

Présentation

Démocratiser l'outil informatique : telle est en fait l'une des préoccupations majeures des décideurs en matière de technologies de l'information et de la communication (TICs). En témoigne la tendance actuelle de ces technologies, à savoir l'orientation de l'utilisation des nouvelles productions technologiques (logicielles et matérielles) dans le sens de faciliter au grand public l'accès à l'information et ce, en prenant en considération les spécificités locales de l'utilisateur : langue, culture, niveau intellectuel, etc.

Les communautés et organisations (nationales et internationales) œuvrant dans le domaine de l'information et de la communication ont bien saisi les enjeux de la valorisation de la langue et de la culture dans les TICs. Les langues naturelles peu dotées sur le plan informatique telles que la langue amazighe, ont vu émerger plusieurs plates-formes et projets de recherche nationaux et internationaux qui favorisent le développement de la langue et de la culture sur la base des nouvelles technologies de l'information.

Le présent numéro de la revue *ⵝⵓⵍⵉⵎⵉⵏⵉ-Asinag*, constitué de douze articles (huit en français, deux en anglais et deux en arabe) exclusivement consacrés au domaine des technologies appliquées à la langue, vise la mise en exergue des avancées réalisées dans ce champ de recherche.

La contribution de Patrick Andries dresse un tableau panoramique à propos de l'introduction de l'amazighe dans les nouvelles technologies à l'échelle internationale. Il retrace les différentes étapes en commençant par le niveau le plus basique qu'est le codage du système de l'alphabet tifinaghe. Ensuite, il aborde l'intégration du clavier amazighe dans les systèmes d'exploitation (Windows, en particulier), les polices de caractères et la prise en charge du tifinaghe dans les langages Web pour arriver au stade le plus avancé. C'est-à-dire le renseignement du registre Unicode CLDR (Common Locale Data Repository) et la reconnaissance du système graphique tifinaghe dans les bibliothèques logicielles.

Dans leur article, Fadoua Ataa Allah et Siham Boulaknadel traitent du rôle des technologies de l'information dans la promotion de l'amazighe au niveau national. La technologie, procédé de sauvegarde du patrimoine culturel, y est considérée comme un vecteur de l'aménagement linguistique de l'amazighe (différentes opérations de codification, gestion des ressources, etc.). Cette technologie offre également des moyens qui facilitent l'enseignement de la langue en permettant la promotion de l'amazighe par l'exploitation de la

technologie éducationnelle, de l'apprentissage à distance et des outils du TAL amazighe. Aussi les auteurs proposent-elles une méthode élaborée pour l'informatisation de l'amazighe à court, moyen et long termes.

Le projet de boîte à outils unifiés de la linguistique computationnelle « SmallCodes », spécialement conçu pour les langues peu dotées, est présenté par Carlo Zoli. Celui-ci explique l'objectif du projet en montrant les démarches des champs disciplinaires mobilisées dans sa recherche et présente une démonstration des performances du système développé. Il propose ensuite une analyse des approches existantes et essaie une projection de son approche sur l'amazighe.

Nora Tizgiri et Henri Hudrisier exposent le projet *Humanité DigitMaghreb* qui vise la construction d'une bibliothèque de corpus multilingues (oraux, littéraires et linguistiques), synergiques et normalisés pouvant garantir la communisation inter-langues. Ils expliquent l'utilisation de la norme historique TEI (Text Encoding Initiative) en vue de réaliser leurs objectifs de départ et soulignent les difficultés rencontrées au niveau des Sciences humaines.

Dans le domaine de la construction des corpus annotés, Mohamed Outahajala *et al.* proposent d'améliorer le rendement de leur étiqueteur morphosyntaxique par l'exploitation des processus CACs (Champs Aléatoires Conditionnels). Sur un texte prétraité, ils expérimentent leur système ainsi augmenté et comparent les résultats obtenus par le choix des données d'auto-apprentissage ayant une valeur de confiance élevée à ceux obtenus sur la base de données aléatoires.

L'article de Violetta Cavalli Sfortza traite deux aspects de la pertinence des textes et leur degré de facilité dans le processus de lecture pour la langue arabe. Il s'agit de la lisibilité du texte, indépendamment de l'apprenant ou du programme d'enseignement, et du degré d'appropriation du texte dépendant des connaissances et du vocabulaire de l'apprenant. Après avoir passé en revue la littérature existant sur la question, l'auteur propose un modèle prédictif de cette pertinence et met en œuvre des outils techniques pour une partie de l'analyse (*e.g.* MADA). Elle essaie une projection sur la lisibilité de la langue amazighe, suite à quoi elle présente les limites de l'étude.

La reconnaissance optique des caractères tfinaghés est soulevée par Ali Rachidi. Ce dernier présente une synthèse des travaux réalisés dans ce domaine au niveau national et procède à leur comparaison à la lumière des bases de données des caractères tfinaghés utilisées.

La reconnaissance de la parole est soulevée par Abenaou *et al.* Les auteurs détaillent les méthodes et algorithmes adoptés pour la synthèse des mots en amazighe en proposant des techniques pour optimiser les données et augmenter la vitesse de la reconnaissance. Ils concluent par la présentation de résultats expérimentaux et une évaluation de la performance du système.

La contribution conjointe de Hammou Fadili et Malika Chakiri expose la conception et l'élaboration d'un modèle d'ontologies de domaines ; modèle enrichi par des liens lexico-sémantiques associés aux fonctions lexicales de la Théorie Sens-Texte (Mel'cuk, 1997) et la notion du contexte. Les auteurs fournissent quelques définitions des outils de travail utilisés tels que les ontologies, les langages informatiques utilisés et la relation contexte-ontologie. Ils abordent la constitution du corpus ontologique et notent des difficultés dans le cas de l'amazighe, auxquelles ils proposent des solutions linguistiques et technologiques.

La contribution de F. Ataa Allah, S. Boulaknadel et H. Souifi est consacrée à l'élaboration d'un *Jeu d'étiquettes morphosyntaxiques* de la langue amazighe. Le travail se fonde sur les recommandations EAGLES qui visent la réutilisation des corpus et la comparaison des langues dans le domaine du traitement automatique du langage naturel.

Le domaine des technologies appliquées à l'apprentissage est au centre de l'article de Robert Bibeau. Y sont traitées les problématiques d'accessibilité et de normalisation des ressources d'enseignement et d'apprentissage, notamment dans le cas de plusieurs instances, utilisant différentes langues, sous divers environnements technologiques. Le profil NORMETIC, variante d'application de la norme IEEE 1484.12.1 (LOM) des métadonnées d'objets d'apprentissage, est conçu pour répondre à ces exigences. Eurêka, banque des ressources d'apprentissage, suffisamment décrite dans le présent article, est un exemple de bases de données compatibles avec le profil NORMETIC.

Le volet en langue arabe contient deux contributions. Dans l'une, Mohamed Lguensat aborde la graphie du tifinaghe et propose une approche pour son aménagement. Il appuie sa thèse par des modèles concrets et souligne les défis liés, d'une part, aux déterminants techniques et informatiques et, d'autre part, aux éléments de la communication visuelle. Dans l'autre, Hassan Jaa et Youssef Ait Ouguengay brossent un panorama général des travaux qui ont accompagné, au cours de la dernière décennie, l'introduction de l'amazighe dans les nouvelles technologies de l'information ; notamment les efforts consentis au niveau national et à l'IRCAM.

Le dossier est agrémenté d'un entretien avec Lahbib Zenkouar et Patrick Andries sur des questions relatives à la promotion de l'amazighe dans les TICs ainsi qu'à la recherche scientifique dans ce domaine.

Le présent numéro comprend également un compte rendu élaboré par Fouad Brigui sur l'ouvrage d'Ahmed Boukous intitulé : *Revitalisation de la langue amazighe : Défis, enjeux et stratégies*, publié par l'Institut Royal de la Culture Amazighe en 2012.

La rubrique *Résumés de thèses* est destinée à faire connaître des travaux académiques récents et inédits sur l'amazighe. Y sont livrés deux résumés

portant sur la reconnaissance automatique des caractères tifinaghes imprimés et manuscrits. Le premier, dû à Youssef Es-saady, traite de deux approches disjointes : l'une, syntaxique, utilise les automates finis ; l'autre, neuronale. L'auteur analyse les résultats des deux approches et conclut à leur pertinence. Le second, de Moustapha Amrouche, propose, d'abord, une modélisation markovienne complémentée par une technique de caractérisation de caractères isolés, et, ensuite, la combinaison d'analyses par chemins discriminants (DP-HMM) et des caractéristiques morphologiques de la graphie amazighe. Le test de rendement et les résultats des deux thèses sont effectués sur une base de données nommée AMHCD, élaborée dans le cadre du travail de Youssef Es-saady.

La Direction et le Comité de rédaction de la Revue tiennent à exprimer leurs plus vifs remerciements à toutes les personnes qui ont apporté une quelconque contribution à la réalisation de ce numéro : Patrick Andries, Fadoua Ataa Allah, Aïcha Bouhjar, Belaïd Bouikhalene, Siham Boulaknadel, El Houssine Bouyakhf, Violetta Cavalli-Sfortza, Abdelkrim El Moukhtari, Lahbib Fouad, Abdelfattah Hamdani, El Mehdi Iazzi, Rachid Laabdalaoui, Mohamed Maamouri, Kamal Naït-Zerrad, Patrice Pognan, Mohamed Yeou, Abdellah Yousfi et Lahbib Zenkouar.

⦿⦿⦿-Asinag