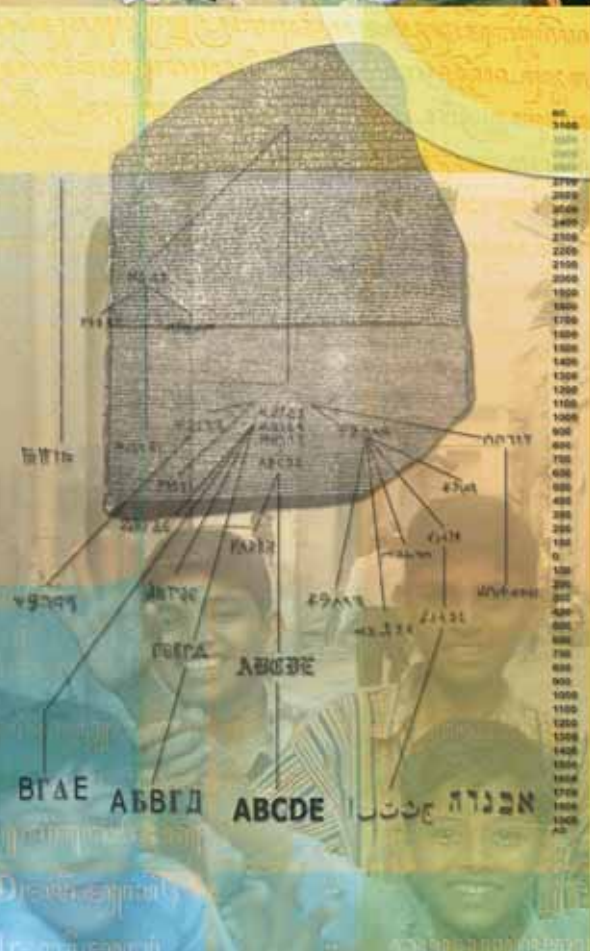


initiative B@BEL

www.unesco.org/webworld/babel



Organisation des Nations Unies
pour l'éducation, la science et la culture



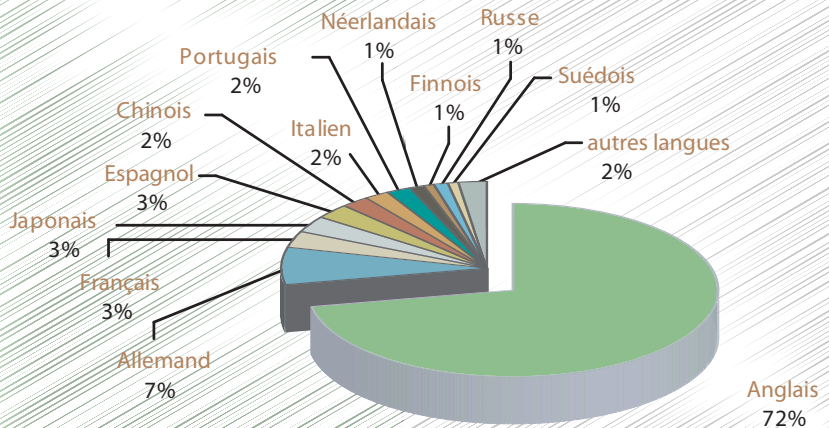
Dans la société du savoir en voie d'émergence, l'information est une ressource clé pour la prise des décisions, l'autonomisation, la participation sociale et le développement humain. Aussi l'aptitude à accéder aux connaissances et aux informations – et à les influencer – sur les réseaux d'information mondiaux tels que l'Internet influera-t-elle sur l'aptitude de chacun à participer à la société et à déterminer la qualité de sa propre vie. La langue, notre premier moyen de communication, est évidemment cruciale dans ce processus. Mais les langues ne sont pas seulement des instruments de communication : elles sont les véhicules de systèmes de valeurs, d'expressions culturelles, un facteur important de l'identité des groupes et des individus. La diversité linguistique est un élément essentiel du patrimoine vivant de l'humanité qu'il faut cultiver et promouvoir.

L'Initiative B@bel contribue à la réalisation de l'objectif de l'UNESCO - construire des sociétés du savoir, favoriser l'accès universel à l'information et le développement impulsé par l'information – par une stratégie à trois volets incluant des actions dans les domaines des politiques, de la sensibilisation et de la mise en œuvre de projets pilotes novateurs. Ces activités menées en collaboration avec des partenaires visent à soutenir l'élaboration de contenus et d'outils multilingues sur les réseaux mondiaux d'information et à protéger les langues en péril et leur héritage de savoirs traditionnels. L'Initiative B@bel est un exemple pratique de l'application de stratégies élaborées dans des instruments normatifs de l'UNESCO adoptés par la communauté internationale comme la Recommandation sur la promotion et l'usage du multilinguisme et l'accès universel au cyberspace.

Alors que de nombreux pays ont deux ou même trois langues officielles, la Constitution de la République d'Afrique du Sud de 1993 reconnaît onze langues officielles.

La plupart des alphabets contiennent de 20 à 30 symboles, mais la complexité relative des systèmes phonétiques de différentes langues conduit à des alphabets de tailles diverses. Le plus petit alphabet est, semble-t-il, celui des Rotokas, en usage dans les Iles Salomon, qui compte onze lettres. Le plus grand est peut-être celui de la langue khmère, qui compte 74 lettres.

Estimation des contenus par langue sur le Web



Source : Proportions of languages on the web in a random sample of web pages. O'Neill, Lavoie et Bennett, 2003.

Principes directeurs et études à l'appui des initiatives en matière de politiques et de sensibilisation

L'éducation dans un monde multilingue

Ce document-cadre a été élaboré par l'UNESCO en collaboration avec un groupe d'éminents experts internationaux dans les domaines de la diversité culturelle et de l'éducation. Outre qu'il incorpore des perspectives régionales importantes sur les enjeux et questions clés, il rappelle et souligne les accords et

engagements internationaux existants des gouvernements en faveur du pluralisme linguistique et de l'éducation multilingue sur leur territoire. Le document est disponible en anglais, arabe, chinois, espagnol, français et russe.

Principes directeurs pour la planification des terminologies nationales

Une communauté linguistique qui n'a pas mis au point de terminologies scientifique et technique pour sa langue est inévitablement conduite à utiliser une autre langue, peut-être étrangère, pour exprimer ces concepts. Ce processus renforce les fractures linguistique et scientifique et peut contribuer à faire considérer comme moins importantes les langues locales, à les marginaliser et à les exclure de la société du savoir au sens large.

La planification des terminologies joue un rôle important pour ce qui est de relever ces défis et de soutenir l'innovation, le partage des connaissances, les activités éducatives et économiques et le processus de développement en général. En apportant aux responsables des politiques, en particulier à ceux des pays en développement, une compréhension plus claire de ce domaine souvent négligé, les principes directeurs cherchent à renforcer le rôle des langues locales dans la société du savoir.

Edition thématique de la Revue internationale des sociétés multiculturelles « Les langues sur l'Internet »

L'Internet offre un excellent exemple de la façon dont les TIC influent sur les profils d'utilisation des langues. Ces évolutions peuvent avoir un certain nombre d'effets sociaux, économiques et autres de grande portée, ce qui fait que les politiques publiques en matière de langues peuvent apporter des réponses utiles à ces nouveaux défis et à ces nouvelles opportunités. Les recherches menées innovent en collectant et en analysant des données empiriques sur les profils d'utilisation des langues sur l'Internet en Afrique, dans le monde arabe, en Asie et en Europe.



Echantillon d'implantations d'écriture

Même si un système peut être techniquement capable de manier de multiples langues/écritures, il est nécessaire de fournir les modules de base (polices, claviers, etc.) permettant d'entrer et d'afficher les données dans ces systèmes d'écriture. Parallèlement

à la mise au point d'outils multilingues, nous avons élaboré des implantations d'écriture pour le Myanmar et les langues panafricaines, de telle sorte que ces outils soient utilisés par les communautés linguistiques concernées.

Principes directeurs pour la mise en place d'un soutien aux systèmes d'écriture

Ce document présente un aperçu des questions, activités et stratégies clés destinées à soutenir le développement des systèmes d'écriture dans l'environnement électronique.

Format d'échange des données sur les écritures complexes

Pour bien assurer la création et la diffusion d'informations, en particulier dans le cas des données sur les écritures complexes, il est besoin d'un format standard d'échange des données.

Droits de propriété intellectuelle et systèmes d'écriture

La mise au point d'implantations informatiques adéquates des systèmes d'écriture et le traitement des données pour les écritures complexes constituent une tâche difficile, exigeant divers types d'expertise et d'investissements. En conséquence, les droits de propriété intellectuelle (DPI) risquent de constituer un obstacle au perfectionnement, à la diffusion et à l'utilisation de ces systèmes par les communautés

linguistiques désavantagées. Ce document explore des modèles et des politiques susceptibles de stimuler le développement et la diffusion des systèmes d'écriture tout en protégeant les intérêts commerciaux et les DPI et en favorisant la création de contenus dans nombre de langues moins répandues.



Guide pour la collecte de données linguistiques par les non linguistes

Traditionnellement, les linguistes ont visité diverses régions du monde pour « découvrir », étudier et enregistrer de « nouvelles » langues. Le nombre limité des linguistes, l'accélération rapide de la disparition des langues menacées, ainsi que d'autres problèmes, font que cette approche est désormais inadéquate. Ce projet cherche à donner aux communautés intéressées parlant les langues menacées les moyens de préserver leur patrimoine linguistique en mettant à leur disposition les méthodologies appropriées pour le documenter. Cela permettra d'enrayer la disparition des langues, d'approfondir l'étude de ces langues et de soutenir ultérieurement l'élaboration et la diffusion de matériels pédagogiques et de contenus autochtones sous forme imprimée et sous forme numérique.



On estime qu'environ 20% des langues du monde n'ont pas de forme écrite. L'émergence de la technologie du multimédia, la gestion de la communication vocale numérisée et des innovations comme les systèmes de reconnaissance vocale et de traduction offrent de nouvelles possibilités pour la préservation du riche patrimoine de ces langues dans l'environnement numérique d'aujourd'hui.

Plus de la moitié des 6000 langues du monde sont menacées et 90% des langues du monde ne sont pas réellement représentées sur l'Internet.



Projets pilotes, ressources en ligne et outils multilingues

Rapport sur le multilinguisme sur l'Internet

Combien de langues sont présentes sur le Web ? Combien de pages existent dans chacune de ces langues ? Quels sont les problèmes à résoudre pour mesurer la diversité linguistique sur l'Internet et comment améliorer la technologie utilisée à cet effet ? Quelles forces façonnent la trame linguistique de l'Internet et comment la situation évolue-t-elle avec le

temps ? Ce document, élaboré en coopération avec l'Institut de statistique de l'UNESCO et incorporé dans la collection de publications sur le Sommet mondial sur la société de l'information, cherche à nous faire mieux comprendre cette question et d'autres questions connexes.

La Plate-forme de gestion de contenus multilingues (MLCM)

En coopération avec Infoterm, le Centre international d'information pour la terminologie (www.infoterm.org) à Vienne, l'Initiative B@bel a mis au point un système de gestion des informations multilingues fondé sur la plate-forme CDS/ISIS de l'UNESCO assurant la compatibilité et la réutilisation. Le logiciel, qui incorpore des fonctionnalités telles que la numérisation

des codes-barres ou la compatibilité avec le système Office de Microsoft, est aussi capable d'effectuer des recherches multilingues dans le réseau local et sur l'Internet. La MLCM, qui est ciblée sur les besoins de gestion des SME, contient des bases de données factuelles et terminologiques étendues et extensibles.

SILA – un navigateur Web multilingue

Le Web est un forum important pour la diffusion et l'échange d'informations et l'interaction, mais beaucoup de communautés linguistiques en sont exclues faute d'outils Web pour leurs écritures. En intégrant la technologie Graphite du SIL pour l'affichage des systèmes d'écriture complexes dans le système d'application Internet Mozilla (navigateur,

courrier électronique, composeur HTML, IRC Chat et carnet d'adresses) et en élaborant des polices de soutien, l'Initiative B@bel crée les outils qui permettent aux communautés linguistiques marginalisées de mettre à profit le Web et d'interagir sur le Web. Dans un premier temps, un soutien est fourni au birman et aux langues africaines.

Version modifiée de WorldPad

WorldPad est un éditeur de texte et un outil d'entrée des données simple qui peut sans accroc traiter les

écritures complexes (non latines) à l'aide du moteur de rendu Graphite mis au point par SIL International.

Commande d'édition Graphite

Il s'agit d'une Commande d'édition Graphite beta pour les applications d'entrée des données qui utilise le moteur Graphite de SIL International et un kit de développement logiciel. Cet outil est particulièrement utile pour les applications dans lesquelles les textes contiennent plusieurs systèmes d'écriture (par

exemple les références bibliographiques comportant les titres d'un ouvrage tels qu'ils figurent dans diverses traductions) et il accepte les systèmes d'écriture de gauche à droite comme de droite à gauche (comme par exemple ceux de l'arabe ou de l'hébreu).



Aujourd'hui, environ 90% des contenus de l'Internet ne sont disponibles que dans 12 langues.

Base de connaissances sur le patrimoine culturel linguistique

Ce projet combine les connaissances thématiques sur les politiques culturelles avec le système de traduction Universal Network Language mis au point par la UNDL Foundation. Une fois achevée, cette base de connaissances offrira une plate-forme comprenant des dictionnaires et des terminologies spécialisées intégrés modulables et une base de données sur les politiques

culturelles accessible en 6 langues (anglais, arabe, espagnol, français, italien, russe). Comme le système est modulaire et contient un kit de formation, un support pour les autres langues et domaines de connaissances peut être intégré créant ainsi une encyclopédie multilingue de connaissances.

Fonds d'enregistrements sonores accessibles sur l'Internet des langues du Caucase en péril

Il y a plus de 40 langues menacées de disparition dans la région du Caucase, dont beaucoup présentent des caractéristiques linguistiques uniques et remarquables. Dans le cadre de ses efforts pour protéger les langues et le riche corpus de savoirs et de traditions qui leur est associé, un fonds d'enregistrements sonores des langues en péril de la région du Caucase accessible sur

l'Internet est en voie d'être constitué. Outre qu'elle aide à sensibiliser à l'urgence de la situation de ces langues, cette ressource permettra aux linguistes d'approfondir l'étude de ces langues et contribuera à la préservation et peut-être à la renaissance du patrimoine linguistique et culturel pour les générations futures.

Projet de dictionnaire abkhaze-géorgien

La Géorgie est abondamment dotée de ressources culturelles et naturelles mais les conflits civils de ces dix dernières années ont fortement endommagé le tissu culturel, social, physique et économique. A l'appui du processus de construction de la paix, une coopération a été instaurée entre les institutions académiques en vue d'élaborer un dictionnaire

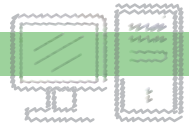
bilingue (abkhaze-géorgien) imprimé et électronique. Outre qu'il développe les ressources multilingues, ce projet, en contribuant à créer des passerelles de communication et à ouvrir/créer des canaux de dialogue, de commerce et d'autres échanges de communication, cherche à contribuer à l'établissement d'une paix durable.



Le dictionnaire polyvalent abkhaze-géorgien en deux volumes



Standards informatiques éthiopiens



Les usagers de l'ordinateur qui souhaitent travailler dans les systèmes d'écriture non latins risquent de se heurter à des difficultés pratiques comme l'incompatibilité des systèmes d'encodage des informations et l'absence de standards pour l'agencement des claviers et la translittération. En conséquence, le partage d'informations entre utilisateurs peut être difficile voire impossible et l'entrée des données comme d'autres processus risquent de devenir inefficients, susceptibles d'erreurs et coûteux, décourageant de ce fait l'utilisation des langues locales dans l'environnement numérique. Face à ce problème,

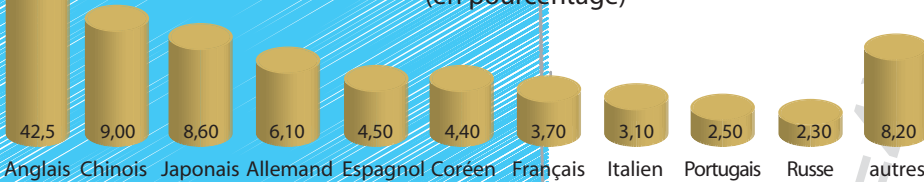
L'Initiative B@bel, de concert avec le gouvernement éthiopien et d'autres partenaires de la communauté internationale, du secteur privé et des milieux académiques, a réalisé une étude en vue d'évaluer et d'identifier les besoins de soutien informatique des langues éthiopiennes et d'élaborer des standards pour l'encodage des caractères éthiopiens, l'agencement des claviers et la translittération de l'amharique en caractères latins. Sur cette base, un plan d'action visant à résoudre ces problèmes et un groupe de travail ont été établis. On s'efforce actuellement d'obtenir les fonds nécessaires pour mettre en œuvre ces changements.

Atlas UNESCO des langues en péril dans le monde

L'Initiative B@bel est en train de mettre au point un prolongement en ligne de cette publication très populaire. Ce produit en ligne apportera de nouveaux aperçus sur la vie socio-linguistique et culturelle des peuples qui utilisent ces langues en péril. Il s'agira essentiellement, dans un premier temps, de fournir des informations sur les communautés africaines parlant des langues en péril.

Plus tard, des fonctionnalités multimédia interactives ainsi que le soutien de la gestion des données SIG (systèmes d'information géographique) y seront intégrés. Nous espérons non seulement enrichir l'expérience des utilisateurs et promouvoir la sensibilisation, mais aussi créer un outil qui permettra de fournir des informations aisément accessibles et régulièrement mises à jour sur la situation de ces langues.

Les populations linguistiques les plus nombreuses sur l'Internet (en pourcentage)



Les dix langues les plus parlées dans le monde (en milliards de locuteurs)

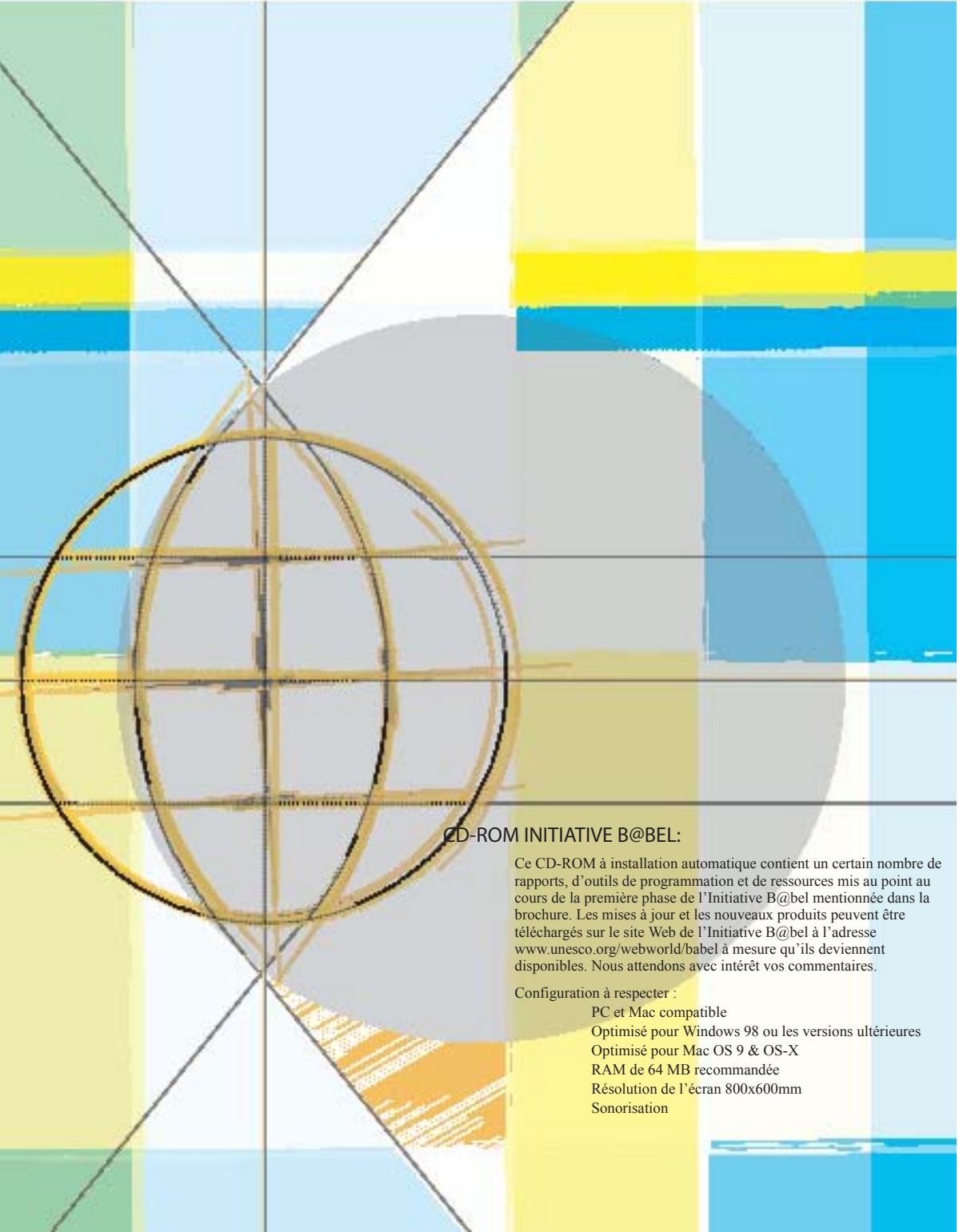


Source :

The Cambridge Encyclopedia of Languages. D. Crystal, 1992. p289



Arrière-plan: langage de codage HTML sur le Web



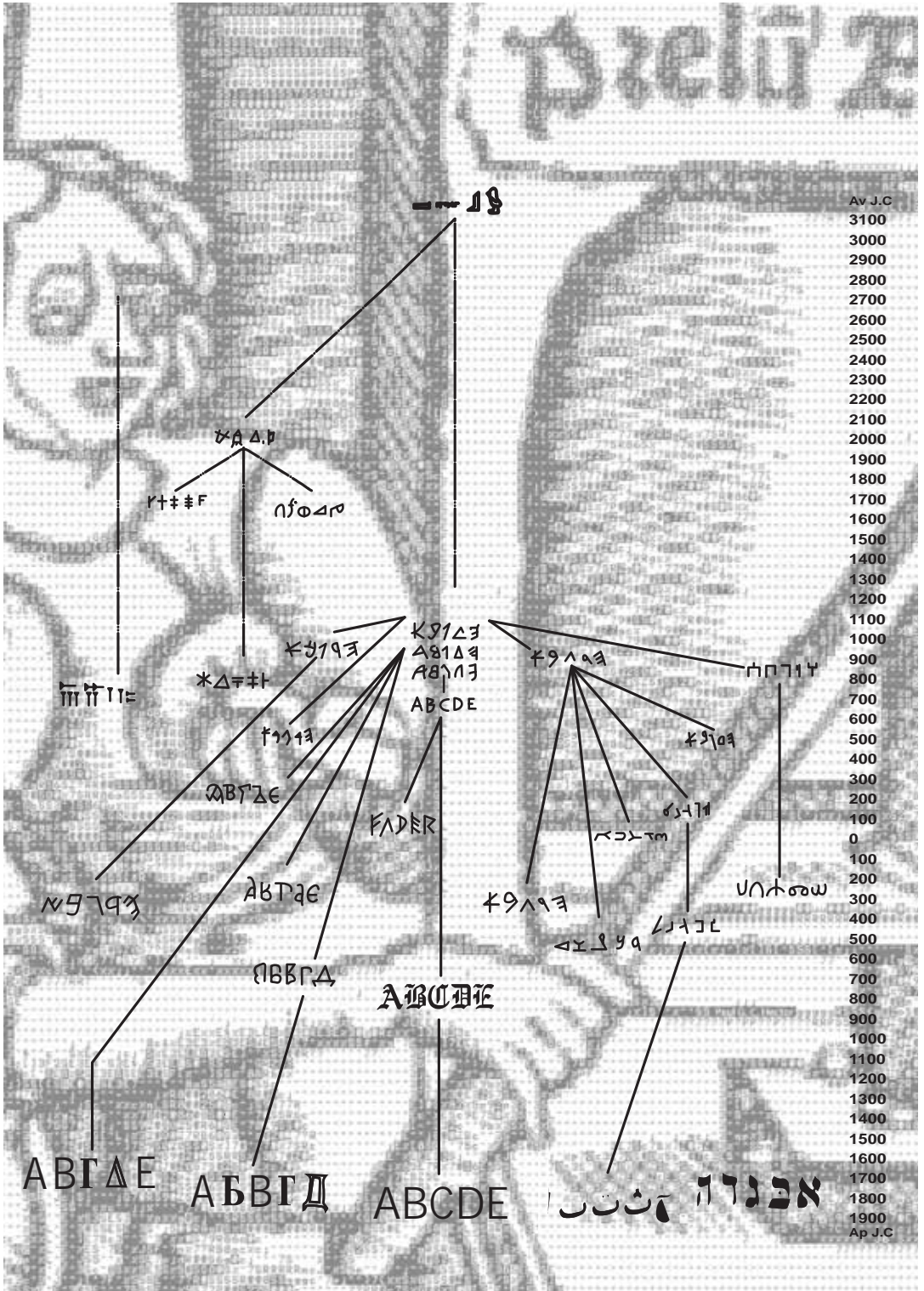
CD-ROM INITIATIVE B@BEL:

Ce CD-ROM à installation automatique contient un certain nombre de rapports, d'outils de programmation et de ressources mis au point au cours de la première phase de l'Initiative B@bel mentionnée dans la brochure. Les mises à jour et les nouveaux produits peuvent être téléchargés sur le site Web de l'Initiative B@bel à l'adresse www.unesco.org/webworld/babel à mesure qu'ils deviennent disponibles. Nous attendons avec intérêt vos commentaires.

Configuration à respecter :

- PC et Mac compatible
- Optimisé pour Windows 98 ou les versions ultérieures
- Optimisé pour Mac OS 9 & OS-X
- RAM de 64 MB recommandée
- Résolution de l'écran 800x600mm
- Sonorisation

« Les alphabets du monde au cours des siècles »



Quelques initiatives à venir:

- Systèmes de traduction multilingues
- Apprentissage des langues en ligne
- De la voix à la parole
- Encodage en script Unicode
- Mise au point de polices et formation des concepteurs de polices
- Soutien linguistique aux personnes handicapées



Organisation des Nations Unies
pour l'éducation, la science et la culture

Références et remerciements :

- The Cambridge Encyclopedia of Languages
- Mary Evans Picture Library, Londres, Angleterre
- Stockholm Challenge
- Atlas UNESCO des langues en péril dans le monde
- UNESCO Photobank
- M. Victor Montviloff
- Lucia Dessart, Christina Human, Nathalie Da Fonseca, Sabine Schaefer, Marina Forzani, Elena Iasyreva
- M. Paul G. C. Hector

Partenaires et soutiens du projet à ce jour :

- Gouvernement de la Principauté de Monaco
- Gouvernement de la République fédérale démocratique d'Éthiopie
- Commission nationale géorgienne pour l'UNESCO
- Agence suisse pour le développement et la coopération /ICT4D/Global Knowledge Partnership
- Académie africaine des langues
- Infoterm
- Institut international de l'alphabétisation
- UNDL Foundation
- Association éthiopienne pour les standards informatiques
- SIL International
- Université d'Aston
- Matenadaran, Institut Mashtots
- Université de Californie (Berkeley)



Thèmes de la conception graphique :

La conception graphique de la brochure incorpore des éléments tels que les couleurs, la Terre et les symboles de diverses origines culturelles qui préparent le lecteur à explorer la communication multilingue. Le sens de la flèche centrale oriente le lecteur dans un voyage des origines du langage vers les médias numériques actuels. Son caractère cyclique implique une dynamique ni passé ni futur mais bien ancré dans l'optique d'un village planétaire où tous ont droit à l'information. Les caractères de divers systèmes d'écriture, constituant le fond de chaque page, nous rappellent la diversité de ces systèmes et la nécessité de les protéger, de les préserver et de les représenter TOUS numériquement. Il y a aussi d'autres motifs que vous êtes invités à découvrir par vous-mêmes comme la rencontre fortuite entre le code lithographique de Gutenberg et le code ASCII,...

Jean-Luc Seyfried, Concepteur freelance
www.bitspersecond.org

Pour tout complément d'information, ou pour participer à nos activités,
veuillez nous contacter à l'adresse suivante :

Secteur de la communication et de l'information Division de la société de l'information

1 rue Miollis 75732 Paris Cedex 15, France
Tél. +33 (0)1 45 68 43 62
Fax +33 (0)1 45 68 55 83

Site : www.unesco.org/webworld/babel
Courriel : Babel@unesco.org