

Franco Lurà - Michele Moretti - Carlo Zoli

DALLA CARTA AL WEB: LA VERSIONE INFORMATICA DEL LESSICO DIALETTALE DELLA SVIZZERA ITALIANA

GENESI E STRUTTURA

L'intenzione di affiancare una versione informatica alla pubblicazione cartacea del 'Lessico dialettale della Svizzera italiana' (LSI), uscita in 5 volumi nel mese di dicembre del 2004, si era manifestata a più riprese già durante la gestazione dell'opera; essa veniva poi convertita in promessa formale, ancorché generica, durante la presentazione ufficiale del dicembre 2004, in occasione della quale fu parimenti assunto l'impegno di compilazione di un indice inverso italiano-dialetti. L'entusiasmo del momento aveva fatto azzardare qualche scadenza temporale, che già le prime verifiche tecniche sulla portata dell'operazione avevano identificato come eccessivamente ottimistica. Con l'idea che tale strumento dovesse consentire la perfetta collocazione di tutti i dati acquisiti e offrire un grado massimo di interrogabilità quanto alla possibilità di reperirne ogni possibile aspetto, si è presto compresa l'inadeguatezza della base informatica (*File-Maker*) sulla quale si era condotta la compilazione dell'opera originale e la necessità di un supporto programmatico di più elevata raffinatezza, quale poteva essere fornito unicamente da informatici professionisti. A partire dal 2006, dopo diversi infruttuosi tentativi, si è identificata come *partner* la ditta *Open Lab* di Firenze; ha preso in carico il progetto Carlo Zoli, responsabile della divisione *Smallcodes* specializzata nel trattamento digitale di lingue regionali e di minoranza e coautore del presente articolo. L'ingegner Zoli, insieme a vari collaboratori, ha curato l'elaborazione del progetto a partire dall'agosto 2007 e ne sta attualmente accompagnando la messa in funzione operativa. Rispetto alle prime previsioni, il compito ha presto presentato diverse difficoltà aggiuntive, dovute in particolare alla frammentata varietà del versante dialettale del repertorio e all'esigenza di elaborare un 'contenitore' informatico entro il quale andare a inserire non già dati linguistici grezzi, ma un'intera opera già rigorosamente definita e strutturata.

L'adesione al progetto *Opensource Smallcodes* ha portato in dote una piattaforma lessicografica appositamente sviluppata per la tutela integrata sul *web*

di minoranze linguistiche e culturali. Già da tempo utilizzata da diversi centri di studio linguistici, essa consiste in un applicativo completamente *web-based* sviluppato in *Java*. L'accesso all'applicazione è possibile tramite un normale *browser*. L'architettura originale è stata poi modificata più volte, e anche in misura importante, per far fronte alle esigenze di adattamento via via incontrate.

Considerati i dati di partenza del LSI (composto da circa 57000 lemmi per oltre 190.000 varianti, da circa 60.000 locuzioni polirematiche e da oltre 200.000 significati), si è dovuto dapprima perfezionarne la migrazione da *FileMaker* al-

The screenshot displays a software interface for a dialectal dictionary. At the top, a blue header reads 'LESSICO DIALETTALE DELLA SVIZZERA ITALIANA'. The main window is divided into several sections:

- Top Bar:** Includes a search bar with the lemma 'campana', a 'Def=1/af=0' field, and buttons for 'Barbara' and 'ok'.
- Left Panel:** Contains navigation and statistics: 'Usa', 'Formato:', 'Introduz:', 'Record: 1', 'Trovati: 10', 'Totale: 192418', and 'Non ordinati'.
- Main Content Area:**
 - lemma:** 'campana'.
 - varianti:** Lists regional variants such as 'campana (Biasca)', 'campane (Medeglia, Robasacco, Fescoggia, Breno)', 'campèna (Bondo)', 'campèna (Olivone, Calpiogna, SottoP.)', 'čampèna (Ludiano)', 'campèna (Gerra Gamb.)', 'čempèna (Biasca)', 'chempane (Gerra Gamb.)', and 'chempèna (Soglio)'.
 - qgr:** 's.f.' and 'rim2'.
 - signif:** A detailed list of meanings, including 'Campana; campano per il bestiame', 'campanello elettrico', 'parafumo della lampada', 'meccanismo degli orologi', and various regional uses like 'rotolare un bambino per gioco' and 'sonare le campane a distesa'.
 - locuz:** A list of idiomatic expressions such as 'A -, con l'estremità tagliente leggermente ripiegata: di succhiello (Rovio)', 'A sòn da -, a tutto spiano, senza ritegno (Russo)', 'Fá (la) / giúgá a -, far roteare un bambino per gioco', 'Fai a -, fatto e finito', 'Fá sonda ra i -, correre il rischio di morire', 'Fá sonda (a) - (a/ e) martèll, suonare le campane a martello', and 'Fá sonda - fèssa, essere malaticcio'.
 - rimfin:** '→ bordón 1, digiún 1, fascia, còrda 1, nós, sòn 1, tèsta, viola'.
- Right Panel:**
 - Stampare scheda ? (0=NO, 1=SI):** A control for printing the entry.
 - Lista argomenti:** A list of related topics, all currently showing '<Nessun accesso>'.
 - Massa:** A button labeled '<Nessun'.
 - Massa pond.:** A button labeled '<Nessun'.
 - Ricerca globale:** A button labeled '<Nessun accesso>'.

Fig. 1 - Schermata del lemma *campana* nella versione originale *file-maker*.

l'applicazione di lessicografia di *Smallcodes*, e poi migliorarne l'implementazione attraverso euristiche specifiche. Questa fase ha comportato 18 mesi di lavoro di progettazione e programmazione ed è stata accompagnata da diversi mesi di analisi e verifiche cui hanno partecipato anche due redattori del Centro di Dialettologia e di Etnografia della Svizzera italiana. Di fatto, questo duplice percorso non può ancora dirsi del tutto concluso, dandosi ancora possibilità e necessità di affinamenti, integrazioni e miglioramenti.

Il LSI, nel suo assemblaggio operativo su *FileMaker* e nella sua versione a stampa su carta, è organizzato per lemmi, a rappresentare i quali viene eletta, fra le varianti disponibili, quella che meglio risponde a determinati e rigorosi criteri di selezione; sotto di essa vengono poi elencate le diverse varianti locali. Alle varianti dotate di tratti fonetici sufficientemente connotati sono attribuite etichette geografiche di punto (i 293 comuni politici in cui era divisa la Svizzera italiana all'inizio del Novecento) o di area (le 64 giurisdizioni di circolo, distretto, regione o cantone che li raggruppavano); le varianti di più ampia diffusione e di poco significativa rilevanza distintiva si presentano invece prive di localizzazione. Ugualmente localizzati sono i significati e le locuzioni polirematiche.

A questo tipo sintetico e riassuntivo di presentazione si contrappone l'ap-

The screenshot shows the 'smallcodes dizionario' interface. At the top, there are navigation links: 'lemmi', 'impostazioni', 'correttore', and 'informazioni'. A search bar contains 'dial. Sv.It.'. The main heading is 'campana (dial. Sv.It.)'. Below this, there are sections for 'varianti locali', 's.f.', and 's.f.pl.', each containing a list of numbered entries with their respective geographical locations in parentheses.

varianti locali	varianti locali
Campana (Biasca), campane (Medeglia, Robasacco, Fescoggia, Breno), çampèna (Ludiano), campèna (Olivone, Calpiogna SottoP.), campèna (Bondo), campèna (Gerra Gamb.), çempèna (Biasca), chempane (Gerra Gamb.), chempèna (Soglio)	
s.f.	
1 Campana (dial. Sv.It.)	1 Campana (dial. Sv.It.)
2 Campano per il bestiame (Isonne, Corzoneso, Leontica, Olivone, Mendr.)	2 Campano per il bestiame (Isonne, Corzoneso, Leontica, Olivone, Mendr.)
3 Campanello elettrico, suoneria (dial. Sv.It.)	3 Campanello elettrico, suoneria (dial. Sv.It.)
4 Parte di oggetti svasata, a forma di campana (dial. Sv.It.)	4 Parte di oggetti svasata, a forma di campana (dial. Sv.It.)
5 Convessità, forma convessa delle testate dello stampo da tegoli (Malc.)	5 Convessità, forma convessa delle testate dello stampo da tegoli (Malc.)
6 Parafumo della lampada (dial. Sv.It.)	6 Parafumo della lampada (dial. Sv.It.)
7 Isolatore che sostiene i fili sui pali della linea elettrica (Castasegna)	7 Isolatore che sostiene i fili sui pali della linea elettrica (Castasegna)
8 Vasca per la nichelatura dei pezzi del meccanismo degli orologi (Arugno)	8 Vasca per la nichelatura dei pezzi del meccanismo degli orologi (Arugno)
9 Giacinto (Cimadera)	9 Giacinto (Cimadera)
10 Varietà di noce (Meride)	10 Varietà di noce (Meride)
11 Varietà di castagno primaticcio e il relativo frutto (Gerra Gamb., Sigirino)	11 Varietà di castagno primaticcio e il relativo frutto (Gerra Gamb., Sigirino)
12 Panetto di zucchero (Rovio)	12 Panetto di zucchero (Rovio)
13 Tipo di gioco infantile che consiste nel sollevarsi a vicenda tenendosi per le braccia uniti schiena contro schiena (Giubiasco, Morbio Inf.)	13 Tipo di gioco infantile che consiste nel sollevarsi a vicenda tenendosi per le braccia uniti schiena contro schiena (Giubiasco, Morbio Inf.)
14 Aitalena sospesa (Lumino)	14 Aitalena sospesa (Lumino)
15 Dente traballante (Rovio)	15 Dente traballante (Rovio)
16 Tintinnio, ronzio nelle orecchie (S. Antonio, Leontica, Olivone, Sorvico, Viganello)	16 Tintinnio, ronzio nelle orecchie (S. Antonio, Leontica, Olivone, Sorvico, Viganello)
17 Opinione, parere, punto di vista, versione (dial. Sv.It.)	17 Opinione, parere, punto di vista, versione (dial. Sv.It.)
18 Serie, ripetizione monotona di azioni, di avvenimenti (Cavergho, Lugano)	18 Serie, ripetizione monotona di azioni, di avvenimenti (Cavergho, Lugano)
19 Allievo che non capisce nulla (Tic.)	19 Allievo che non capisce nulla (Tic.)
s.f.pl.	
1 Notizie, voci, chiacchiere, dicerie (Biasca, Cavergho, Gravesano)	1 Notizie, voci, chiacchiere, dicerie (Biasca, Cavergho, Gravesano)
2 Candele di moccio (Olivone, Sigirino, Bosco Lug., Rovio)	2 Candele di moccio (Olivone, Sigirino, Bosco Lug., Rovio)

Fig. 2a - Prima parte del lemma *campana* nella nuova versione informatica: visualizzazione sinottica.

proccio 'esplosivo' della versione informatica: in questa a ogni etichetta geografica (compresa quella più generale corrispondente all'intera Svizzera italiana) viene fatta equivalere una varietà di lingua (la più generale designata come dial. Sv.It.), e ogni forma, dotata di un senso, di una locuzione o di un collegamento di rimando, costituisce lemma a sé stante.

Lo 'spezzettamento' delle schede originali e la loro atomizzazione trova comunque una rappresentazione riassuntiva del tutto affine all'impostazione che vigeva nell'opera di partenza; tale visualizzazione sinottica, detta 'anteprima', costituisce inoltre il modello della pagina che nella versione di *front end* (= di pubblica consultazione) verrà presentata all'utenza.

campana (dial. Sv.It.)

locuzioni

1A *campana* Con l'estremità tagliente leggermente ripiegata: di succhiello (Rovio)

2A *són da campana* A tutto spiano, senza ritegno (Russo)

3<>|Fá (la)/giúgá a campana Far roteare un bambino per gioco tenendolo sotto le ascelle (S. Antonio, Verscio) ◊ Dondolarsi appeso a una corda (Cavergho)

4Fai a campana Fatto e finito, di prima categoria (Montagnola)

5<>|pl. Sentii i campana Essere seminato poco profondamente: di ortaggio (Giubiasco, Brione Verz.)

6pl. Fèe nêe in pée i campana Suonare le campane a distesa (Gerra Gamb.)

7pl. Mandá in pée i campana Suonare le campane a distesa (Sornico)

8Fá soná ra/ i campana Corrire il rischio di morire (Sornico) ◊ Suscitare meraviglia, essere inaspettato (Sornico)

9pl. Sonaa campana Dire, divulgare (Roveredo Grig.)

10<>|pl. Tirá (l) campana Cercare appoggio, protezione (Lumino, Biasca, Morbio Inf.)

11<>|Costaa pisséi la còrda dala campana Essere maggiore la spesa del guadagno (Verscio, Cavigliano)

12<>|Da campana a/ e martèll Grave, tragico (Intragna) ◊ Truce, sfrontato, da fufante: dell'espressione del volto (Minusio, Certara)

13<>|Soná (a) campana (a/ e) martèll Suonare le campane a martello. (dial. Sv.It.)

14<>|campana balórda Brutta notizia (Brissago, Gandria) ◊ Diceria (Gandria)

15campana che sóna pòch bén Brutta notizia (Gandria) ◊ Diceria (Gandria)

16<>|campana da bòsc Campana che annuncia l'inizio della corsa per il possesso degli appezzamenti boschivi su cui falciare il fieno selvatico (circ. Airole)

17pl. campana dal cùgliaa Campana che suonano a mezzogiorno (Pedrinete)

18<>|campana dal dotòr Campana che si suona per comunicare alla popolazione l'arrivo in paese del medico (Cavergho, Rovio)

19Soná da campana féssa Essere malaticcio (Viganello)

20campana d'Giann l'av Campana dal suono garrulo e noioso (Cavergho)

21<>|Sunè la(n) campana (granda) Palesare fatti, divulgare notizie (SottoP.)

22<>|pl. campana da légn Strumenti di legno che si suonano durante la settimana santa in sostituzione delle campane (Ble.)

23<>|Soná (da) campana róta Circolare voci poco gradite, pettegolezzi (Verscio, Cavigliano, Intragna, Sornico)

24campana stórna Brutta notizia (Gandria) ◊ Diceria (Gandria)

25Podé véss metüü sótt a na campana da védro Avere salute fragile, cagionevole. (Davesco-Soragno)

26Podé staa sótt a na campana da védro Avere salute fragile, cagionevole. (Verscio)

rimandi

- burdún tacaa a una campana (Mendrisio)
- deegión di campann (Caviano)
- fascia da campan(a) e martèll (Comologno, Gravesano, Certara)
- gòrda di campènn (Calpiogna)
- nés a campana (Torricella-Taverne)
- vièla a campana (Sornico)
- Nu manchè ün suna da campèna (Soglio)
- Végh la tèsta quader facia a campana (S. Domenica)

ricarica

Fig. 2b - Seconda parte del lemma *campana* nella nuova versione informatica: visualizzazione sinottica.

In altre parole, è possibile considerare la versione informatica come l'insieme di 358 diversi vocabolari dialettali (corrispondenti a 293 località + 64 aree + 1 area generale), ognuno in linea di principio autonomo, strettamente collegati e consultabili anche come se fossero uno solo.

Il lemma della versione cartacea (che nella stampa era evidenziato dal neretto) mantiene tuttavia, *primus inter pares*, una funzione di rappresentanza ed è pertanto detto 'capolemma'. Ogni altra informazione linguistica, in particolare ogni variante di pronuncia, è però analogamente consegnata ai singoli lemmi locali.

Quando una stessa lingua presenta più varianti di pronuncia, una di queste viene posta in evidenza e definita come principale; l'insieme della forma principale in una località e delle sue eventuali forme alternative costituisce un'unità indivisibile. È questo insieme a possedere un apparato di significati, rimandi, locuzioni, ecc. e non le singole forme che si trovano su un piano di parità.

Le varie lingue sono disposte in una struttura gerarchica organizzata ad albero, che da una lingua primaria (detta anche 'lingua radice'), corrispondente all'intera area indagata della Svizzera italiana, dirama progressivamente fino a tutte le lingue puntuali (dette anche 'lingue foglia' in accordo con la terminologia tecnica delle strutture-dati ad albero), corrispondenti ai singoli comuni; fra questi due capi si interpongono diverse lingue areali intermedie, collocate su vari gradi (fino a 5) di inclusione progressiva. Tutto il materiale linguistico documentato (sia esso variante di pronuncia, significato, locuzione polirematica, ecc.) viene collocato in questa struttura gerarchica. Ogni livello di questa gerarchia partecipa a una catena di eredità (con un numero maggiore o minore di ascendenti a seconda del livello): in forza di questo meccanismo di eredità ogni livello raccoglie, oltre alle proprie particolarità esclusive (ad esempio sensi specifici di un lemma in una data località), anche quanto gli perviene in dote da ciascuno dei livelli sovrastanti. Sono in particolare i significati a venire ereditati per questa via, mantenendo comunque la certificazione della loro provenienza d'origine.

Chiariamo questo meccanismo di eredità tramite un esempio:

Es.1: Il capo-lemma *campana*, tra le sue varianti locali, ha il lemma *campèna* di Ludiano. Vediamo i significati di questo lemma locale:

campèna (Ludiano)

ereditati da dial. Sv.It.

- 1 *Campana*
- 2 *Campanello elettrico, suoneria*
- 3 *Parte di oggetti svasata, a forma di campana*
- 4 *Parafumo della lampada*
- 5 *Opinione, parere, punto di vista, versione*

ereditati da Tic.

- 6 *Allievo che non capisce nulla*

ereditati da Ble.

- 7 *Campano per il bestiame*

ereditati da circolo di Malvaglia

- 8 *Varietà di noce*

TIPICI DI LUDIANO

9 Serie, ripetizione monotona di azioni, di avvenimenti

“Ereditati da dial. Sv.It.” significa semplicemente che quei significati sono comuni, per le varietà locali del lemma *campana* a tutte le località della Svizzera Italiana. Similmente, “ereditati da Ble.” significa che quel significato è comune a tutte le varianti di *campana* della valle di Blenio. Come si vede, dei 9 significati del lemma *campèna* di Ludiano, solo uno è tipico della specifica località: gli altri 8 sono generalizzati ad aree più o meno vaste.

Quando per un dato punto d'inchiesta o lingua non risulta esplicitamente documentata nessuna forma specifica, questa deve venir desunta dalle attestazioni presenti nei diversi gradi della sua catena ascendente; in altre parole deve essere estrapolata da attestazioni di aree più vaste, all'interno delle quali si trova il nostro punto d'inchiesta; se tra queste si presenta un'unica forma candidata, essa viene attribuita per ipotesi certa. Se le forme candidate presenti sulla catena ascendente sono più d'una, viene selezionata e attribuita al nostro punto di inchiesta quella detta della ‘migliore ipotesi’, vale a dire la forma attestata (e, se più d'una, la principale) sul grado della catena più prossimo alla lingua in questione.

Il sistema ha la facoltà di eseguire automaticamente dei raggruppamenti riassuntivi, ad esempio quando un senso risulta condiviso da un numero sufficientemente elevato di lingue diverse ma geograficamente vicine; in questo caso esso viene trasferito a una lingua areale di livello superiore, dalla quale potrà poi irradiarsi per eredità sui rispettivi discendenti; se la lingua intermedia interessata non risulta già dotata in precedenza di una variante formale attestata o ipotetica, gliene viene attribuita una sotto forma di ‘puro tipo lessicale’, normalmente mutuato sul capolemma. In altre parole, si tratta di un meccanismo di **promozione geografica ragionata**, che informalmente abbiamo chiamato ‘algoritmo dei pesi’, che consente di attribuire ad aree **le più vaste possibili** dei fenomeni linguistici che nei materiali si trovano attestati puntualmente, in un numero **sufficiente** di punti tra loro **sufficientemente** vicini. È chiaro che questo criterio di sufficienza è stato deciso dalla redazione del LSI caso per caso e area per area, tenendo conto della geografia dialettale di ogni area e delle sue implicazioni qualitative e quantitative; ma, una volta impostato, agisce automaticamente e aiuta il redattore ad attribuire il fenomeno linguistico al giusto livello geografico. Per esempio, un'accezione di un dato lemma può essere promossa e considerata ‘pan-grigionese’ se la si trova attestata in almeno quattro comuni grigionesi situati in ognuna delle quattro aree dialettali principali dei Grigioni italiani.

In linea teorica, i fenomeni linguistici che potrebbero venire geograficamente promossi sono sia di tipo fonetico (pronunce particolari, evoluzioni caratteristiche dovute a fenomeni di fonetica storica) sia di tipo semantico (accezioni, locuzioni polirematiche). In questa realizzazione del LSI si è deciso di utilizzare l'*algoritmo dei pesi* solo per fatti semantici.

Il cosiddetto 'spezzettamento' messo in atto nella versione informatica non è unicamente funzionale all'ordinamento e alla presentazione puntuale di tutte le informazioni linguistiche riassunte nel LSI, ma permette anche ricomposizioni dei dati diverse e estremamente interessanti; in particolare il sistema è in grado di generare diversi dizionari 'di zona' (vale a dire di aree o di singole località), la cui tipologia varia a seconda della collocazione, sulla gerarchia delle lingue, dell'area o della località presa in esame.

Per la lingua radice primaria, che corrisponde all'area più vasta d'indagine (cioè l'intera Svizzera italiana), sono disponibili tre vocabolari, intesi come tre diverse 'views' ('punti di vista', nel gergo dei *data-base*) sulla banca-dati:

1a) il **vocabolario globale della Sv.It.**: raccoglie tutto il materiale linguistico attestato nell'intera area e attribuito, generalmente, alla lingua radice o, specificamente, a tutte le altre lingue figlie

campana (dial. Sv.It.)	
varianti	campana (Biasca), campane (Medeglia, Robasacco, Fescoggia, Breno), campèna (Olivone, Calpiogna, SottoP.), çampèna (Ludiano), campèna (Bondo), campène (Gerra Gamb.), çempèna (Biasca), chempàne (Gerra Gamb.), chempèna (Soglio)
locali	
sc.	<ol style="list-style-type: none"> 1 Campana (dial. Sv.It.) 2 Campano per il bestiame (Isona, Corzoneso, Leontica, Olivone, Mendr.) 3 Campanello elettrico, suoneria (dial. Sv.It.) 4 Parte di oggetti svasata, a forma di campana (dial. Sv.It.) 5 Convessità, forma convessa delle testate dello stampo da tegoli (Maic.) 6 Parafumo della lampada (dial. Sv.It.) 7 Isolatore che sostiene i fili sui pali della linea elettrica (Castasegna) 8 Vasca per la nichelatura dei pezzi del meccanismo degli orologi (Arogno) 9 Giacinto (Cimadèra) 10 Varietà di noce (Meride) 11 Varietà di castagno primaticcio e il relativo frutto (Gerra Gamb., Sigirino) 12 Panetto di zucchero (Rovio) 13 Tipo di gioco infantile che consiste nel sollevarsi a vicenda tenendosi per le braccia uniti schiena contro schiena (Giubiasco, Morbio Inf.) 14 Altalena sospesa (Lumino) 15 Dente traballante (Rovio) 16 Tintinnio, ronzio nelle orecchie (S. Antonio, Leontica, Olivone, Sorvico, Viganello) 17 Opinione, parere, punto di vista, versione (dial. Sv.It.) 18 Serie, ripetizione monotona di azioni, di avvenimenti (Caveragno, Lugano) 19 Allievo che non capisce nulla (Tic.)
scip.	<ol style="list-style-type: none"> 1 Notizie, voci, chiacchiere, dicerie (Biasca, Caveragno, Gravesano) 2 Candele di moccio (Olivone, Sigirino, Bosco Lug., Rovio)
locuzioni	<ol style="list-style-type: none"> 1 A campana Con l'estremità tagliente leggermente ripiegata: di succhiello (Rovio) 2 A sòn de campana A tutto spiano, senza ritegno (Russo) 3 <> Fà (la) l'gìgà a campana Far roteare un bambino per gioco tenendolo sotto le ascelle (S. Antonio, Verscio) ◊ Dondolarsi appeso a una corda (Caveragno) 4 Fai a campana Fatto e finito, di prima categoria (Montagnola) 5 <> pì. Sentì i campana Essere seminato poco profondamente: di ortaggio (Giubiasco, Brione Verz.) 6 Fà sonà ra/ i campana Correre il rischio di morire (Sorvico) ◊ Suscitare meraviglia, essere inaspettato (Sorvico) 7 pì. Sonà campana Dire, divulgare (Roveredo Grig.) 8 <> pì. Tirà (i) campana Cercare appoggio, protezione (Lumino, Biasca, Morbio Inf.) 9 <> Costa pissèi la còrda dela campana Essere maggiore la spesa del guadagno (Verscio, Cavigliano) 10 <> Da campana a/ e martèll Grave, tragico (Intragna) ◊ Truce, sfrontato, da furfante: dell'espressione del volto (Minusio, Certara) 11 <> Sonà (a) campana (a/ e) martèll Suonare le campane a martello. (dial. Sv.It.) 12 <> campana balòrda Brutta notizia (Brissago, Gandria) ◊ Diceria (Gandria)

Fig. 3 - Anteprima (parziale) del vocabolario globale della Sv.It. per *campana*.

1b) il **nucleo base del lessico della Sv.It.**: raccoglie tutto il materiale linguistico specificamente attribuito alla lingua radice (e quindi portato in dote, 'propagato' con termine tecnico, a tutte le altre lingue)

dal hard core - nucleo base del lessico dial. Sv.It. campana (dial. Sv.It.)

s.f.

- 1 - Campana (dial. Sv.It.)
- 2 - Campanello elettrico, suoneria (dial. Sv.It.)
- 3 - Parte di oggetti svasata, a forma di campana (dial. Sv.It.)
- 4 - Parafumo della lampada (dial. Sv.It.)
- 5 - Opinione, parere, punto di vista, versione (dial. Sv.It.)

locuzioni

- 1 - <>|Soná (a) **campane** (a' e) martèll| Suonare le campane a martello. (dial. Sv.It.)

ricostruisci l'ordine delle polirematiche ricarica

Fig. 4 - Anteprima del vocabolario 'nucleo base del lessico della Sv.It.' per *campana*.

1c) il **patrimonio lessicale condiviso della Sv.It.**: raccoglie tutto il materiale linguistico specificamente attribuito alla lingua radice, oltre ai tipi lessicali condivisi nell'intera area.

Per le **lingue areali intermedie** sono disponibili quattro vocabolari:

2a) il **vocabolario globale della regione**: raccoglie tutto il materiale linguistico attestato nell'intera regione e attribuito alla lingua regionale o a tutte le altre lingue sue figlie, oltre a quello ricevuto in dote dai suoi ascendenti. In altre parole, raccoglie il materiale attestato nella regione, nelle aree più vaste in cui la regione è collocata e nelle località interne alla regione stessa.

°campana (Ble.)

forme diffuse in tutta l'area campana (dial. Sv.It.)

varianti locali campèna (Olivone), campèna (Ludiano), chempèna (circ. Malvaglia)

s.f.

- 1 - Campana (dial. Sv.It.)
- 2 - Campanello elettrico, suoneria (dial. Sv.It.)
- 3 - Parte di oggetti svasata, a forma di campana (dial. Sv.It.)
- 4 - Parafumo della lampada (dial. Sv.It.)
- 5 - Opinione, parere, punto di vista, versione (dial. Sv.It.)
- 6 - Allievo che non capisce nulla (Tic.)

1 Campano per il bestiame (Ble.)

2 Varietà di noce (circ. Malvaglia)

3 Tintinnio, ronzio nelle orecchie (Leontica, Olivone)

4 Serie, ripetizione monotona di azioni, di avvenimenti (Ludiano)

s.f.Pl.

1 Candele di moccio (Olivone)

locuzioni

- 1 <>|Soná (a) **campane** (a' e) martèll| Suonare le campane a martello. (dial. Sv.It.)
- 2 <>|ol. **campana da legn**| Strumenti di legno che si suonano durante la settimana santa in sostituzione delle campane (Ble.)

ricarica

Fig. 5 - Anteprima del vocabolario globale della regione di Ble., cioè della valle di Blenio, per *campana*.

2b) il **vocabolario caratteristico dei dialetti della regione**: raccoglie tutto il materiale linguistico attestato nell'intera regione e attribuito alla lingua regionale o a tutte le altre lingue sue figlie. In altre parole, raccoglie il materiale attestato esplicitamente nella regione e nelle località interne alla regione stessa.

campana (Ble.)

varianti locali: campèna (Olivone), çampèna (Ludiano), chempène (circ. Malvaglia)

sl.

- 1 Campano per il bestiame (Ble.)
- 2 Varietà di noce (circ. Malvaglia)
- 3 Tintinnio, ronzio nelle orecchie (Leontica, Olivone)
- 4 Serie, ripetizione monotona di azioni, di avvenimenti (Ludiano)

sl.PL.

- 1 Candele di moccio (Olivone)

locuzioni

- 1 >|pl. *campana de légn* Strumenti di legno che si suonano durante la settimana santa in sostituzione delle campane (Ble.)

ricarica

Fig. 6 - Anteprima del vocabolario caratteristico della regione di Ble., cioè della valle di Blenio, per *campana*.

2c) il **vocabolario specifico della regione**: raccoglie tutto il materiale linguistico specificamente attribuito alla lingua regionale.

campana (Ble.)

sl.

- 1 Campano per il bestiame (Ble.)

locuzioni

- 1 >|pl. *campana de légn* Strumenti di legno che si suonano durante la settimana santa in sostituzione delle campane (Ble.)

ricarica

Fig. 7 - Anteprima del vocabolario specifico della regione di Ble., cioè della valle di Blenio, per *campana*.

campana (Ble.)

forme diffuse in tutta l'area: campana (dial. Sv.It.)

sl.

- 1 - Campana (dial. Sv.It.)
- 2 - Campanello elettrico, suoneria (dial. Sv.It.)
- 3 - Parte di oggetti svasata, a forma di campana (dial. Sv.It.)
- 4 - Parafumo della lampada (dial. Sv.It.)
- 5 - Opinione, parere, punto di vista, versione (dial. Sv.It.)
- 6 - Allievo che non capisce nulla (Tic.)

1 Campano per il bestiame (Ble.)

locuzioni

- 1 >|Soná (a) *campana* (a/e) *martèll* Suonare le campane a martello. (dial. Sv.It.)
- 2 >|pl. *campana de légn* Strumenti di legno che si suonano durante la settimana santa in sostituzione delle campane (Ble.)

ricarica

Fig. 8 - Anteprima del vocabolario di *koinè* della regione di Ble., cioè della valle di Blenio, per *campana*.

2d) il **vocabolario di *koinè***: raccoglie tutto il materiale linguistico specificamente attribuito alla lingua regionale, oltre a quello ‘ricevuto in eredità’ dalle aree più vaste in cui la regione è collocata.

Per le singole località (le cosiddette ‘lingue-foglia’ dell’albero gerarchico delle lingue) sono disponibili 2 vocabolari:

3a) il **vocabolario globale della località**: raccoglie tutto il materiale linguistico specificamente attribuito alla lingua locale, oltre a quello ‘ricevuto in eredità’ dalle aree più vaste in cui la località è collocata.

dal vocabolario globale della località Leontica: °campana

forme diffuse in tutta l'area campana (dial. Sv.It.)

s.f.

- 1 - Campana (dial. Sv.It.)
- 2 - Campanello elettrico, suoneria (dial. Sv.It.)
- 3 - Parte di oggetti svasata, a forma di campana (dial. Sv.It.)
- 4 - Parafumo della lampada (dial. Sv.It.)
- 5 - Opinione, parere, punto di vista, versione (dial. Sv.It.)
- 6 - Allievo che non capisce nulla (Tic.)

1 - Campano per il bestiame

2 - Tintinnio, ronzio nelle orecchie

locuzioni

- 1 - <>|Sonà (a) campana (a/ e) martè|| Suonare le campane a martello. (dial. Sv.It.)
- 2 - <>|La campana da legn| Strumenti di legno che si suonano durante la settimana santa in sostituzione delle campane (Ble.)

ricostruisci l'ordine delle polirematiche ricarica

Fig. 9 - Anteprima del vocabolario globale della località di Leontica per *campana*

3b) il **vocabolario specifico della località**: raccoglie tutto il materiale linguistico specificamente attribuito alla lingua locale.

dal vocabolario specifico della località Leontica: °campana (Leontica)

s.f.

- 1 - Campano per il bestiame
- 2 - Tintinnio, ronzio nelle orecchie

ricostruisci l'ordine delle polirematiche ricarica

Fig. 10 - Anteprima del vocabolario specifico della località di Leontica per *campana*

Una normalizzazione delle traduzioni, che nel sistema sono in linea di principio indipendenti dalle ‘definizioni’ delle varie accezioni, permetterà anche l’impostazione di dizionari inversi italiano-dialetti; dei nove tipi di dizionari diretti si è ritenuto di invertirne soltanto alcuni, non tanto perché l’inversione degli altri sia impossibile, quanto perché si determinerebbero delle viste sui dati che risulterebbero di scarsa utilità o di difficile comprensione per l’utente.

I dizionari inversi che ci sono sembrati interessanti sono di un unico tipo per

le lingue puntuali: “come si dice nel dialetto di X”, specularlo al vocabolario diretto globale della località, 3a); di due tipi per le lingue areali (cioè per tutta la Svizzera italiana o le sue sotto-aree): “come si dice nei dialetti dell’area X”, specularlo al vocabolario diretto globale della Sv.It., 1a), o della regione, 2a), e “come si dice nella *koinè* dell’area X” – specularlo al vocabolario diretto patrimonio lessicale condiviso, 1c), per la lingua primaria, al vocabolario diretto di *koinè*, 2d), per le lingue intermedie.

LA LUNGA VITA DIGITALE

In ogni applicazione, e soprattutto in quelle di tipo storico-linguistico, è cruciale una prospettiva di lunga disponibilità dei dati, strettamente legata ai formati digitali di memorizzazione. È esperienza ormai comune trovarsi ad avere dati illeggibili e quindi inutilizzabili soltanto 10 o 15 anni dopo la realizzazione del prodotto *software* destinato teoricamente a gestirli. La storia del trattamento digitale dei dati linguistici è ancora troppo breve perché ci si possa riferire fruttuosamente al passato, ma oggi risulta assolutamente impensabile pretendere di tramandare il sapere in formato digitale ai decenni o ai secoli futuri (ciò che per la documentazione affidata alla stampa non pone particolari problemi), senza l’adozione di accorgimenti anche molto importanti e complessi.

D’altra parte, la potenza di analisi, di ricerca e di confronto fornita dagli strumenti di trattamento automatico del linguaggio è troppo grande per pensare di potervi rinunciare.

Pertanto, a nostro parere, per affrontare questo fondamentale (e spesso trascurato) problema è necessario adottare alcuni strumenti concreti.

Sul versante dei dati è essenziale la scelta di formati aperti e ampiamente documentati; il che, purtroppo, spesso significa non commerciali, e quindi privi del supporto garantito da grandi gruppi industriali di riferimento, e in alcuni casi meno performanti.

Sfortunatamente questo punto, assolutamente necessario, non è di per sé sufficiente. Anche i dati memorizzati in formati aperti e documentati devono poi essere gestiti tramite degli applicativi per poter essere concretamente utilizzati e analizzati. Infatti, sul versante degli applicativi, l’unica chance concreta è quella della continua manutenzione (la cosiddetta ‘manutenzione evolutiva’). In parole semplici, un applicativo *software* di livello industriale-professionale vive (e quindi scrive e legge i dati che ci interessano) unicamente se è costantemente messo a punto, non solo quanto a funzionalità, ma soprattutto quanto a compatibilità con gli ambienti tecnologici e con i sistemi operativi, essi pure soggetti a incessante e rapidissima evoluzione.

I costi di manutenzione, relativi cioè all’adeguamento continuo ai nuovi standard, vanno così considerati di gran lunga superiori a quelli di avvio ini-

ziale e di sviluppo del sistema; anzi, asintoticamente, questi costi iniziali tendono a ridursi proporzionalmente fino a diventare praticamente trascurabili.

Date tali premesse, risulta chiaro che questa manutenzione continua, questo accompagnamento costante del sistema verso il futuro, richiede uno sforzo che può essere sostenuto soltanto da grandissimi gruppi industriali o da estese comunità di interesse, costituite da una massa critica di utenti che possa unire i necessari sforzi economici, progettuali e di risorse umane. Un piccolo gruppo di utenti, un singolo gruppo di ricerca risente necessariamente, dopo pochi mesi o anni, dell'affievolirsi della propria capacità di investimento e di interesse. Solo passando continuamente il testimone ad altri utenti, e delegando loro gli investimenti per la continua e necessaria evoluzione, è possibile garantire quell'adeguamento tecnologico costante che consente, in ultima analisi, la piena disponibilità nel tempo dei propri dati.

In sintesi:

a) la lunga vita digitale di un sistema software è di interesse primario in ogni settore, e in particolare in quello delle cosiddette discipline umanistiche. La durata della vita dei dati è intimamente connessa e dipendente dalla durata della vita delle applicazioni

b) la lunghezza della vita utile di un *software*, e conseguentemente dei dati da esso trattati, è direttamente proporzionale al numero di utenti che ne fanno abitualmente uso.

È ritenuto uno dei concetti fondamentali dell'ingegneria del *software* il fatto che una parte non trascurabile del lavoro di *testing* di un programma, in particolare se sviluppato nel settore di interesse della linguistica, venga fatto dagli utenti stessi (oltre e più che dai programmatori), e spesso già in piena fase operativa.

Un'affermazione del genere verrebbe considerata quantomeno bizzarra in molti altri campi dell'ingegneria: risulterebbe a dir poco imprudente delegare buona parte della verifica funzionale di un nuovo modello di automobile ai suoi primi 1000 acquirenti e sarebbe addirittura criminale far provare la stabilità di un ponte ai primi 100 treni che lo attraversano. Nel mondo del *software* questo è invece quello che avviene normalmente, per ragioni che in questa sede non è possibile viscerare; basti dire che questa modalità di testaggio permette di raggiungere un compromesso accettabile tra costi e qualità. Un test preventivo integrale ed economicamente sostenibile (considerato che tale procedura assorbe una percentuale pari al 70%-80% del lavoro totale) potrebbe applicarsi solo a sistemi che hanno prospettive di mercato con ordini di grandezza tali da situarli irrimediabilmente al di fuori della portata di tutto il mondo delle scienze umane, quindi, e ad ancor maggior ragione, della dialettologia e della ricerca linguistica.

Sono essenzialmente queste le premesse (oltre ad altre istanze, presso certi settori, di tipo anarcoide-libertario: non meno interessanti ma meno pertinenti in questa sede) che, nella ricerca di soluzioni possibili, hanno condotto al par-

digma del cosiddetto 'software libero', una branca del quale è l'*opensource*. I suoi principi essenziali, e il modello tecnico ed economico che li informa, possono essere così riassunti: chi accede al sistema lo utilizza, a titolo tendenzialmente gratuito, nello stato in cui si trova in quel momento, ma può (e moralmente deve) contribuire, secondo le sue possibilità e capacità, al suo miglioramento nelle direzioni sopra specificate; questo si perfezionerà in particolare fornendo contributi finanziari destinati a nuovi sviluppi, di cui poi beneficerà progressivamente l'intera comunità di utenti, e contributo tecnico di programmazione, secondo gli standard del progetto.

Il progetto deve in tal modo essere considerato unico, anche se in realtà si compone, e sempre di più si dovrà comporre, di diversi moduli, non necessariamente tutti attivi o tutti di uguale interesse per i vari gruppi di utenti che vi aderiscono.

Lo sviluppo di *software* di alta qualità e di alta complessità, anche non finalizzati a scopi lucrativi, rimane comunque un'attività costosa, in particolare perché richiede risorse umane altamente qualificate; la continuazione e la stabilità di qualsiasi progetto non potrebbe pertanto ragionevolmente basarsi solo sui contributi di volontari che vi parteciperebbero unicamente solo durante il loro tempo libero.

Perciò questo modello economico deve parallelamente spostare lo sforzo di reperimento di fondi dal campo puramente commerciale a quello, non meno problematico, del *fund raising* rivolto alla ricerca di partner, finanziatori, sponsor. Naturalmente questo approccio garantisce maggiore imparzialità e trasparenza e permette a chi ha le competenze e gli interessi di collaborare o anche di utilizzare il sistema (al punto di sviluppo in cui si trova) senza investimenti rilevanti o perfino a costo zero.