

Kartu-Verbs : un système d'informations logiques de formes verbales fléchies pour contourner les problèmes de lemmatisation des verbes géorgiens

Mireille Ducassé*

*IRISA-INSA Rennes
mireille.ducasse@irisa.fr,

Résumé. La langue géorgienne possède un système verbal complexe, à la fois agglutinant et flexionnel, avec de nombreuses irrégularités. Les formes fléchies d'un verbe peuvent être très différentes les unes des autres. Il faut une bonne connaissance de la grammaire géorgienne pour remonter à l'infinitif (le lemme d'accès des dictionnaires le plus fréquent). L'accès aux dictionnaires pour les débutants est, de ce fait, très difficile. De plus, il n'y a pas de consensus parmi les lexicographes du Géorgien sur les lemmes qui représentent un verbe dans les dictionnaires, ce qui complexifie encore davantage les accès. Nous proposons Kartu-Verbs, une base de formes fléchies de verbes géorgiens accessible par un système d'informations logiques. Cette démonstration¹ montre comment, à partir de n'importe quelle forme fléchie, on peut trouver le lemme pertinent pour accéder à n'importe quel dictionnaire. Kartu-Verbs peut, ainsi, être utilisé comme une interface aux dictionnaires géorgiens.

1 Introduction

Le Géorgien est une langue caucasienne, langue maternelle d'environ 5 millions de personnes. Il a son propre alphabet. La grammaire géorgienne a un système verbal complexe. Certains problèmes sont illustrés ci-dessous du point de vue d'un débutant (pour plus de détails, voir par exemple Anderson (1984); Tuite (1998); Assatiani et Malherbe (2011); Gérardin (2016)). Il existe de nombreux verbes irréguliers et la langue est à la fois agglutinante et flexionnelle. La conjugaison peut modifier toute partie de la forme des verbes. Par exemple, le verbe « travailler » (mushaoba - მუშაობა), à la première personne du pluriel du présent donne « vmushaobt » (ვმუშაობთ). Notez le « v » au début du verbe pour marquer la première personne et le « t » à la fin pour marquer le pluriel. La première personne du singulier du futur donne « vimushaveb » (ვიმუშავებ). Un « i » a été inséré après le marqueur « v » de la

1. Cette démonstration complète et reprend des éléments de l'article suivant. Ducassé, M. (2020). Kartu-verbs : A semantic web base of inflected Georgian verb forms to bypass Georgian verb lemmatization issues. In Z. Gavriilidou, M. Mitsiaki, et A. Fliatouras (Eds.), Proceedings of XIX EURALEX Congress