

Une base de données au service de la toponymie corse

1. Préambule : la toponymie de la Corse

Les toponymes de la Corse ont une forme orale qui s'inscrit dans le cadre dialectal traditionnel de l'île. Depuis le Moyen-Âge, le toscan a offert, de par son statut véhiculaire, sa graphie aux noms de lieux insulaires dans la documentation écrite. Traditionnellement, les locuteurs, en possession des deux codes, étaient en mesure de restituer dans la forme corse les noms qui étaient graphiquement toscans. Dans les énoncés en langue française ou en contexte officiel, jusqu'à quelques années de ça, la forme toscane était employée dans la plupart des cas pour les noms des villes, régions et hameaux principaux.

L'équilibre corse-toscan s'est progressivement rompu avec la francisation linguistique de l'île. Partant de la forme toscane du nom traditionnel, le contact avec le français conduit à une altération des toponymes : la prononciation tend à s'effectuer de plus en plus en fonction de la valeur des graphèmes du français. On opposera ainsi les noms prononcés selon la tradition corse¹ et toscane, face à une prononciation inspirée du français, fait qui n'était, jusqu'à une vingtaine d'années de ça, réservé qu'aux noms des principales villes de l'île. Par exemple :

| Graphie toscane | forme corse | forme altérée |
|-----------------------|--------------|---------------|
| (a) <i>Murato</i> | [mur'atu] | [myrato] |
| (b) <i>Salice</i> | [u z'aldʒɛ] | [salis] |
| (c) <i>Macinaggio</i> | [matʃin'aʒu] | [masinaʒ] |
| (d) <i>Ferringule</i> | [fɛr'iŋgule] | [farinɔ] |

Selon les schémas prosodiques du français, le déplacement d'accent sur la dernière syllabe tend à se généraliser (exemples *supra*) et les noms à s'abrégier, particulièrement les noms composés :

| Graphie toscane | forme corse | forme altérée |
|--------------------------|------------------|---------------|
| (e) <i>Pietra Serena</i> | [p'etra z'erena] | [pjetra] |
| (f) <i>Bustanico</i> | [bust'aniku] | [bys] |

¹ Nous donnons ici une transcription moyenne de la prononciation corse.

On observe également des réinterprétations de toponymes, selon un exemple connu *e Canne* (formé depuis *canna* ‘roseau’) devenu *les Cannes*, ou encore des traductions comme *San Fiurenzu* > *Saint Florent*, *Stretta di u Soli* > *Rue du Soleil*. En outre, l’urbanisation importante que connaît la Corse s’accompagne bien entendu d’un développement important de néotoponymes, souvent en français (*Les Résidences de... Les Jardins de... Les Terrasses de...*).

Ces évolutions, liées au contact des variétés, sont accrues par l’abandon des campagnes, les mutations démographiques et le recul constant de la pratique linguistique du corse comme vernaculaire, ainsi que la perte du toscan comme variété véhiculaire. Ces faits conduisent à une altération profonde voire à la disparition du patrimoine toponymique de l’île.

Du côté des études scientifiques relatives à la toponymie corse, au-delà du faible nombre d’études d’intérêt notable, le bilan dressé (Medori 2009) a montré qu’il était indispensable de disposer de données fiables, issues de la tradition orale, pour mener des analyses scientifiques sur les noms de lieux de la Corse.

Par ailleurs, les collectivités locales – particulièrement la *Collectivité Territoriale de Corse* (CTC) – se montrent désireuses, depuis quelques années, de disposer d’un outil répondant à des besoins divers tels que l’adressage, la signalétique ou les documents d’urbanismes, mais surtout la valorisation de la toponymie corse qui fait partie des actions en faveur de la langue corse dans le cadre de la *Cartula di a Lingua Corsa* (mise en œuvre par la CTC auprès d’institutions et de partenaires privés).

2. Le CESIT Corsica : création d’une base de données toponymiques

La situation que nous avons décrite succinctement, a conduit à la création, en 2009, d’une association Loi 1901, le *Comité d’Etudes Scientifiques et Informatiques de la Toponymie de la Corse*, CESIT Corsica. Cette association a été fondée par des scientifiques (linguistes, informaticiens, historiens), et par des amateurs d’onomastique; elle s’est dotée d’un Conseil Scientifique afin d’assurer la qualité de sa démarche. Elle s’est donnée pour but d’inventorier, par le biais d’enquêtes de terrain réalisées auprès de locuteurs natifs, les toponymes traditionnels de la Corse. Le recueil des toponymes s’appuie sur la cartographie des données cadastrales (cadastre napoléonien et rénové), et de l’IGN. Si le travail de l’association s’articule essentiellement autour du recueil des toponymes oraux, certains membres de l’association dédient leur temps aux recherches en archives.

Après un séjour d’étude auprès de l’*Atlante Toponomastico del Piemonte Montano* à Turin grâce au soutien de la *Corsica Ferries*, le chantier de la base de données toponymiques a été entamé. Conçue sur la plateforme DynMap pendant l’année 2010 par la société *Cyrnéa Info Géographie*, avec le concours d’un financement de la *Collectivité Territoriale de Corse*, elle a été mise en ligne en juillet 2011 à l’adresse suivante : <<http://www.cesitcorsica.org>>.

Désormais, nous allons présenter brièvement les contenus de la base de données². Celle-ci accueille le dépouillement des noms de lieux qui sont géolocalisés. En préambule, il faut préciser que cinq catégories de toponymes sont distinguées:

- microtoponymes ;
- macrotoponymes ;
- noms de communes ;
- petits choronymes (vallées, chaînes montagneuses) ;
- choronymes (régions).

Pour des raisons techniques, les microtoponymes sont matérialisés par des points, tandis que les macrotoponymes et les choronymes sont matérialisés par des polygones. Les fonds cartographiques, rendus accessibles par une convention avec l'*Office de l'Environnement de la Corse* sont

- le Top 25 (IGN) ;
- le cadastre rénové ;
- la photographie aérienne de 2002.

Pour chaque toponyme saisi, les informations consignées sur les fiches sont les suivantes (macrostructure) :

- forme phonique et orthographique ;
- ethnonyme corse pour les lieux habités (régions, communes, hameaux, quartiers) ;
- informations INSEE (nom de commune, toponyme et référence) ;
- métadonnées : noms de l'enquêteur, de l'informateur, et du collecteur pour les sources écrites (cadastres, archives), dates ;
- données cadastrales et du Plan terrier de la Corse (formes et références dont numéros de parcelles ou de rouleau) ;
- données d'archives (publiées et inédites) ;
- analyse linguistique (classification, étymologie, reconstruction du lemme et du signifié originels) ;
- référent ;
- documents sonores, textuels, iconographiques.

Le public n'a accès qu'à des fiches simplifiées³ avec nom corse, nom du cadastre napoléonien, ethnonymes. Il faut préciser que lorsque la forme corse d'un toponyme n'a pas été recueillie oralement, et quelle qu'en soit la raison (aléas de l'enquête ou perte du nom dans la mémoire collective), c'est le nom du cadastre napoléonien qui figure sur l'en-tête de la fiche ; il est noté avec un astérisque pour marquer le fait qu'il n'est pas attesté à l'oral (**Pietre Ritte*, **Aja al Chioso*).

² Prochainement en accès multilingue : français, corse, italien et anglais.

³ Le site fait actuellement l'objet d'un travail d'ergonomie afin d'améliorer l'accessibilité et la maniabilité des données par le public. Il est à noter que le public peut aussi suggérer des toponymes par le biais d'un forum avec carte interactive

Les modules de travail internes sont, comme le montre la liste *supra*, beaucoup plus complexes que ne le suggère l'accès externe. L'analyse linguistique des toponymes fait toujours l'objet d'essais et de réflexions qui seront soumis prochainement au conseil scientifique pour évaluation. L'objet est, à terme, de pouvoir conduire des recherches complexes telles que les :

1. continuateurs d'un étymon ;
2. correspondants onomastiques d'un élément lexical y compris s'il a disparu de l'usage ;
3. formes onomastiques qui correspondent à un signifié d'appellatif originel.

La géolocalisation des données et le développement à venir de modules de symbolisation devraient permettre d'appréhender la base de données comme un outil géolinguistique. Il devrait être possible, ainsi, d'établir par les traces toponymiques⁴, l'ancienne extension d'une unité lexicale par comparaison avec des cartes d'atlas linguistiques.

3. Quelques résultats

En raison de la forte dégradation de la situation linguistique de la Corse, le CESIT Corsica a choisi de faire des enquêtes de terrain sa priorité. A ce jour, au niveau quantitatif, une soixantaine de communes a été couverte (en totalité ou partiellement)⁵, et les données sont saisies progressivement. Les 5000 premières fiches ont fait l'objet d'une première phase de correction et la vérification de chaque forme phonique et graphique est en cours par un dépouillement de contrôle des enregistrements.

D'un point de vue qualitatif, voire scientifique, le collectage des toponymes oraux, confronté à l'analyse linguistique, permet de réviser certaines données offertes par les cadastres, révisions confortées également par l'observation des référents. On peut citer par exemple :

| Cadastre | | Forme corse | | Étymon |
|---------------------------|----|----------------------|---|-------------------------------------|
| (g) <i>La Mosa</i> | vs | <i>Lamosa</i> | < | LAMA REW 4862 'marais' ou 'ronce' ; |
| (h) <i>Areto</i> | vs | <i>Laretu</i> | < | LAURUS REW 4943 'laurier' ; |
| (i) <i>Campo ai Piedi</i> | vs | <i>Campu li Peri</i> | < | PĪRUS REW 6525 'poirier' ; |
| (j) <i>Acqua Degna</i> | vs | <i>Acque Tigne</i> | < | TĪNEA REW 'misérable' ; |
| (k) <i>Pentanelli</i> | vs | <i>Pantanelli</i> | < | *PALTA REW 6177 'vase, marais'. |

Concernant la dimension diachronique, quelques faits peuvent déjà être mentionnés, telle l'observation de l'extension géographique du type lexical *valdu* 'forêt' aujourd'hui employé dans une aire géolinguistique restreinte du centre ouest (BDLC) alors que les attestations toponymiques montrent sa présence sur tout le territoire

⁴ Dans cette perspective voir Pfister (1999) et des applications pour la Corse notamment in Medori (2013).

⁵ La Corse compte 366 communes.

insulaire, ce qui est d'ailleurs démontré par la documentation médiévale⁶. On peut observer aussi la substitution d'une forme lexicale à une autre, la forme ancienne étant figée par les noms de lieux. C'est le cas de *Salettu* < SALĪCTUM REW 7534 'sau-laie', dont P. Aebischer (1963, pp. 169-181) a pu démontrer qu'il s'est éteint au Moyen-Âge, spécialement en Toscane. On peut évoquer également le témoignage de l'usage passé de **àciaru* < lat. **ACER* (LEI s.v.) 'érable sycomore' qui a disparu de la pratique linguistique contemporaine. Enfin, du point de vue du signifié de l'appellatif originel, les données toponymiques suggèrent la reconstruction de certains signifiés comme *'source' pour toponymes de type *Polla* (signifié attesté en Italie, DELI s.v. *pollare*) ou *Prunu*, les substantifs corses *polla* (avec dérivés) et *prunu* connaissant aujourd'hui d'autres signifiés.

4. Conclusion

La base de données du CESIT Corsica répond à un besoin scientifique et présente un potentiel important en terme de transfert vers la société civile, puisqu'elle est amenée à être utilisée dans le cadre de révisions cartographiques, de signalétique, d'adressage mais aussi, par les analyses qui devraient se développer prochainement, en terme d'outil pour la connaissance du territoire (géographie historique, urbanisme). Sur le plan scientifique, l'absence de données fiables était jusqu'alors préjudiciable aux études toponymiques relatives à l'île. Le CESIT Corsica espère, par la mise à disposition du produit de ses enquêtes, contribuer à la connaissance du trésor onomastique de l'île ainsi qu'à sa conservation – restitution, la démarche du CESIT répondant en tout point aux recommandations de l'UNESCO pour la sauvegarde des noms géographiques.

CESIT Corsica

Stella RETALI-MEDORI
 Francescu Maria LUNESCHI

Références bibliographiques

- Aebischer, Paul, 1963. *Miscelánea Paul Aebischer*, Barcelona, Instituto internacional de cultura románica, Biblioteca filológica-histórica 9.
- BDLC = *Banque de Données Langue Corse*, éd. par Marie José Dalbera-Stefanaggi *et al.* <bdlc.univ-corse.fr>.
- Dalbera-Stefanaggi, Marie-José / Retali-Medori, Stella (ed.), 2013. *Castagni è puddoni, le lexique corse de la castanéiculture*.
- DELI = Cortelazzo, Manlio / Zolli, Paolo, 1999. *Dizionario Etimologico della Lingua Italiana*, Bologna, Zanichelli.
- Pfister, Max, 1999, « L'importanza della toponomastica per la storia della lingua nella Galloromania e nell'Italoromania », *RION* V/2, 449-464.

⁶ Voir Dalbera-Stefanaggi / Retali-Medori (2013).

Retali-Medori, Stella, 2009. «Toponymie corse: études et matériaux», *RION* XV/1, 71-88.

Retali-Medori, Stella, 2013. «Éléments gallo-italiens et gallo-romans dans les parlers corses»,
RLiR 77, 121-138.