The context of use of e-dictionaries for the minority languages of Italy (case study)

Carlo Zoli[1], Silvia Randaccio[2]

[1,2] Smallcodes, Firenze
E-mail: carlo.zoli | silvia.randaccio @smallcodes.com

1. Written, or on their way to written use, minority languages of Italy

Italian law 482/99 recognizes 12 minority languages: "the language of the Albanian, Catalan, Germanic, Greek, Slovenian and Croatian *populations* and *those speaking* French, Franco-Provençal, Friulian, Ladin, Occitan and Sardinian". We will briefly describe the current situation regarding online lexicographical projects - when they are not occasional - limiting to those with (at least some) scientific ambition. We will also analyze the most common claims from users and the lessons we are learning from our privileged point of view, i.e. the point of view of the developers of the lexicography platform shared by many of these e-dictionares.

We will list briefly: URL | technological partner.

1.1 Francoprovençal

1.1.1. Piedmont varieties: (so called "Genre" ortography):

http://tresorfrp.chambradoc.it | SMALLCODES

2.1.2. Val d'Aosta varieties: ("Brel" orthography)

http://patoisvda.org/it/index.cfm/ | local agency

1.2. Occitan

1.2.1 Common standard

http://diccionarioc.chambradoc.it | SMALLCODES

1.2.2. Brigasco sub-standard

(work in progress) | SMALLCODES

1.2.3. Guardia Piemontese sub-standard

(work in progress) | SMALLCODES

1.3. Friulian

1.3.1 Official standard

http://www.arlef.it/struments/grant-dizionari-talian-furlan | Serling

1.4. Ladin

1.4.1. Common (Dolomitan) standard

http://vll.smallcodes.it | SMALLCODES

http://dls.ladintal.it | SMALLCODES

1.4.2. Gherdëina sub-standard

http://vll.smallcodes.it | SMALLCODES

1.4.3 Badia sub-standard

http://moling.ladinternet.it | SMALLCODES

1.4.4 Fascian sub-standard

http://dilf2.ladinternal.it | SMALLCODES

1.4.5. Anpezo sub-standard

(work in progress) | SMALLCODES

1.4.6 Fodom sub-standard

(work in progress) | SMALLCODES

1.4.7 Collese sub-standard

(work in progress) | SMALLCODES

1.4.8 Ladino-Venetian dialectal family (Agordino, Comelico, Oltrechiusa)

http://www.istitutoladino.it/index.php | local web agency

1.6. "Germanic"

1.6.1 Cimbrian standard

http://www.zimbarbort.it | SMALLCODES

1.6.2 Mòcheno standard

http://kib.ladintal.it/ | SMALLCODES

1.6.3 Sappada standard

http://sappada.smallcodes.org | SMALLCODES

2. Contexts of use of electronic dictionaries for Regional or Minority Languages are peculiar

2.1 The pioneers, i.e. the *manufacturers* of a formal language

The first purpose of the lexicography for minority languages (at least, for those that enjoy official recognition, at different levels) is essentially a purpose of language policy; dictionaries play a key role, they are always necessary as a first step (though they are obviously not sufficient) in the process of corpus / status / acquisition planning. But, unlike what happens for the lexicography of "big" "national", "majority" languages, in the case of minority languages the user's needs focus (in the fist stage we are discussing) primarily on *status* and *corpus*, and less on *acquisition*. Essentially, users deal with an early stage of existence of the written language, and for this reason almost only language professionals (translators - especially of bureaucratic / administrative documents - and teachers) use electronic resources. These resources (dictionaries) are almost invariably bilingual, and the most widely used ones - even if this is a counter-intuitive fashion - are those **from** the majority language → **to** the minority language.

Early users of newly standardized languages, being language professionals, can give a strong feedback to lexicographers. Two are the needs that these users have and therefore two are the kind contributions and feedback that they can give to linguists and lexicographers:

A) **terminology**: users need essentially **terms** for their "official" translations: technical words for administrative life, school, textbooks, etc; Such a tool for the creation of terminology allows users to operate a fundamental procedure if we want the minority language to be employed in school teaching and administration: in fact, languages which do not have a written tradition normally lack of technical lexicon. We believe that this creation should be participative and we have therefore developed a tool for terminology making with different levels of approval of proposed

terms, in order to involve different levels of competences among users (from teachers to language experts, from administrative staff to professional translators). The normal flow is: discussion of translation proposals among power-users → pre-publication on line → final approval → *go public.* The contribution of users in this field is therefore essentially of the type "(ii) Indirect user contributions, which occur in different forms of explicit feedback (e.g., by e-mail or web forms)" (Abel and Meyer, 2013). This is because these newly standardized languages need some sort of centralised control, especially in the field of terminology.

The second area of contribution is a B) **local variants**: a newly standardized language normally comes form a chaotic past of former local / spontaneous spellings, that where often the only available spellings before the reach of (a sudden and often unexpected) officialdom. An extremely interesting case of a transformation of this previous chaos into a standardized spelling (even if contrary to the original intents of linguistic experts) is the case we are experiencing in Switzerland and in the Francoprovençal area of Piedmont; a situation that can arise, we suspect, every time the language does not have a historical prestige, or if there is a lack of will among the cultural elites for the creation of a unified written language. These languages are currently developing big dialectal dictionaries (see for examples lsi.ti-edu.ch/lsi/) that record in detail all local form of single entries, using as feedback the collection among speakers. Once there forms are collected they necessarily have to be grouped (usually sharing the meaning and, **by definition**, the etymology) into one macro-entry; in other words, lexicographers need a macro-entry that somehow encompasses, and covers, all the internal variants of that lexical type. These macro-entries, usually chosen or designed with criteria of graphic simplicity and etymological transparency, end to be chosen as referential forms for texts that want to go beyond purely local levels. This is empirically evident if one analyses written production in these languages. Somehow paradoxically, what some scholars do not want - or are unwilling - to do, i.e. to create a standard written language, it is done by users who end up electing as standard a "medium language" that they find "etymologically transparent" and "graphically simple", but that was designed for totally different purposes. Users end up making use of a *de-facto* standard language, simply choosing the macro-entries of a dictionary that was designed for merely dialectal purposes. This says a lot about the features one should have in mind when designing spellings of newly written languages!

2.2. Where I need it, when I need it

As far as "common" users of electronic dictionaries of minority languages are concerned, the experience of Smallcodes platform shows that the hurdle is not (yet) the use of vocabulary, but the very usage of the written form of the language, which is invariably seen, especially at the beginning, as awkward and unusual. As we have shown, users need and claim an approach that is totally different to the one used with traditional paper dictionaries and also to the one used with traditional web

portals or CD Roms. This approach can be synthesized with the sentence: "where I need it, when i need it." Users are increasingly starting to access dictionaries in a very direct way, because the necessity appears only when they are reading or writing a text (both in their native or foreign language).

What may seem like a small effort (open a browser or a new tab on the browser, go to the desired website, look up the word) during a dictionary search, can not only discourage the use the electronic tool, but also prevent the very possibility to write in a language in which one is not accustomed to writing. The user must be able to check the e-dictionary simply by clicking on the word while reading or writing. This awareness of needs come from our direct and constant relationship with around 25 working groups which are using our platform (es editors, as beneficiaries and both), and it is really valuable because, as stated before, our user are always language experts.

Smallcodes is answering to this requests by making its lexicography platform directly accessible in web pages, web forms, e-books and writing tools such as Open / Libre Office. Being these listed tools so various and different, the link with our platform we are going to develop must be accessible through web-service without any need for installation (apart from a small add-on). In this way, the only requirement will be an internet connection and the tool will automatically link the digital instruments in usage with the related dictionary, with a pop-up window that appears at the user's click on any word. Once the window is open, the users can view the meaning and occurrences of the word and close the pop-up window. Alternatively, they can expand the research and enter the front-end of the related online dictionary by simply clicking on any word link of the pop-up window. As seen, speakers of minority languages are those who need an authoritative dictionary more than anyone. The need for a dictionary is thus very high when they are in the act of producing or reading texts. If the tool "dictionary" is available without interruptions or obstacle, the writing or reading process can be really simplified, giving the minority language a better chance to be spread, used and promoted.