULE O ANNEAU INTERNE
U+2C7B	; □	# LETTRE LATINE PETITE CAPITALE CULBUTÉE E

LATIN ÉTENDU D

Ajouts pour lettres égyptologues

U+A723	; □	# LETTRE MINUSCULE LATINE ALEPH EGYPTOLOGIQUE
U+A725	; □	# LETTRE MINUSCULE LATINE AYIN EGYPTOLOGIQUE

Ajouts pour lettres mayanistes

U+A727	; □	# LETTRE MINUSCULE LATINE ENG
U+A729	; □	# LETTRE MINUSCULE LATINE TEZ
U+A72B	; □	# LETTRE MINUSCULE LATINE ED
U+A72D	; □	# LETTRE MINUSCULE LATINE CUATRILLO
U+A72F	; □	# LETTRE MINUSCULE LATINE CUATRILLO VIRGULE

Ajouts pour lettres médiévalistes

U+A730	; □	# LETTRE LATINE PETITE CAPITALE F
U+A731	; □	# LETTRE LATINE PETITE CAPITALE S
U+A733	; □	# LETTRE MINUSCULE LATINE AA
U+A735	; □	# LETTRE MINUSCULE LATINE AO
U+A737	; □	# LETTRE MINUSCULE LATINE AU
U+A739	; □	# LETTRE MINUSCULE LATINE AV
U+A73B	; □	# LETTRE MINUSCULE LATINE AV BARRÉ HORIZONTAL
U+A73D	; □	# LETTRE MINUSCULE LATINE AY
U+A73F	; □	# LETTRE MINUSCULE LATINE RÉFLÉCHI C POINT
U+A741	; □	# LETTRE MINUSCULE LATINE K BARRÉ
U+A743	; □	# LETTRE MINUSCULE LATINE K BARRÉ DIAGONALEMENT
U+A745	; □	# LETTRE MINUSCULE LATINE K BARRÉ ET DIAGONALE
U+A747	; □	# LETTRE LATINE PETITE CAPITALE L BROKEN
U+A749	; □	# LETTRE MINUSCULE LATINE L BARRÉE HAUTE
U+A74B	; □	# LETTRE MINUSCULE LATINE O BARRÉ SUPERPOSÉ
U+A74D	; □	# LETTRE MINUSCULE LATINE O BOUCLÉ
U+A74F	; □	# LETTRE MINUSCULE LATINE OO
U+A751	; □	# LETTRE MINUSCULE LATINE P BARRÉ TRAVERS DES
U+A753	; □	# LETTRE MINUSCULE LATINE P BARRÉ PARAFE
U+A755	; □	# LETTRE MINUSCULE LATINE P QUEUE D'ÉCUREUIL
U+A757	; □	# LETTRE MINUSCULE LATINE P BARRÉ TRAVERSANT DES
U+A759	; □	# LETTRE MINUSCULE LATINE Q BARRÉ TRAVERSANT DIAGONALEMENT
U+A75B	; □	# LETTRE MINUSCULE LATINE R ROTUNDA
U+A75D	; □	# LETTRE MINUSCULE LATINE RUM ROTUNDA
U+A75F	; □	# LETTRE MINUSCULE LATINE V BARRÉ DIAGONALEMENT
U+A761	; □	# LETTRE MINUSCULE LATINE VY

U+A763	; □	# LETTRE LATINE MINUSCULE Z WISIGOTHIQUE
U+A765	; □	# LETTRE MINUSCULE LATINE THORN BARRÉ
U+A767	; □	# LETTRE MINUSCULE LATINE THORN BARRÉ DIAGONALE
U+A769	; □	# LETTRE MINUSCULE LATINE VEND
U+A76B	; □	# LETTRE MINUSCULE LATINE ET
U+A76D	; □	# LETTRE MINUSCULE LATINE IS
U+A76F	; □	# LETTRE MINUSCULE LATINE CON
U+A771	; □	# LETTRE MINUSCULE LATINE DUM
U+A772	; □	# LETTRE MINUSCULE LATINE LUM
U+A773	; □	# LETTRE MINUSCULE LATINEMUM
U+A774	; □	# LETTRE MINUSCULE LATINE NUM
U+A775	; □	# LETTRE MINUSCULE LATINE RUM
U+A776	; □	# LETTRE LATINE PETITE CAPITALE RUM
U+A777	; □	# LETTRE MINUSCULE LATINE TUM
U+A778	; □	# LETTRE MINUSCULE LATINE UM

Lettres pour insulaires et celticistes

U+A77A	; □	# LETTRE MINUSCULE LATINE INSULAIRE D
U+A77C	; □	# LETTRE MINUSCULE LATINE INSULAIRE f
U+A77F	; □	# LETTRE MINUSCULE LATINE INSULAIRE G
U+A781	; □	# LETTRE MINUSCULE LATINE CULBUTÉE L
U+A783	; □	# LETTRE MINUSCULE LATINE INSULAIRE G
U+A785	; □	# LETTRE MINUSCULE LATINE INSULAIRE R
U+A787	; □	# LETTRE MINUSCULE LATINE INSULAIRE T

Lettres orthographiques pour glottales

U+A78C	; ' □	# LETTRE MINUSCULE LATINE SALTILLO
--------	-------	------------------------------------

Lettres épigraphiques du romain ancien

U+A7FB	; □	# LETTRE ÉPIGRAPHIQUE LATINE RÉFLÉCHI F
U+A7FC	; □	# LETTRE ÉPIGRAPHIQUE LATINE RÉFLÉCHI P
U+A7FD	; □	# LETTRE ÉPIGRAPHIQUE LATINE RENVERSÉ M
U+A7FE	; □	# LETTRE ÉPIGRAPHIQUE LATINE I LONGA
U+A7FF	; □	# LETTRE ÉPIGRAPHIQUE LATINE ARCHAÏQUE M

Symboles phonétiques

U+A78E	; □	# LETTRE MINUSCULE LATINE L HAMEÇON RÉTROFLEXE SANGLÉE
--------	-----	--

Lettres yanalf

U+A791	; □	# LETTRE MINUSCULE LATINE N CRAMPONNÉ
--------	-----	---------------------------------------

Lettres latviennes pour l'orthographe d'avant 1921

U+A7A1	; □	LETTRE MINUSCULE LATINE G BARRÉE OBLIQUEMENT
U+A7A3	; □	LETTRE MINUSCULE LATINE K BARRÉE OBLIQUEMENT

U+A7A5 ; LETTRE MINUSCULE LATINE N BARRÉE OBLIQUEMENT
U+A7A7 ; LETTRE MINUSCULE LATINE R BARRÉE OBLIQUEMENT
U+A7A9 ; LETTRE MINUSCULE LATINE S BARRÉE OBLIQUEMENT

Ajouts pour L'UPA

U+A7FA ; # LETTRE LATINE PETITE CAPITALE CULBUTÉE M