

Repli d'une langue vers ou autre, ou demander son chemin à Helsinki

Stéphane Bortzmeyer
<stephane+blog@bortzmeyer.org>

Première rédaction de cet article le 12 janvier 2009

<https://www.bortzmeyer.org/demander-son-chemin-a-helsinki.html>

Chaque fois que quelqu'un qui n'a pas encore étudié la question se penche sur le choix des langues pour un logiciel, en fonction des demandes d'un utilisateur, il se dit « Mais pourquoi, lorsque la langue demandée n'est pas disponible, ne pas lui en fournir une autre, assez proche? ». Pour comprendre pourquoi c'est une idée difficile à appliquer, prenons le cas d'un étranger qui demande son chemin à Helsinki.

Si l'étranger parle finnois, pas de problème, il peut utiliser cette langue, officielle et majoritaire en Finlande. Et sinon? Selon la vision qu'on adopte, la meilleure seconde langue pourra être :

- Le suédois, dira celui qui a une vision politique ou administrative des langues. C'est la seconde langue officielle du pays (et un des finlandais les plus célèbres, Linus Torvalds l'a comme langue maternelle <<http://www.catb.org/~esr/faqs/linus/>>).
 - L'estonien, dira le linguiste, c'est la langue la plus proche du finnois, au sein de cette famille des langues finno-ougriennes (le hongrois en fait également partie mais il est bien plus éloigné).
 - l'anglais (ou peut-être l'allemand ou même le russe), dira le pragmatique. C'est ce qui multiplie les chances de se faire comprendre, en pratique.
- Toutes ces visions sont correctes, bien que parfois peu réalistes. Toutes sont implémentables (des bases comme CLDR <<http://www.unicode.org/cldr/>> nous donnent les langues officielles et des normes comme ISO 639-5 <<https://www.bortzmeyer.org/iso-639-5.html>> nous donnent une partie du classement linguistique). Mais si on a une vision pragmatique, il faut se faire à l'idée que le choix est complexe et ne peut pas être résolu par des algorithmes simples comme ceux du RFC 4647¹, qui se limite aux cas où on distribue une autre variante de la même langue, un problème un peu plus simple.

Ainsi, un site Web où l'utilisateur, via la configuration de son navigateur, demande des pages en breton choisira plutôt de lui servir du français puisque tous les locuteurs du breton, aujourd'hui, parlent français. Mais ce genre d'information ne peut pas être déduite algorithmiquement, elle nécessite une base assez riche et toujours changeante. Question subsidiaire : quelle base publique existant stocke cette information sur le breton?

(On la trouve indirectement dans CLDR, qui note le breton comme parlé en France et dans aucun autre pays ou bien sur Ethnologue.com <http://www.ethnologue.com/show_language.asp?code=bre> mais qui brouille l'information en notant que le breton est également parlé aux USA.)

1. Pour voir le RFC de numéro NNN, <https://www.ietf.org/rfc/rfcNNN.txt>, par exemple <https://www.ietf.org/rfc/rfc4647.txt>