

CLAUDIO BECCARI & TOMMASO GORDINI

L'ARTE

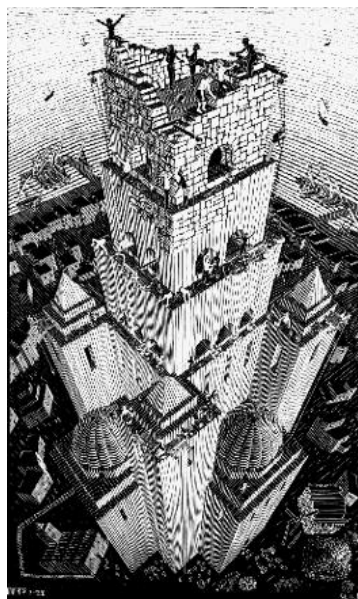
DI SCRIVERE
IN DIVERSE LINGUE CON

$(X_{\exists})^{\text{LAT}}_{\text{EX}}$



CLAUDIO BECCARI *&* TOMMASO GORDINI

*L'*ARTE DI SCRIVERE IN DIVERSE LINGUE
CON $(X_3)\text{\LaTeX}$



EDIZIONE 2015

Gruppo Utilizzatori Italiani di \TeX e \LaTeX


Claudio Beccari · Tommaso Gordini
L'arte di scrivere in diverse lingue con X_YL^AT_EX
Copyright © 2015

COLOPHON



Questo lavoro è stato realizzato su OS X usando la classe ClasseArteLingue di Claudio Beccari, una rielaborazione della mai pubblicata ClasseArteLaTeX dello stesso autore e vagamente ispirata allo stile degli *Elementi dello stile tipografico* di Robert Bringhurst.


I nomi commerciali, i loghi e i marchi registrati menzionati nella guida appartengono ai rispettivi proprietari, i pacchetti e le relative documentazioni ai rispettivi autori.


La copertina riproduce l'incisione *Tassellazione del piano con uccelli* di Maurits Cornelis Escher.

Il frontespizio riproduce la medesima incisione presente sulla copertina e la litografia *Torre di Babele*, dello stesso autore (tutte le riproduzioni delle opere di Escher contenute nella guida sono tratte da  ESCHER).

CONTATTI

 GUIT · Sito ufficiale del 

 claudio dot beccari at gmail dot com · Scrivi a Claudio Beccari

 illinguista1972 at gmail dot com · Scrivi a Tommaso Gordini

La citazione è un utile sostituto dell'arguzia.

– Oscar Wilde

Dedicato a tutti gli appassionati di \LaTeX e \XeLaTeX .

INDICE

PREFAZIONE XI

INTRODUZIONE XV

1	GESTIONE DELLE LINGUE	1
1.1	I file di descrizione di una lingua	1
1.2	Divisione in sillabe	3
1.2.1	I file dei pattern	5
1.2.2	I file per caricare i pattern	7
1.3	Le tastiere	8
2	COMPORRE IN VARIE LINGUE	13
2.1	Comporre in italiano	13
2.2	Comporre in tedesco	16
2.2.1	Con pdf \LaTeX	16
2.2.2	Con XeLaTeX	16
2.3	Comporre in latino	17
2.3.1	con pdf \LaTeX	17
2.3.2	con xelatex	18
2.4	Comporre in greco	19
2.4.1	Con pdfLaTeX	21
2.4.2	Con XeLaTeX	24
2.5	Composizione in russo	25
2.5.1	Con pdf \LaTeX	25
2.5.2	Con xelatex	26
2.6	Comporre in spagnolo e messicano	27
2.7	Comporre in portoghese e in brasiliano	27
2.8	Commenti	28
3	COMPORRE IN LINGUE ESOTICHE	29
3.1	Introduzione	29
3.2	Comporre in ebraico	29
3.3	Comporre in arabo	30
3.4	Comporre in giapponese e cinese	32
3.4.1	In orizzontale	32
3.4.2	In verticale	36
3.5	Comporre in coreano	37
4	LE TRADIZIONI TIPOGRAFICHE DI ALCUNE LINGUE	39
4.1	Tipografia italiana	41
4.2	Tipografia francese	43
4.3	Tipografia anglo-americana	44
4.4	Tipografia tedesca, austriaca, svizzera	46
4.5	Tipografia latina	48
4.6	Tipografia greca	50
4.7	Tipografia spagnola e messicana	52

4.8	Tipografia catalana	54
4.9	Tipografia portoghese e brasiliana	55
4.10	Tipografia rumena	56
4.11	Commenti	58
ELENCO DEI SITI INTERNET		61
BIBLIOGRAFIA		63
INDICE ANALITICO		67

ELENCO DELLE FIGURE

1.1	Schema della tastiera italiana sulle piattaforme Windows	8
1.2	Schema della tastiera USA	10
2.1	I font greci con codifica LGR	21
3.1	Finestra di dialogo per impostare la tastiera giapponese	34
3.2	Il menu Tastiera in cui selezionare la tastiera Hiragana	34
3.3	La parola nihongo	35

ELENCO DELLE TABELLE

1.1	Pattern usati per dividere in sillabe <i>dell'istruzione</i>	6
2.1	Traslittezzazione latino-greca secondo la codifica LGR	20
2.2	Corrispondenza fra i caratteri latini e i diacritici greci	23
2.3	Le lingue che polyglossia può gestire	28

PREFAZIONE

Questo testo è un complemento della guida *l'Arte di scrivere con L^AT_EX* che da alcuni anni rappresenta la guida per eccellenza per chi in Italia si avvicina a L^AT_EX per la prima volta. Con il passare del tempo quel manuale si è arricchito di altri contenuti, diventando un punto di riferimento anche per utenti esperti.

Crescendo in questo modo, però, *L'Arte* si è appesantita, lievitando nel numero di pagine e di informazioni al punto di allontanarsi dalla sua vocazione iniziale di guida per i principianti. Perciò ci è sembrato opportuno staccarne le parti più specialistiche e dotarle di vita propria, sottraendole ai vincoli imposti dagli autori alla guida originale e permettendo così di approfondirle con maggiore respiro. Certo: questa e altre guide che dovessero essere pubblicate in questa veste tipografica non pretendono di esaurire gli innumerevoli aspetti della composizione tipografica con L^AT_EX. Chi volesse approfondire la materia, può trovare guide generali e tematiche liberamente scaricabili nel sito del G_UIT (Gruppo utilizzatori Italiani di T_EX e L^AT_EX).

Da qualche tempo a questa parte, *L'Arte* contiene un'appendice sulla composizione di testi in latino e greco (le lingue classiche di interesse per latinisti, grecisti e linguisti in generale) scritta, di fatto, dagli stessi autori del presente manuale. I quali, complici i numerosi cambiamenti nella gestione delle lingue e la mancanza di una guida vera e propria per muovere i primi passi nel mondo di X_YL^AT_EX, hanno pensato di trattare più approfonditamente l'argomento staccandolo dal lavoro originale per non appesantirlo eccessivamente.¹

In realtà, dietro a questo nuovo manuale non c'è solo il gusto di scrivere un'altra guida: semplicemente, non sarebbe stato possibile inserire nell'*Arte* la trattazione della scrittura in alfabeti non latini per motivi tecnici insormontabili. Innanzitutto, *L'Arte* è composta con P_DF_LA^T_EX, un motore di tipocomposizione in grado di gestire solo font a 256 caratteri e dunque inadatto a comporre in alfabeti che contengono un numero superiore di simboli. In secondo luogo, gli esempi di codice sorgente per alcune lingue diventano impossibili senza i font OpenType, tipicamente codificati secondo le direttive del consorzio Unicode: il pacchetto listings, che gestisce tutti i codici presenti nell'*Arte*, è incompatibile con font di questo tipo. I font OpenType possono contenere decine e decine di alfabeti per tutti i tipi di lingue, comprese quelle con i caratteri più esotici (ideogrammi compresi) e centinaia e centinaia di simboli di vario genere. Teoricamente, ogni font potrebbe contenere alcuni milioni di caratteri, anche se nessuno dei font esistenti contiene ancora tutti i glifi che la codifica Unicode potrebbe identificare.

Ecco allora la necessità di accompagnare la pubblicazione dell'Edizione 2015 dell'*Arte* con questa nuova guida, destinata a coloro che hanno necessità di comporre testi in lingue diverse dall'italiano e dalle più note e usate lingue del mondo occidentale.

Torino-Verona, 15 dicembre 2015

Gli Autori

¹ Esiste, però, un ottimo breve documento introduttivo a X_YL^AT_EX: [Gregorio, 2011a].

RINGRAZIAMENTI

Lorenzo Pantieri, autore dell’*Arte di scrivere con L^AT_EX*, da cui la presente guida scaturisce, scriveva:

Desidero ringraziare innanzitutto i membri dello staff del G_JT e poi tutti quelli che hanno discusso con me sul forum del Gruppo, in particolare Fabiano Busdraghi, Gustavo Cevolani, Rosaria D’Addazio, Massimiliano Dominici, Gloria Faccanoni, Gianluca Gorni, Maurizio Himmelmann, Jerónimo Leal, Lapo Filippo Mori, Andrea Tonelli, Emiliano Giovanni Vavassori ed Emanuele Vicentini, per l’insostituibile aiuto fornito nella redazione di questo lavoro, le spiegazioni dettagliate, la pazienza e la precisione nei suggerimenti, le soluzioni fornite, la competenza e la disponibilità: grazie mille, ragazzi! Rivolgo un ringraziamento davvero particolare al professor Enrico Gregorio, per i suoi impagabili insegnamenti e per avermi concesso l’onore di scrivere la prefazione alla guida.

È evidente che anche noi siamo in debito nei confronti di tutte le persone citate da Lorenzo, perché senza di esse questo manuale non esisterebbe. Inoltre, noi siamo riconoscenti a Lorenzo stesso per l’identico motivo: senza l’*Arte*, questa guida non esisterebbe né avrebbe avuto senso.

Dato che copertina, frontespizio e bibliografia mantengono lo stesso stile dell’*Arte* (anche se con leggere differenze) anche noi ringraziamo Ivan Valbusa per i suoi consigli grafici e per aver creato lo stile con cui anche la bibliografia in fondo al manuale è composta.

Lorenzo Pantieri ringraziava André Miede per lo stile ClassicThesis e Daniel Gottschlag, i cui consigli gli avevano permesso di scrivere ArsClassica, il ‘vestito’ dell’*Arte*. Tuttavia, questa guida è stata composta con una classe apposita che non si appoggia più ai due stili appena menzionati sebbene, riveduti e modificati, ne conservi diversi aspetti compositivi.

Di nostra iniziativa ringraziamo i frequentatori del Forum del G_JT, che ci hanno stimolato a lavorare in questo campo fino a convincerci che i tempi per esporre il frutto dei nostri lavori in modo coerente erano maturi. Il tutto nello spirito dell’*Arte*, ai cui lettori esperti questa guida si rivolge.

Torino-Verona, 15 agosto 2015

Gli autori

INTRODUZIONE

*Abbiamo visto che la programmazione è un'arte,
perché richiede conoscenza, applicazione, abilità e ingegno,
ma soprattutto per la bellezza degli oggetti che produce.*

– Donald Ervin Knuth

Come si è detto nella Prefazione, questa guida è destinata agli utenti di `PDFLaTeX`, `XLaTeX` e `LuaLaTeX` che già conoscono nei particolari le parti rilevanti dell'Arte di scrivere con *LaTeX*, la cui lettura è propedeutica a quella di questo lavoro. Pertanto, nei capitoli seguenti non si ripeteranno le informazioni lì contenute, che si danno per conosciute.

Probabilmente, i lettori si sono già fatti una buona esperienza con il programma di composizione `PDFLaTeX`, ma molto verosimilmente non hanno grande dimestichezza con `XLaTeX` e `LuaLaTeX`. Non se ne sentano intimiditi: il markup di `LaTeX` è sempre lo stesso, quindi un file sorgente che viene composto correttamente con `PDFLaTeX` quasi certamente lo sarà anche con `XLaTeX`, a meno di non usare quei pochi pacchetti e comandi specifici del primo. Le eventuali modifiche da apportare al codice per renderlo compatibile con `XLaTeX` sono di solito poche cose sistemabili nei proverbiali 'due minuti'. La stessa compatibilità si riscontra tra `XLaTeX` e `LuaLaTeX`: quest'ultimo presenta in più la particolarità che permette di usare il linguaggio di scripting Lua, ma l'utente che non ne deve sfruttare obbligatoriamente le funzionalità probabilmente si troverà meglio con `XLaTeX`. Perciò, quello che in questa guida si dirà a proposito di `XLaTeX` vale anche per `LuaLaTeX`; le poche differenze saranno menzionate esplicitamente dove serve.

Le differenze più marcate fra i due programmi di composizione `PDFLaTeX` e `XLaTeX` sono le seguenti.

1. `PDFLaTeX` è in grado di gestire solo polizze di caratteri (comunemente chiamate con il nome collettivo di font) contenenti al massimo 256 segni (tra alfabetici, non alfabetici, interpuntivi, cifre e simboli di altro genere). `XLaTeX`, invece, può gestire i font OpenType e TrueType, non importa che siano quelli contenuti in ogni distribuzione completa del sistema `TeX` o quelli costituenti la dotazione di caratteri propria del sistema operativo installato nel proprio computer o, infine, quelli scaricabili dalla Rete o acquistati dall'utente. I font OpenType possono contenere un numero enorme di segni, che si rivelano particolarmente utili nella composizione di testi scritti (totalmente o solo in parte) in alfabeti diversi da quello 'latino esteso'.
2. `PDFLaTeX` può sfruttare pienamente le funzionalità del pacchetto `microtype`, grazie alle quali si possono raggiungere livelli di perfezione nella composizione tipografica difficilmente ottenibili con altri programmi. In sostanza, esse consistono nelle seguenti due caratteristiche.

PROTRUSIONE La protrusione consiste nel lasciare sporgere dai margini destro e sinistro della gabbia del testo alcuni piccoli segni otticamente poco ingombranti come il punto e la virgola, o parti di

segni come le grazie delle aste dei caratteri ‘graziati’ (cioè dotati di grazie). In questo modo, le righe di testo vengono riempite ancora meglio di come farebbe di solito qualunque programma di composizione del sistema \TeX , con notevoli vantaggi anche per l’occhio del lettore: il ‘grigio’ del capoverso composto appare più omogeneo e, in linea di massima, si evitano anche i piccoli ‘ruscelli’ bianchi che talvolta serpeggiano fra righe di testo consecutive.

ESPANSIONE L’espansione consiste nell’allargare o nel contrarre ogni riga di caratteri di una certa percentuale rispetto alla sua larghezza naturale, in modo da agevolare la giustificazione di un intero capoverso. Questa percentuale deve essere davvero minima, in modo che l’occhio non se ne accorga.

Ora, \XeTeX non è completamente compatibile con microtype, perché può sfruttare solo la prima delle due caratteristiche appena descritte. Si tratta di perfezionismi, a ben vedere, che però a lungo andare creano dipendenza (positiva!) e si finisce per rammaricarsi di non poterli applicare sempre. \LuaTeX , invece, non presenta questi limiti.

La possibilità di usare i font OpenType, al contrario, è molto più importante, tanto da risultare irrinunciabile per comporre in molte lingue, composte in caratteri sia latini sia non latini.

Nei capitoli che seguono verranno esposti più o meno approfonditamente i seguenti concetti.

IL PRIMO CAPITOLO descrive come gestire la composizione in diverse lingue sia con \PDFTeX sia con \XeTeX .

IL SECONDO CAPITOLO descrive come comporre in diverse lingue usando sia \PDFTeX sia \XeTeX ; non si limita alla scrittura con l’alfabeto latino esteso, ma descrive anche come comporre in greco e in cirillico.

IL TERZO CAPITOLO spiega come si gestiscono alcune lingue ‘esotiche’ che si scrivono da destra a sinistra o dall’alto al basso, con alfabeti non latini o con sistemi di ideogrammi.

IL QUARTO CAPITOLO descrive molto sommariamente le tradizioni tipografiche di alcuni paesi, dalle virgolette alle enumerazioni; dalla spaziatura della punteggiatura, alla composizione dei numeri ordinali; eccetera. Non verranno discussi tutti questi dettagli per ciascuna tradizione tipografica ma si cercherà di mettere in evidenza alcuni punti salienti.

In tutti i casi, le indicazioni verranno date assumendo l’uso di \XeTeX , ma non si mancherà di evidenziare le peculiarità che lo distinguono da \PDFTeX o \LuaTeX .

Anche per questa guida vale quanto Lorenzo Pantieri scrive nell’Introduzione all’*Arte*: non state leggendo un manuale su \XeTeX , ma solo una sistemazione più o meno ordinata dell’esperienza che ci siamo fatti in questo campo. Di noi due, CB è un ingegnere che si interessa da molto tempo ai programmi del sistema \TeX , in particolare per quanto riguarda la gestione delle lingue, specie alcune di quelle che si scrivono con alfabeti non latini. TG, invece, è un insegnante che ha accumulato non poca esperienza producendo materiale scritto per i propri allievi e rivedendo nella forma e nel contenuto le bozze di alcune delle guide disponibili nel sito del \G\IT .

1

GESTIONE DELLE LINGUE

I programmi di composizione $\text{PDF}\text{L}\text{A}\text{T}\text{E}\text{X}$, $\text{X}\text{E}\text{L}\text{A}\text{T}\text{E}\text{X}$ e $\text{L}\text{u}\text{a}\text{L}\text{A}\text{T}\text{E}\text{X}$ gestiscono la composizione in lingue diverse dall'inglese (nella variante americana, che è l'impostazione predefinita in mancanza di altre indicazioni) mediante i due strumenti seguenti.

1. Un pacchetto che permette di specificare la lingua principale e dominante del documento in lavorazione e le eventuali altre lingue subalterne a essa.
2. La definizione delle regole di cesura in fine di riga che, con i primi due programmi sopra elencati, devono essere precaricate nei rispettivi file di formato.

Il primo compito viene svolto dai pacchetti `babel` e `polyglossia`. Sebbene sia predisposto per lavorare prevalentemente con $\text{PDF}\text{L}\text{A}\text{T}\text{E}\text{X}$ (infatti richiede praticamente solo font a 256 caratteri), `babel` funziona anche con $\text{X}\text{E}\text{L}\text{A}\text{T}\text{E}\text{X}$, ma a patto di non caricare font OpenType. Il pacchetto `polyglossia`, invece, può lavorare solo con $\text{X}\text{E}\text{L}\text{A}\text{T}\text{E}\text{X}$ e $\text{L}\text{u}\text{a}\text{L}\text{A}\text{T}\text{E}\text{X}$ e, oseremmo dire, solo con font codificati in Unicode a cui si accede con la transcodifica UTF-8. Poiché questa è la codifica predefinita per i due programmi appena nominati, *non* serve specificarla nel preambolo del documento come opzione al pacchetto `inputenc` (che, dunque, *non* va caricato). Essenziale, invece, è assicurarsi di lavorare con un editor in grado di salvare i file sorgente nella codifica UTF-8.

Il secondo compito è molto più complesso e non è alla portata di tutti. L'utente, perciò, deve fidarsi del lavoro degli specialisti in grado di scrivere i moduli contenenti le regole di cesura in fine di riga nel linguaggio nativo di TEX .

Oggi, a metà 2015, i programmi del sistema TEX sono predisposti per comporre in un'ottantina di lingue ed è molto difficile che fra di esse manchi proprio quella con cui si vuole scrivere. Nel caso in cui si trovassero delle cesure errate, tuttavia, non bisogna perdersi d'animo. L'inconveniente può essere superato con vari strumenti alla portata dell'utente, ma non bisogna mai dimenticare di informare il curatore dei moduli di cesura dell'errore rinvenuto mediante una semplice email all'indirizzo elettronico normalmente contenuto nei file di *pattern* (che verranno descritti sommariamente più avanti). Prima di farlo, a ogni modo, si consiglia di documentarsi accuratamente sulle regole di sillabazione della lingua in questione su una grammatica o su un dizionario che riporti di ogni lemma la divisione corretta: è sempre meglio non essere sicuri che una cosa che sembra un errore lo sia davvero.

1.1 I FILE DI DESCRIZIONE DI UNA LINGUA

Ogni lingua è descritta mediante un file che per `babel` presenta la seguente forma generale:

```
<nome inglese della lingua in tutte minuscole>.ldf
```

e per polyglossia quest'altra:

```
gloss-<nome inglese della lingua in tutte minuscole>.ldf
```

Talvolta un solo file permette di chiamare una stessa lingua con nomi diversi oppure usa nomi diversi per specificare caratteristiche diverse della lingua.

Ora si vedrà nel dettaglio come impostare lingue e varietà di lingua con i due pacchetti in questione.

Con polyglossia, il comando che specifica una lingua da usare nel documento accetta un argomento facoltativo nel quale se ne possono indicare le varianti; l'elenco completo di lingue, varianti e opzioni è contenuto nella documentazione del pacchetto; la tabella 2.3 a pagina 28 indica solo le lingue gestibili da polyglossia, ma per i dettagli di ciascuna bisogna consultare la suddetta documentazione.

Per esempio, per un documento in greco si può scrivere:

```
\setmainlanguage[variant=<variante>]{greek}
```

dove *<variante>* può assumere i seguenti valori:

- mono per selezionare la variante moderna monotonica;
- poly per la variante moderna politonica; oppure
- ancient per la variante antica politonica.

Tutte e tre influenzano sia l'ortografia delle parole fisse (come *Ἀναφορὴς* in greco monotónico a fronte di *Ἀναφορῆς* in greco moderno politónico) sia la scelta delle regole di cesura in fine di riga.

Con babel la situazione è simile, tranne che per gli aspetti seguenti.

1. Si possono scegliere la lingua o le lingue in cui comporre il documento come opzioni da passare alla dichiarazione di classe o al comando che carica babel. Le due scritture seguenti sono equivalenti:

```
\documentclass[... ,english,...]{book}
\usepackage[...]{babel}
```

oppure

```
\documentclass[...]{book}
\usepackage[... ,english,...]{babel}
```

2. Le eventuali varianti si possono esprimere come modificatori delle lingue passate a babel con la sintassi *<lingua>.<variante>*:

```
\usepackage[latin.classic,greek,...]{babel}
```

o con comandi a parte:

```
\usepackage[latin.classic,greek,...]{babel}
\setlanguageattribute{greek}{polutoniko}
```

3. La lingua principale del documento è quella elencata per ultima o quella espressa come valore nell'opzione *main=<lingua>*, la cui posizione è libera. Le due scritture seguenti sono equivalenti:

```
\usepackage[main=italian,latin.classic,greek]{babel}
\usepackage[latin.classic,greek,italian]{babel}
```

Una volta specificate le lingue da usare, quella principale viene impostata automaticamente come predefinita, ma nulla vieta di comporre brani più o meno lunghi in un'altra lingua. A questo scopo l'utente ha a disposizione i seguenti strumenti, identici per i due pacchetti. Si tratta di due comandi:

```
\selectlanguage{⟨lingua⟩}
⟨testo⟩
```

e

```
\foreignlanguage{⟨lingua⟩}{⟨testo⟩}
```

e di un ambiente:

```
\begin{otherlanguage}{⟨lingua⟩}
⟨testo⟩
\end{otherlanguage}
```

Si noti quanto segue.

1. Il comando `\selectlanguage` e l'ambiente `otherlanguage` cambiano tutte le impostazioni del documento in funzione della nuova *⟨lingua⟩* temporaneamente selezionata, comprese le parole fisse (per esempio, a seconda dei casi *Capitolo* diventa *Chapter*, *Chapitre*, *Κεφάλαιον*, eccetera).
2. Dell'ambiente `otherlanguage` esiste la versione asteriscata `otherlanguage*` che si limita a cambiare font, se necessario, e ad attivare la sillabazione per la *⟨lingua⟩* indicata.

Prima di terminare questo paragrafo, è opportuno segnalare ancora che le regole di cesura in fine di riga *non vengono impostate* da polyglossia o da babel, ma *vengono solo selezionate* fra quelle già caricate nel file di formato. Quindi, se il formato non le contiene perché si sta lavorando con una distribuzione obsoleta o parziale del sistema T_EX, si sappia che la composizione del documento verrà eseguita con regole di sillabazione 'sostitutive', che di solito sono quelle del predefinito inglese americano. Non è bello comporre in italiano sillabando le parole in inglese: è, anzi, veramente comico!

Se invece un forestierismo entrato nell'uso corrente risultasse diviso male, lo si può scrivere correttamente sillabato nell'argomento del comando `\hyphenation` e alla successiva compilazione il problema non si presenterà più. Per esempio, in un contesto italiano la parola *download* viene divisa in *down-load*. Scrivendo nel preambolo del documento `\hyphenation{down-load}`, invece, se ne assicurerà la corretta sillabazione a ogni occorrenza verso la fine di una riga.

1.2 DIVISIONE IN SILLABE

Nella composizione tipografica, la divisione in sillabe delle parole per la cesura in fine di riga è un'operazione importante. Va detto che in tipografia si riesce a comporre *a pacchetto* (cioè con entrambi i margini perfettamente giustificati) molto meglio che a mano, anche se gli amanuensi che precedettero l'introduzione della tipografia meccanica erano bravissimi a farlo anche scrivendo con la penna d'oca.

Oggi esistono stili diversi di composizione, e con quello detto *in bandiera* (con un solo margine giustificato) non è quasi mai necessario ricorrere alla cesura di una parola. Tuttavia, questo tipo di composizione, rarissimo nei libri, è più adatto per stampati informali, depliant pubblicitari e simili.

Per comporre a pacchetto, quindi, bisogna conoscere le regole *grammaticali* di divisione in sillabe, da distinguere dalle regole *tipografiche* di cesura in fine di riga. Non è una cosa strana: la grammatica si basa su regole astratte valide in ogni caso e indipendenti da che cosa e come si sta scrivendo; la tipografia, invece, deve fare i conti con la giustezza del blocco di testo, con i font usati, eccetera, ma più di tutto con la comodità di lettura per il lettore. Ovviamente entrano in gioco anche criteri estetici: i word processor, se non adeguatamente configurati, compongono senza eseguire la cesura, il che spesso comporta la formazione di spazi interparola eccessivi e molto antiestetici.

La tipografia, dunque, deve evitare questi inestetismi e avere regole che obbediscano alla grammatica da un lato, ma che permettano anche di conservare la bellezza dello stampato professionale dall'altro. Ecco perché le regole tipografiche della cesura in fine di riga sono più rigide di quelle grammaticali. È bene ricordarsene per non doversi stupire di certe scelte dei programmi di composizione del sistema \TeX , per noi strane solo in apparenza.

Di seguito si riassumono le regole tipografiche valide per tutte le lingue che si scrivono con alfabeti (sono esclusi, perciò, gli ideogrammi).

1. Ogni sillaba contiene o una vocale o un gruppo di vocali che formino un unico suono, anche modulato; di solito questi gruppi si chiamano dittonghi, trittonghi e simili.
2. Un gruppo di consonanti a inizio parola non si stacca dalla vocale successiva.
3. Un gruppo di consonanti a fine parola non si stacca dalla vocale precedente.
4. Un gruppo di consonanti che formi un unico suono non si divide fra le consonanti che lo formano; di solito questi gruppi di consonanti si chiamano digrafi, trigrafi, eccetera.
5. Le consonanti modificate con segni diacritici si trattano come tutte le altre consonanti.
6. Le vocali modificate con segni diacritici normalmente sono escluse dalla formazione dei gruppi inseparabili di vocali; dunque, greco a parte, dove alcune vocali formano dittongo con quella che le precede, non formano dittonghi e simili. Solo la dieresi può spezzare questo legame.

Ogni lingua possiede le proprie eccezioni e classifica le lettere dell'alfabeto in modo diverso da quello di altre lingue. Per esempio, nelle grammatiche di croato la *r* può essere classificata come vocale: si pensi a parole come *smrt* o *Trst* (nelle quali l'accento insiste proprio sulla *r*) che a noi sembrano impronunciabili perché prive di vocali. In italiano esiste la *s* scempia (o 'impura'), cioè una *s* seguita da consonante diversa, dalla quale normalmente non si divide anche se non forma un digrafo o simili. Poi esistono le eccezioni delle eccezioni, per cui due *s* consecutive seguite da una consonante diversa appartengono entrambe alla sillaba precedente la consonante, contraddicendo la regola precedente (come in *massmediologo*, da sillabare *mass-me-dio-lo-go*). Inoltre, la tradizione tipografica italiana non divide nemmeno gli iati, come in *croa-to*, ma permette di separare gruppi di tre o quattro vocali in base alla presenza di dittonghi o trittonghi al loro interno, come in *pa-io*, *ma-ieu-ti-ca*, *gua-io*.

In inglese in generale, e nella varietà americana in particolare, di solito la vocale tonica non si separa dalla consonante seguente, la quale, a propria volta, va separata da quanto segue: perciò si ha *liq-uid*, che a noi sembra assolutamente errato. Peggio ancora: una parola come *record* si pronuncia diversamente a seconda che sia un sostantivo (*rècord*, da sillabare *rec-ord*) o un verbo (*recòrd*, da sillabare *re-cord*). Non esistendo gli accenti in inglese, le due parole risultano omografe e nessun programma di sillabazione riesce a distinguerle.

Non è un caso che il sistema T_EX non commetta errori nel sillabare parole strettamente italiane (ma anche molte parole di origine straniera come *leishmaniosi*, *maxwelliano*, *newyorkese*, *taylorismo*, *wagneriano*, eccetera), mentre in inglese le parole divise in modo scorretto sono frequenti (inizialmente erano circa il 10%). Per fortuna, nel tempo si sono raccolte moltissime eccezioni e le si sono già memorizzate nel file di formato, per cui molti errori oggi non si manifestano più. Tuttavia, le parole omografe ma non omofone presentano ancora gli inconvenienti descritti sopra.

Altri casi di sillabazione difficile se non impossibile sono quelli basati sulla divisione etimologica: in italiano è permessa, anche se sempre più rara, ma in altre lingue è assolutamente obbligatoria. Per esempio in tedesco, dove le parole composte sono innumerevoli e formano stringhe di lettere talvolta molto lunghe. Non è quindi un caso se le complicate regole di sillabazione del tedesco richiedono file contenenti migliaia di regole, a fronte delle poche centinaia contenute nei file per l'italiano.

1.2.1 I FILE DEI PATTERN

Per eseguire una cesura in fine di riga, il sistema T_EX segue il seguente ordine.

1. Nel file di descrizione di una lingua si trova specificato il numero minimo di lettere di cui possono comporsi la sillaba iniziale e quella finale di una parola. La cesura tipografica *non* viene eseguita se la prima o l'ultima sillaba grammaticale risultano più brevi dei seguenti valori:
 - 2 e 3 (impostazione predefinita);
 - 2 e 2 (per italiano e numerose altre lingue neolatine);
 - 1 e 1 (per il greco).

Nonostante che la sillabazione tipografica per l'italiano sia semplice, con i due valori appena visti una parola come *idea* non può essere sillabata, sebbene la grammatica dica che la sua divisione sia *i-de-a*. In greco antico la stessa parola sarebbe *ιδέα* e i valori indicati dicono che la si può dividere tranquillamente in *ι-δέ-α*, sia tipograficamente sia grammaticalmente.

2. Viene poi esaminata la lista delle eccezioni, sia quelle precaricate nel file di formato, sia quelle aggiunte dall'utente mediante `\hyphenation`.
3. Infine, se non ha trovato la parola nella lista delle eccezioni, il programma attiva l'algoritmo di sillabazione mediante i *pattern*.

I pattern sono sequenze di lettere di varia lunghezza in cui si possono scomporre le parole. Si noti quanto segue.

1. Prima di ogni sequenza, fra una lettera e l'altra, e alla fine della sequenza ci sono delle cifre da 0 a 9 (lo zero si può sottintendere).
2. Una cifra dispari indica che la divisione è permessa, una cifra pari indica che è vietata.
3. Ogni stringa di lettere e cifre forma un pattern.

4. Se fra le stesse due lettere presenti in pattern distinti si trovano cifre diverse, prevale la cifra maggiore.

Merita fare un esempio perché, sebbene la funzione dei pattern sia descritta correttamente nel capoverso precedente, il meccanismo si capisce molto meglio se lo si simula sulla pagina; si capisce anche molto bene perché il programma di composizione si comporti così e non diversamente.

La tabella 1.1 mostra il processo di divisione in sillabe del sintagma *dell'istruzione*. Poiché per il programma di composizione l'apostrofo è un carattere valido per formare una parola, l'intera stringa è considerata una *parola*. La cosa è possibile in italiano solo assegnando all'apostrofo un codice di lettera minuscola diverso da zero. In italiano, in francese, in catalano, e in diverse altre lingue l'apostrofo dopo una consonante e prima di una vocale marca l'elisione vocalica, quindi deve essere considerato alla stessa stregua di una vocale.

Di seguito si riporta la spiegazione della tabella.

1. La prima riga della tabella contiene la parola da dividere. Si noti la presenza di un punto prima della prima lettera e dopo l'ultima: per il pattern quel segno rappresenta l'inizio o la fine di una parola.
2. Nella seconda riga appaiono i pattern di una sola lettera contornati da zeri che non sono presenti nel file `hyph-it.tex`, perché il valore zero è quello sottinteso.
3. La terza e la quarta riga riportano i pattern di una lettera effettivamente presenti nel file dei pattern debitamente incolonnati sotto le corrispondenti lettere della parola da sillabare.
4. La quinta riga contiene i pattern di due lettere presenti nella parola in questione debitamente incolonnati sotto le corrispondenti stringhe.
5. Il file dei pattern non ne contiene altri utili per questa parola, per cui si può tirare una riga e riportare in ogni colonna di cifre quella che ha il valore più alto; in ogni colonna di lettere, la lettera comune dell'intera colonna. Il risultato ottenuto è la stringa

de2l1l'2i1s2t2ru1zio1ne

che conserva solo le cifre maggiori di zero.

In questo modo, si vede a colpo d'occhio che i possibili punti di cesura si trovano in corrispondenza delle cifre dispari, quindi il programma può dividere la parola come *del-l'i-stru-zio-ne*. Corretto!

L'algoritmo di sillabazione appena descritto deve essere molto veloce, perciò i pattern devono essere registrati nelle memorie interne del programma sotto forma di una struttura informatica speciale che ne consenta la ricerca in modo efficientissimo. Ecco spiegato perché i pattern devono essere caricati

.d	e	l	l	'	i	s	t	r	u	z	i	o	n	e.														
.d	0	e	0	l	0	l	0	'	0	i	0	s	0	t	0	r	0	u	0	z	0	i	0	o	0	n	0	e.
		1	l			'	2	1	s	2		1	r					1	z							1	n	
			1	l							1	t																
		2	l	0	l							t	2	r														
.d	0	e	2	l	1	l	0	'	2	i	1	s	2	t	2	r	0	u	1	z	0	i	0	o	1	n	0	e.

Tabella 1.1: Pattern usati per dividere in sillabe *dell'istruzione*

una volta per tutte nel file di formato, e non caricati di volta in volta. L'utente di $\text{Lua}\text{\LaTeX}$, invece, sa per esperienza che componendo esclusivamente in inglese il programma di composizione è abbastanza veloce, ma con altre lingue i pattern relativi vengono caricati in memoria solo al momento di comporre il documento, il che rallenta la partenza del programma. Con $\text{pdf}\text{\LaTeX}$ e $\text{X}\text{\LaTeX}$ questo non succede, ma se l'installazione del sistema $\text{T}\text{\LaTeX}$ è parziale, è possibile che il file di formato sia stato creato inizialmente solo con i file di pattern effettivamente disponibili, ciò che potrebbe escludere completamente la sillabazione o permetterla solo con i pattern per l'inglese.

Il sistema $\text{T}\text{\LaTeX}$ dispone del comodo pacchetto `testhyphens` per controllare come il programma di composizione sillaba le parole. Esso definisce la dichiarazione `\testhyphens` e l'ambiente `checkhyphens`, che stampano nel documento, una per riga, l'elenco di parole divise secondo le regole della lingua corrente. per esempio, in italiano si ha:

<code>{\testhyphens</code>	del-l'i-dea-le
<code>dell'ideale maxwelliano</code>	max-wel-lia-no
<code>macro"istruzione</code>	ma-cro-istru-zio-ne
<code>download</code>	do-wn-load
<code>hardware software</code>	hard-ware
<code>}</code>	soft-ware

Come si vede, la parola *download* è sillabata in modo errato. Basta indicare l'eccezione con `\hyphenation{down-load}` per ottenere il risultato corretto:

<code>\hyphenation{down-load}</code>	del-l'i-dea-le
<code>{\testhyphens</code>	max-wel-lia-no
<code>dell'ideale maxwelliano</code>	ma-cro-istru-zio-ne
<code>macro"istruzione</code>	down-load
<code>download</code>	hard-ware
<code>hardware software</code>	soft-ware
<code>}</code>	

L'ambiente `checkhyphens` si usa nello stesso modo, ma non c'è bisogno di usare le graffe. Si ricordi, però, di scrivere `\end{checkhyphens}` sempre su una riga a sé.

1.2.2 I FILE PER CARICARE I PATTERN

Va detto che scrivere i pattern per la sillabazione di una lingua è un'operazione difficile per vari motivi: per esempio, i file vanno confezionati e caricati diversamente a seconda che si usino `babel` e $\text{pdf}\text{\LaTeX}$ o `polyglossia` e $\text{X}\text{\LaTeX}$. Un modo ancora diverso è richiesto dal tandem `polyglossia` e $\text{Lua}\text{\LaTeX}$.

Per questi motivi, per ogni lingua gestibile con i tre programmi principali basati su \LaTeX devono esistere ben tre file per caricare i pattern nel file di formato e con le opportune righe di caricamento contenute in tre file distinti: `language.dat`, `language.def` e `language.dat.lua`. In questa sede non è il caso di descrivere ulteriormente questi file e i corrispondenti file locali, nel caso che un utente voglia definirsi gli strumenti per un'ulteriore lingua. Le informazioni necessarie, anche se sommarie, si trovano nella documentazione della distribuzione del sistema $\text{T}\text{\LaTeX}$ installata sul proprio calcolatore. La quale contiene anche il modello delle righe necessarie per ogni lingua, da replicare *mutatis mutandis* nei file locali.

Va ricordato, però, che l'utente normale raramente deve preoccuparsi di queste cose, perché l'ottantina di lingue per le quali il sistema $\text{T}\text{\LaTeX}$ è già attrezzato probabilmente riesce a soddisfare la maggior parte delle esigenze. Tuttavia...

No, non si vuole affermare che tutti i lettori debbano inventarsi una nuova lingua per cui creare file di descrizione e file di pattern, oltre ai file necessari

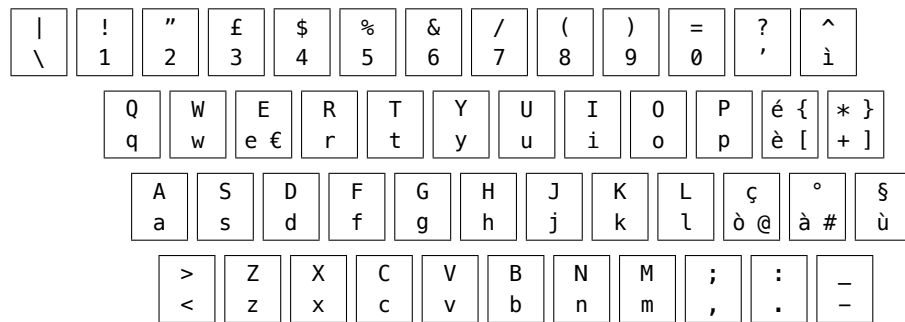


Figura 1.1: Schema della tastiera italiana sulle piattaforme Windows

per caricare i pattern nei file di formato, per rigenerare infine i file di formato. Chi scrive (CB) cura questi aspetti per diverse lingue presenti fra le ottanta indicate sopra (tra le quali l'italiano) ed è quindi al corrente delle difficoltà appena menzionate. Ma se qualche linguista volesse scrivere un saggio sull'ostrogoto usato dai Franchi nell'epoca carolingia, avrebbe bisogno di creare almeno i file dei pattern. La cosa migliore da fare, dopo aver letto la teoria nel *T_EXbook*, sarebbe quella di aprire qualche file di sillabazione per altre lingue, riconoscibili per la forma generale `hyph-⟨sigla della lingua⟩.tex`: imparerebbe facilmente la tecnica. Il difficile è creare l'insieme minimo e sufficiente di pattern per la lingua in questione, che va conosciuta abbastanza bene per evitare di fare un buco nell'acqua.¹

1.3 LE TASTIERE

La tastiera costituisce un problema ricorrente per l'utente di L^AT_EX in particolare e dei comuni word processor in generale. Tanto per capirci, quella italiana non è adatta nemmeno per scrivere in italiano; al che si aggiunga che in Italia sembrano non esserci negozi di informatica – dal piccolo esercizio di quartiere alle grandi catene di distribuzione – in grado di vendere calcolatori dotati di tastiere più versatili.

Perché la tastiera italiana è inadatta per lavorare con i programmi del sistema T_EX e per scrivere in italiano?

Alla prima parte della domanda si può rispondere brevemente: i quattro segni ' , ~, {, } non sono accessibili direttamente. Solo su alcune macchine la tastiera riporta le graffe come 'maiuscole' delle parentesi quadre, che a propria volta richiedono l'uso del tasto `AltGr`. Ma per gli altri due segni bisogna ricorrere a combinazioni di tasti che spesso richiedono di digitare il relativo codice numerico ASCII sul tastierino numerico: `Alt + 96` per il backtick e `Alt + 126` per la tilde (si veda la figura 1.1).

La seconda parte della domanda, invece, richiede una spiegazione di carattere dattilografico. In Italia, la tastiera italiana da computer è stata disegnata in modo da replicare le tastiere delle macchine da scrivere più avanzate esistenti all'inizio degli anni '80 del secolo scorso. Quelle tastiere non prevedevano i tasti morti per gli accenti, ma solo tasti per le lettere minuscole

¹Durante la scrittura di questa guida CB ha scoperto che la lingua occitana è selezionabile solo con polyglossia e non con babel. Andando più a fondo egli si è accorto che il modulo `gloss-occitan` definisce solo le parole infisse e la data, ma non seleziona nessun set di pattern di sillabazione, se non altro perché non esiste nessun file nel sistema T_EX che contenga questi pattern; andando ancora più a fondo egli ha scoperto che polyglossia in questi casi seleziona i pattern della pseudolingua `nohyphenation` che, come dice il nome, non divide in sillabe nessuna parola. L'occitano perciò viene composto molto male da un punto di vista tipografico. È possibile che al momento della pubblicazione di questa guida il problema sia stato risolto.

accentate e alcuni segni di dubbia utilità in italiano, come quello per la lettera ç (non in uso nella nostra lingua), il segno § (forse già allora desueto), il segno ° (utile forse nella scrittura di lettere commerciali, anche se l'autore (CB) addirittura dubita che lo si usasse del tutto).

In compenso mancavano le vocali maiuscole accentate, dattilograficamente riprodotte con la combinazione 'vocale maiuscola non accentata + apostrofo' (E', scorretto, anziché Ê, corretto). Allora bisognava fare di necessità virtù, ma l'avvento dei primi PC permise di rimediare a questi errori sistematici. In italiano, infatti, l'apostrofo posposto a una lettera e seguito da uno spazio segnala un'*apocope*, come quando si scrive *po'* al posto di *poco*. Di fatto, dunque, A' corrisponde all'apocope di *Ai* o *Agli*, E' è l'apocope di *Ei* o *Egli*, eccetera. A questo punto si potrebbe osservare che oggi simili apocopi non vengono quasi più usate in italiano, ma qui stiamo parlando di tipografia, che serviva e serve non solo, ma anche, per comporre edizioni critiche di testi italiani dei secoli passati, quando queste ortografie erano vive e vegete. In ogni caso, con i font oggi disponibili sui calcolatori non si vede la necessità di ricorrere alle modalità di scrittura correnti oltre trent'anni fa.

Oggi bisogna poter programmare anche in linguaggi non così specialistici come poteva succedere negli anni '80, perciò una qualunque tastiera *deve* permettere all'utente di accedere a tutti i caratteri ASCII semplicemente dai propri tasti: senza premerne mai più di due, evitando di doverne imparare i relativi codici per combinazioni di tasti più complesse. Tutte le altre tastiere nazionali permettono non solo questo, ma anche di scrivere i caratteri specifici di lingue diverse da quella per cui la tastiera è stata predisposta: si pensi alla presenza dei tasti morti per inserire i diacritici. Per esempio, la tastiera svizzera permette di scrivere correttamente in tedesco, francese, italiano e romancio senza doverne cambiare il driver. Chissà perché con la tastiera italiana, predisposta per una sola lingua, si riesce a scrivere correttamente in italiano solo intrecciando le dita.

La tastiera USA estesa per le moderne macchine Mac dispone dei tasti per tutti i caratteri ASCII, dei tasti morti per i segni diacritici e permette di inserire i segni più diversi con semplici combinazioni di tasti. La figura 1.2 nella pagina successiva ne mostra il layout. dispone di 47 tasti che consentono di inserire il 47 segni disegnati nella parte inferiore di ogni tasto; premendo il tasto dele maiuscole si compongono gli altri 47 segni disegnati nella parte superiore di ciascun tasto; si aggiunga il tasto per lo spazio; e si hanno in totale 95 segni da inserire (spazio compreso) cioè i 95 segni definiti "stampabili" nella definizione dello standard ASCII (che ne contiene in totale 128, numerati da 0 a 127), cioè i caratteri con indirizzi (decimali) da 32 a 126 compresi.

Nello standard ASCII dei 128 caratteri disponibili i primi 32 segni, numerati da 0 a 31, sono caratteri di controllo delle vecchie telescriventi e comprendono anche il segno di tabulazione, Tab, il segno di avanzamento della carta di una riga, <LF>, il segno di ritorno del carrello di stampa, <CR>, e il controllo di cancellazione dal carattere precedente, Del. I segni immessi con i tasti Tab e Del, con nomi simili ma spesso in lettere minuscole o non abbreviati, sono presenti fra i tasti non letterali ancora oggi; i caratteri ASCII <CR> e <LF> sono usati dai moderni sistemi operativi in modo diverso per marcare il fine record delle linee di testo effettivamente inseriti nel file sorgente quando si preme il tasto invio o enter o return. I programmi di composizione del sistema T_EX trattano i caratteri che arrivano nel file sorgente in modo coerente, trattano i tabulatori come spazi e i fine record come spazi se sono semplici e come fine capoverso se sono almeno due di seguito; gli altri caratteri vengono completamente ignorati.

Quando si usano codifiche a 8 bit, per esempio Latin 1, o la codifica UNICODE, le tastiere devono essere attrezzate per inserire altri segni non

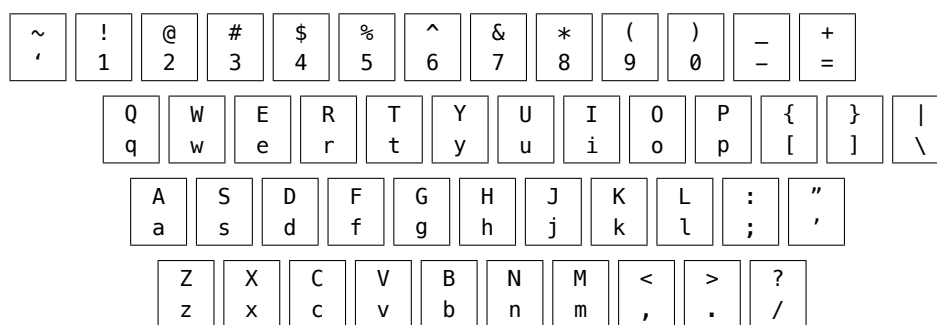


Figura 1.2: Schema della tastiera USA

disegnati sui tasti. In teoria il numero di segni inseribili potrebbe essere di altri 95 segni ogni volta che si preme uno dei tasti di controllo `control`, `command`, `Alt`, `AltGr`, eccetera, quindi ogni tastiera avrebbe enormi possibilità se i driver fossero progettati in modo accorto. Di fatto con l'uso al massimo di due tasti di controllo, si potrebbero duplicare i caratteri inseribili dalla tastiera portandolo 190 segni, ma questo avviene molto raramente.

Infatti i tasti di servizio sono quasi sempre gli stessi, non importa se la tastiera è predisposta per Windows o per un altro sistema operativo. L'unica vera differenza sta nella coppia di tasti (identici) a cavallo della barra spaziatrice: sulle attuali tastiere Windows essi raffigurano un'icona tonda divisa in quattro spicchi di quattro colori diversi e servono ad aprire il pannello Start. Sulle moderne tastiere per (Mac) OS X, riportano il solito simbolo ⌘ completato con il nome in lettere, `command` o `cmd` a seconda della lingua per cui la tastiera è predisposta.

Su entrambi i tipi di tastiera, i due tasti immediatamente all'esterno di quelli appena descritti riportano la scritta `Alt`. Sulle tastiere Mac si legge su entrambi anche `alt` oppure `option`, mentre su quelle Windows (e su quella italiana in particolare) il tasto a destra recita `AltGr` e svolge compiti diversi dal tasto `Alt` di sinistra.

Infine, immediatamente a sinistra dei tasti `Alt` e `option` c'è il tasto `Ctrl` (Windows) oppure `control` (Mac).

Tutti questi tasti appena descritti svolgono funzioni particolari a seconda del sistema operativo installato nel calcolatore, ma in linea generale modificano la funzione degli altri tasti normali, esattamente come i tasti `shift` e `caps lock` impostano le maiuscole temporaneamente o in modo permanente.

Nella figura 1.1 a pagina 8 i segni che si ottengono premendo `AltGr` sono riportati nella parte destra di ogni tasto; nelle figure 1.1 a pagina 8 e 1.2 i segni che si ottengono tenendo premuto `shift` sono riportati nella parte superiore del tasto; nei tasti con quattro segni, quello superiore destro si ottiene premendo `shift` + `AltGr`.

Sulle macchine Windows si possono configurare temporaneamente con profitto anche altri tasti di controllo, di solito per lavorare con programmi specifici. La totale sicurezza di poter usare *qualunque* programma, però, si ha solo con la combinazione già vista `shift` + `AltGr`.

Sulle macchine Mac, invece, una tastiera italiana ben usata presenta all'utente problemi molto minori per almeno due motivi.

1. Si può configurare il sistema in modo da far comparire nella parte destra della barra dei menù il menù "Tastiera" indicato dalla bandierina corrispondente o da una lettera 'A' bianca su sfondo nero; cliccando su questo simbolo si apre un dialogo di scelte che contiene la voce "Visore

Tastiera". Cliccando sopra quest'ultimo, sullo schermo comparirà la tastiera effettivamente in uso, con la quale è immediato vedere che cosa succede premendo i tasti di controllo.

2. Dopo un po' di uso sistematico del visore, si imparerà a usare in modo avanzato i vari tasti di controllo per ottenere tutti i segni che la macchina è in grado di usare, ma che non sono disegnati sui tasti.

Si noti che la cosa funziona anche quando si selezionano driver di tastiera diversi come quelli per il greco politonico, l'ebraico, l'arabo, il giapponese, eccetera.

Usando la tastiera mostrata nella figura 1.2 nella pagina precedente su un Mac si possono sfruttare anche gli accenti morti per scrivere in italiano. Entrambi gli autori di questa guida la usano, tanto che le combinazioni **Alt** + **'** per attivare l'accento grave e **Alt** + **e** per quello acuto sono diventate quasi automatiche. Per altre lingue come il francese e il tedesco sono utili le combinazioni **Alt** + **u** per la dieresi, e **Alt** + **6** per il circonflesso. Altre come **Alt** + **o** per *ø*, **Alt** + **s** per *ß*, **Alt** + **q** per *œ*, **Alt** + **'** per *æ* si ricordano con facilità, ma all'occorrenza c'è sempre il Visore Tastiera. I caporali aperti «' e chiusi '» richiedono le combinazioni **Alt** + **** e **Alt** + **shift** + **** rispettivamente: un po' meno mnemoniche, ma presto automatiche usandole spesso.

Si noti che anche con i sistemi operativi Windows 7 (forse) e 8 (certamente) è possibile visualizzare la tastiera sullo schermo e caricare diversi driver di tastiera, selezionabili dalla relativa icona a destra nella barra inferiore.

Infine, anche lo schermo *touch screen* ('schermo sensibile al tocco') dei vari tablet in circolazione permette le operazioni appena viste. Anzi, spesso i tablet più grandi sono sufficientemente avanzati da permettere di cambiare tastiera (chi scrive (CB) ha visto effettivamente un'utente orientale passare dalla tastiera giapponese a quella francese con estrema semplicità).

Questo lungo excursus sulle tastiere mostra chiaramente il punto più o meno debole, a seconda dei sistemi operativi, per scrivere in lingue diverse da quella nazionale per la quale la tastiera è stata pensata. Per scrivere in lingue che usano l'alfabeto latino esteso, *T_EX* viene in soccorso con i comandi per i diacritici che accettano come argomento lettere non accentate; in questo modo si possono inserire molti più segni diacritici che con qualunque tastiera nazionale, ma la cosa è di piccola consolazione, perché il file sorgente diventa poco leggibile.

Che cosa consigliare? È difficile dare consigli perché le soluzioni possibili dipendono dal sistema operativo, dall'hardware in uso, e dalla disponibilità a lavorare con due tastiere. Il lettore che arriverà alla fine della lettura di questa guida, tenga presente che si è scritto anche in giapponese e coreano usando soltanto un MacBook Pro normalissimo, con la tastiera USA, facendo uso di diversi driver nonché del visore di tastiera. Quindi il collo di bottiglia della tastiera è facilmente superabile. Si noti; si sarebbero potute fare cose simili con un hardware windows, e una qualunque variante del sistema operativo Linux; non sarebbe stato altrettanto facile con una versione dei sistemi operativi Windows privi della applicazione per usare una tastiera virtuale interattiva; fortunatamente questo scoglio dovrebbe essere superato con i sistemi operativi recenti.

2

COMPORRE IN VARIE LINGUE

Dopo aver descritto i principali problemi presentati dalle tastiere, ora si vedrà come comporre praticamente in alcune lingue che presentano particolarità.

Si ricordi che è sempre bene registrare i propri file sorgente sul disco fisso con la codifica UTF-8 indipendentemente dal motore di composizione scelto. Usando `PDFLATEX`, nel preambolo *deve* essere presente questa riga:

```
\usepackage[utf8]{inputenc}
```

Con `XYLATEX` e `LuaLATEX`, come si è detto, è sufficiente impostare l'editor in uso per lavorare in UTF-8.

Non si scordi mai di scegliere i font per il proprio documento. Se si desiderano quelli preimpostati (i Computer Modern), lo si faccia a ragion veduta e non per dimenticanza. In particolare, con `PDFLATEX` si ricordi sempre di specificare una codifica dei font che contenga già le lettere accentate e di scegliere una collezione di font che le contenga effettivamente: i font predefiniti del sistema `TEX` *non contengono le lettere accentate!* Come è noto, i Computer Modern le 'costruiscono' sovrapponendo alla lettera base il particolare diacritico: il risultato è accettabile da un punto di vista grafico, ma inaccettabile tipograficamente parlando, poiché tale operazione impedisce la sillabazione della parola dalla lettera accentata in poi. Va da sé che riducendo la possibilità di eseguire la cesura in fine di riga, la composizione tipografica risulta meno professionale.

2.1 COMPORRE IN ITALIANO

La composizione in italiano con `babel` e `PDFLATEX` è già stata abbondantemente discussa nella guida *L'arte di scrivere con L^AT_EX*: qui la si riprenderà solo per approfondire alcune funzioni avanzate di `babel` e per mostrare le differenze che polyglossia e `XYLATEX` stabiliscono con quanto è già noto.

Infatti, nonostante il modulo di `babel` per l'italiano non preveda la possibilità di specificare modificatori o attributi, tuttavia definisce tre comandi per attivare e talvolta per disattivare alcune funzionalità particolari. Si ricorda che la sillabazione grammaticale dell'italiano è fonetica ma in tipografia si usa relativamente spesso anche quella etimologica, specialmente per i termini tecnico-scientifici e i composti.

In realtà, i pattern di sillabazione per l'italiano sono predisposti per funzionare anche con sillabe iniziali e finali di una sola lettera, ma si sa che la lunghezza minima specificata nel file `italian.ldf` è di due. Ciò produce una conseguenza importante: in contesti particolari (come un ambiente o un gruppo) e per scopi particolari si può abbassare il valore predefinito mediante i seguenti due comandi:

```
\lefthyphenmin=1  
\righthyphenmin=1
```

Questa impostazione estrema potrebbe effettivamente ‘rimediare’ alcuni tratti di composizione in colonne molto strette, ma si noti che un suo uso estensivo renderebbe il testo piuttosto faticoso da leggere.

Di seguito si descrivono tre funzionalità avanzate di babel.

1. La possibilità di usare la divisione etimologica è assicurata da un comando per marcare il punto di giunzione dei membri del composto: il doppio apice ASCII `"`, che per poter garantire questa e altre funzionalità va prima dichiarato *attivo*. Poiché questa sua proprietà può risultare dannosa in certe circostanze, nel modulo `italian.ldf` è definito il comando `\setactivedoublequote`, da dare tra la chiamata di babel e `\begin{document}`. Una volta attivato, per disattivarlo bisogna ricorrere ai comandi di babel

```
\shorthandoff{\langle lista di caratteri \rangle}
\shorthandson{\langle lista di caratteri \rangle}
```

Siccome in italiano solo il doppio apice ASCII può essere attivo, la `\langle lista di caratteri \rangle` si riduce al solo `"`.

Quando è attivo, il doppio apice si comporta come una macro e come tale può ricevere una definizione che lo fa comportare in modo diverso a seconda del proprio argomento. Le sue azioni sono le seguenti.

- a) `"` seguito da una lettera dell’alfabeto non accentata inserisce un marcatore di separazione tra i membri di un composto. In quel punto la cesura è permessa, ed entrambi i membri rimangono divisibili. Per esempio: scrivendo nel sorgente `macro"istruzione`, si ottengono i seguenti punti di cesura: *ma-cro-i-stru-zio-ne*¹.
- b) `"|` svolge la stessa azione, ma va dato se seguito da una lettera con diacritici. Per esempio: *kilo"|örsted*.
- c) `"-` stampa un trattino copulativo, che non impedisce la sillabazione di quanto precede e segue. Per esempio: *italo"-francese*.
- d) `"< e ">` producono rispettivamente i caporali aperti ‘<’ e chiusi ‘>’ eliminando gli eventuali spazi che li separano dal proprio contenuto. A differenza di quella francese, la tradizione tipografica italiana non prevede spazi fra i caporali e il loro contenuto. Per esempio: `"<_Il_canzoniere_>` produce «Il canzoniere».
- e) `""` produce le doppie virgolette alte aperte `"` per supplire alle manchevolezze della tastiera italiana. In realtà, per stampare le virgolette alte semplici aperte ‘ e chiusa ’, L^AT_EX offre già di suo due comandi (praticamente sconosciuti perché la loro descrizione è in genere omessa nelle guide): `\lq` e `\rq`. Va da sé che per ottenere quelle doppie basta ripeterli. Dunque, con P_DF_EL^AT_EX (ma non con X_LL^AT_EX) le due scritture

```
""Il canzoniere""
```

```
\lq\lq Il canzoniere\rq\rq
```

producono entrambe *“Il canzoniere”*.

- f) `"/` produce la barra diritta di opposizione tra due parole, permettendo la sillabazione di quanto viene prima e dopo. È equivalente al comando `\slash`, più lungo da scrivere. Per esempio: *input/output* è ottenuto con `input"/output`.

¹La divisione mostrata si può avere quando almeno `\lefthyphenmin=1`. Normalmente quel parametro vale 2, quindi la divisione dei due membri rimane *ma-cro-istru-zio-ne*.

Si raccomanda di non abusare del carattere attivo appena descritto, ma di riservarlo a quei rari casi in cui servisse nella revisione delle bozze.

2. Il pacchetto `siunitx` può manipolare in qualunque modo i valori numerici delle misure e le unità di misura, che è in grado di comporre correttamente secondo le norme ISO grazie a comandi potentissimi. Tuttavia, se l'utente sa bene come scriverle o il documento ne richiede pochissime, `siunitx` può risultare inutilmente oneroso. Risolve il problema il comando `\unit`, che stacca correttamente le unità di misura dal loro valore numerico mediante uno spazio fine non separabile. Il comando ha un'argomento che va riempito con i simboli ISO e i loro eventuali esponenti o con i segni di moltiplicazione (`\cdot`, `\cdot`) e di divisione (`/`). Il comando funziona sia in modo matematico sia in modo testo (ma in questo secondo caso gli esponenti vanno esplicitamente scritti in modo matematico). Le norme ISO richiedono apici e pedici che non rappresentano delle variabili in tondo anche in modo matematico: i comandi `\ap` e `\ped` rispondono a questa esigenza, ma possono funzionare anche in modo testo. Nel modulo `italian.ldf`, la funzionalità di `\unit` è disattivata per impostazione predefinita: per attivarla, basta dare `\setISOcompliance` nel preambolo dopo aver caricato `babel`.
3. Il pacchetto `siunitx` gestisce perfettamente anche il separatore decimale ed è in grado di cambiare quello anglosassone, costituito dal punto, in quello 'normale' costituito dalla virgola. Quest'ultima operazione, però, richiede di impostare sia il pacchetto, sia il comando `\num` per ogni numero con una parte decimale. Più sbrigativamente, il modulo `italian.ldf` può attivare o disattivare la funzionalità della *virgola intelligente*, che in modo matematico, e solo in quel modo, fa sì che la virgola capisca da sola se è subito seguita, senza spazi interposti, da una cifra esplicita o implicita: se lo è, si comporta come separatore decimale; altrimenti, con o senza spazi interposti, si comporta da segno di punteggiatura. Tale funzionalità, inattiva per impostazione predefinita, può essere attivata o disattivata con i comandi `\IntelligentComma` e `\NoIntelligentComma`.

Si faccia attenzione alla clausola «se è subito seguita da una cifra»: infatti, nel caso in cui si voglia usare la virgola come separatore seriale in una sequenza numerica, bisogna lasciare uno spazio dopo ogni virgola. Negli esempi seguenti se la virgola è dichiarata 'intelligente', la virgola seriale deve essere seguita da uno spazio, altrimenti se è subito seguita da una cifra, la considera un separatore decimale. Al contrario se la virgola non è dichiarata 'intelligente', in un numero fratto essa è considerata una virgola seriale e la spaziatura delle cifre è chiaramente errata. Si noti che con la virgola seriale non è necessario lasciare uno spazio dopo la virgola seriale (ma certamente non è sbagliato farlo) quando la virgola è seguita da una macro il cui sviluppo non rappresenta una cifra.

a) Virgola seriale.

<code>\IntelligentComma</code>	
<code>\$\forall i = 0, 1, 2, \dots, \infty\$</code>	$\forall i = 0, 1, 2, \dots, \infty$
<code>\$\forall i = 0,1,2, \dots, \infty\$</code>	$\forall i = 0,1,2, \dots, \infty$

b) Separatore decimale.

<code>\IntelligentComma</code>	
<code>\$\pi = 3,1415\$</code>	$\pi = 3,1415$
<code>\NoIntelligentComma</code>	
<code>\$\pi = 3,1415\$</code>	$\pi = 3, 1415$

Con \LaTeX , si possono *parzialmente* attivare le funzioni del doppio apice passando l'opzione `babelshorthands` (inattiva per impostazione predefinita) al comando che carica la lingua italiana, come si è fatto in questo stesso documento:

```
\setmainlanguage[babelshorthands]{italian}
```

2.2 COMPORRE IN TEDESCO

2.2.1 CON PDFLATEX

Per scrivere in tedesco con `babel` e \PDFLATEX occorre solo selezionare il nome giusto per la lingua tra uno dei seguenti quattro.

german Seleziona l'ortografia tedesca in vigore prima del 1996, utile per riprodurre fedelmente un testo precedente quella data.

ngerman Seleziona la 'nuova ortografia', introdotta nel 1996.

swissgerman Seleziona la varietà svizzero-tedesca.

austrian Seleziona la varietà austriaca.

Ecco un esempio:

Alle Metaphysiker sind demnach
von ihren Geschäften feierlich
und gesetzmäßig so lange
suspendiert, bis sie die Frage:
Wie sind synthetische
Erkenntnisse a priori möglich?
genugtuend beantwortet haben.

Alle Metaphysiker sind demnach
von ihren Geschäften feierlich und
gesetzmäßig so lange suspendiert,
bis sie die Frage: Wie sind syntheti-
sche Erkenntnisse a priori möglich?
genugtuend beantwortet haben.

2.2.2 CON XELATEX

Con `polyglossia` e \XeLaTeX le cose sono un po' più facili. La lingua viene selezionata come principale o accessoria mediante i due comandi

```
\setmainlanguage{german} % lingua principale

oppure

\setotherlanguage{german} % lingua secondaria
```

che accettano diverse opzioni con la sintassi $\langle \textit{chiave} \rangle = \langle \textit{valore} \rangle$. descritte qui di seguito con l'avvertenza che il *valore* predefinito è il primo indicato.

variant = `german, austrian, swiss`.

spelling = `new` (seleziona ortografia e sillabazione riformate nel 1996), `old`.

latesthyphen = `false, true` (carica nuovi insiemi di pattern sperimentali).

babelshorthands = `false, true` (rende attivo il doppio apice).

Con la codifica UTF-8 obbligatoria per i file sorgente di \XeLaTeX e disponendo di una tastiera tedesca con cui scrivere direttamente le lettere già sormontate dalla *umlaut*, il doppio apice attivo per immettere la dieresi è sostanzialmente inutile, anche se potrebbe aiutare gli utenti che non dispongono di una tastiera adeguata o facilitare il riutilizzo di sorgenti scritti parecchio tempo fa. Tuttavia, il tedesco è un po' speciale anche per la sillabazione, come si può osservare nei seguenti casi di utilizzo del doppio apice attivo.

1. "ck serve per sillabare una parola nella forma *k-k*.
2. "ff serve per sillabare una parola nella forma *ff-f*. Analogamente fanno "ll, "mm, "nn, "pp, "rr e "tt;
3. "- introduce un punto di cesura che permette la sillabazione anche del resto della parola.
4. "" messo dopo il trattino che separa i due membri di un composto permette la sillabazione anche di quanto segue il trattino.
5. "~ inserisce un marcatore di parola composta senza inserire nessuna cesura.
6. "= inserisce un marcatore di parola composta nella quale si può avere una cesura e permette la sillabazione dei membri del composto.

script = latin, fraktur (compone le didascalie con i font *Fraktur*).

Questi non sono gli unici usi del carattere attivo " in tedesco; l'elenco completo per il tedesco si trova nel paragrafo 6.8 della documentazione di polyglossia. Sono interessanti in particolare i comandi per inserire le virgolette tedesche "' e "" che producono „ ", o quelle svizzere "< e "> che producono « » senza spazi fra i caporali e il loro contenuto.

2.3 COMPORRE IN LATINO

2.3.1 CON PDFLATEX

Per la lingua latina, babel e PDFLATEX offrono una varietà piuttosto ampia di strumenti. Di seguito si descrivono i modificatori e gli attributi previsti dal modulo `latin.ldf`.

modern Equivale a non specificare nulla e seleziona il latino moderno, in cui l'ortografia delle parole fisse è priva di dittonghi legati, le lettere *u* e *v* sono usate come in italiano, la sillabazione segue le regole fonetiche della nostra lingua.

medieval Seleziona il latino in uso nel Medioevo, con dittonghi legati (*æ* e *œ*), la *u* al posto della *v* e la *V* al posto della *U* nelle parole fisse e la sillabazione fonetica adatta a questo tipo di ortografia. Attenzione: anche il testo del documento deve essere composto con l'ortografia medievale, perché babel non esegue nessuna conversione ortografica.

classic Rappresenta il latino che si sarebbe scritto nell'età classica (I secolo a.C. – I secolo d.C.) avendo a disposizione anche l'alfabeto minuscolo.² Si usa la *u* al posto della *v*, la *V* al posto della *U*, e non ci sono i dittonghi legati. Dato che la lettera *u* si pronunciava *u*, *v* o *û*, la sillabazione, che selezionando questo attributo è etimologica, cerca di capire da sola in quali circostanze quel segno valeva da consonante.³

²Com'è noto, il latino veniva scritto tutto maiuscolo, senza spazi interparola e, almeno nelle prime fasi della sua evoluzione, con andamento boustrophedon. Iscrizioni pompeiane a parte, tracce di un alfabeto latino minuscolo compaiono in numerosi codici ricopiati da amanuensi di varie epoche, ma una vera e propria scrittura minuscola codificata si ha solo con la cosiddetta *rustica*, nata nel V secolo d.C., quando l'Impero romano d'Occidente era sull'orlo del collasso, se non addirittura già scomparso. Nella parte orientale dell'Impero, invece, si scriveva in greco.

³Quest'opera di interpretazione spesso ha successo, ma chi scrive (CB) crede che i latinisti potrebbero trovare molti esempi di sillabazione scorretta. Siccome chi scrive è anche l'autore delle regole di sillabazione, nonostante non sia un latinista, sono gradite tutte le segnalazioni di errore, meglio se corredate da qualche riga di spiegazione.

ecclesiastic serve per comporre il latino ecclesiastico, usato per testi teologici o devozionali della Chiesa Cattolica; esso è sostanzialmente latino moderno ma usa anche i dittonghi legati, usa gli accenti per marcare l'accento tonico sulle parole che il clero di diverse nazionalità e lingue proprie potrebbe altrimenti pronunciare con una ritmica diversa; usa una certa dose di spaziatura francese con i segni di interpunzione alti; compone le note al piede in stile francese. Questi modificatori sono presenti nel modulo per la lingua latina di babel in versione 3.5. Dovrebbero essere presenti anche in polyglossia, ma solitamente vi appaiono con un certo ritardo rispetto al modulo per babel; nel capitolo 4 viene detto di più in merito a come ottenere l'accento sul dittongo legato 'œ' che non è presente in nessun font OpenType.

withprosodicmarks Permette di attivare i segni di breve \breve e lunga \longchar , utili nella scrittura di grammatiche e dizionari, specie in latino moderno. Tuttavia, per evitare conflitti con i caratteri attivi $\^$ e $\=$ per marcare breve e lunga rispettivamente, il modulo `latin.ldf` definisce anche i due comandi `\ProsodicMarksOn` e `\ProsodicMarksOff` che attivano e disattivano questa funzionalità. questo modificatore non è disponibile con polyglossia, perché i font OpenType permettono di introdurre direttamente le vocali marcate con i segni di lunga e breve senza bisogno di ricorrere a macro.

Esistono anche due pacchetti, destinati specialmente al clero della Chiesa Cattolica, che permettono di scrivere il latino liturgico per messali, breviari e simili (pacchetto *liturgic*), e quello ecclesiastico (pacchetto *ecclesiastic*) necessario solo con babel per mantenere una retrocompatibilità con versioni precedenti del modulo per il latino; questo pacchetto *non* funziona con polyglossia e i font OpenType. polyglossia provvede alla composizione del latino ecclesiastico sfruttando alcuni comandi poco noti del nucleo di xetex che consentono di specificare gli spazi prima e dopo categorie di segni; in sostanza la variante *ecclesiastic* del latino da comporre con \XeTeX è totalmente autonoma e non richiede l'intervento di pacchetti esterni. Per ogni approfondimento si rinvia alla loro documentazione.

2.3.2 CON XELATEX

Con polyglossia e \XeTeX la lingua latina si specifica con i soliti comandi che accettano le opzioni seguenti con la sintassi $\langle chiave \rangle = \langle valore \rangle$. Il primo $\langle valore \rangle$ è sempre quello predefinito.

variant = *modern, medieval, classic, ecclesiastic*⁴.

Si noti che le opzioni appena menzionate sono disponibili solo a partire dalla \TeX Live 2015; chi scrive non ha modo di controllare se lo siano anche per \MiKTeX 2.9.

Attualmente il modulo per il latino non accetta ulteriori opzioni, per altro inutili scrivendo il testo sorgente con la codifica UTF-8: segni prosodici, accenti tonici e dittonghi accentati del latino ecclesiastico si possono immettere direttamente da una tastiera adatta su Windows, OS X e probabilmente anche sulle varie incarnazioni di Linux.

Qui si presenta un esempio di latino nella varietà classica tratto dai *Commentarii de bello Gallico* di Giulio Cesare.

⁴Siccome mi occupo dei moduli per il latino (CB) sia per babel sia per polyglossia ho un controllo diretto per l'aggiornamento dei moduli per babel; invece per polyglossia se ne occupa il curatore di quel pacchetto a cui pervengono i contributi dei curatori dei diversi moduli; ne raccoglie gli aggiornamenti e solitamente esegue un solo aggiornamento globale all'anno; perciò è possibile che la variante *ecclesiastic* sia disponibile solo a partire dal 2016. nel frattempo si può procedere come indicato nel capitolo-4.

Flumen est Arar, quo per
fines Haeduum et Sequanorum
in Rhodanum influit,
incredibili lenitate, ita ut
oculis in utram partem fluat
iudicari non possit.
Id Heluetii ratibus ac
lintribus iunctis transibant.

Flumen est Arar, quo per fines Hae-
duorum et Sequanorum in Rhoda-
num influit, incredibile lenitate, ita ut
oculis in utram partem fluat iudica-
ri non possit. Id Heluetii ratibus ac
lintribus iunctis transibant.

Per quanto riguarda il latino ecclesiastico un qualunque libro di laudi o simili testi devozionali potrebbe contenere il *Pater Noster*. La presenza degli accenti assicura che tutti i lettori, spesso con lingue materne molto differenti fra loro, possano pronunciare le parole all'unisono, con ritmo e accentazione identici. L'esempio che segue è composto in latino ecclesiastico senza sfruttare ancora la variante ecclesiastica per poliglossia, ma mostra come anche senza disporre di quella variante sia possibile ottenere già un buon risultato grazie all'uso dei font OpenType.

Pater noster qui es in cælis\\
santificétur Nomen Tuum\\
advénia Regnum Tuum\\
fiat voluntas Tua\\
sicut in cælo et in terra.\\
Panem nostrum\\
cotidiánum da nobis hódie\\
et dimítte nobis débíta
nostra,\\
sicut et nos\\
dimíttimus debitóribus
nostris.\\
Et ne nos indúcas in
tentatiónem\\
sed líbera nos a Malo. Amen.

Pater noster qui es in cælis
santificétur Nomen Tuum
advénia Regnum Tuum
fiat voluntas Tua
sicut in cælo et in terra.
Panem nostrum
cotidiánum da nobis hódie
et dimítte nobis débíta nostra,
sicut et nos
dimíttimus debitóribus nostris.
Et ne nos indúcas in tentatiónem
sed líbera nos a Malo. Amen.

2.4 COMPORRE IN GRECO

Il greco, come è noto, si scrive con un alfabeto diverso da quello latino, ma presenta anche altre importanti differenze, riassunte di seguito.

1. C'è il greco moderno monotonic, che prevede un solo tipo di accento (in genere con la forma di un accento acuto) e la dièresi, ma non usa gli spiriti.
2. C'è il greco moderno politonic, che usa tutta la collezione di accenti, spiriti, dièresi e iota sottoscritto e adscritto, ma esplicita gli iati con la dièresi anche se non sempre serve. Per esempio: *ἀύλος* non richiederebbe la dièresi sulla *υ*, ma questa varietà preferisce scrivere *ᾠλος*.
3. C'è infine il greco antico, che si scrive come il greco moderno politonic, ma usa la dièresi solo se è strettamente necessario. Ovviamente in greco antico il lessico e la sillabazione sono diverse da quelli del greco moderno politonic e, ovviamente, da quello monotonic.

Tre ortografie diverse richiedono tre diversi nomi di lingua da specificare a babel e polyglossia e un totale di sei file di pattern: tre per `pdfLATEX` e tre per `X3LATEX` e `LuaLATEX`.

SCRIVERE DIRETTAMENTE IN GRECO

Con `X3LATEX` e disponendo di una tastiera adatta, si può introdurre direttamente il testo greco nel file sorgente senza problemi.

Con \LaTeX qualche problema c'è, perché un unico font non può contenere i circa 256 glifi del latino esteso e altrettanti di greco politonico (ce ne vorrebbero circa 300 solo per il greco, ma bisogna fare di necessità virtù). Come si sa, \LaTeX può gestire solo font con 256 segni, quindi per comporre in greco direttamente bisogna necessariamente cambiare font, il che significa essenzialmente cambiare codifica ed eventualmente famiglia.

Per poter scrivere nel sorgente in greco con \LaTeX , qualunque delle tre ortografie appena descritte si sia scelta, il preambolo del documento *deve* contenere la riga

```
\usepackage[utf8]{inputenc}
```

altrimenti la cosa non è possibile.⁵

SCRIVERE IN GRECO CON LA TRASLITTERAZIONE

Se l'utente non dispone di un sistema per introdurre il testo greco direttamente, quando compone con \LaTeX (e solo con \LaTeX) può usare la translitterazione latina. Ne esistono almeno due diverse: quella che fa riferimento alla codifica LGR (*Local GReek encoding*) e quella chiamata 'Beta code'.

La prima fa corrispondere a ogni carattere latino un carattere greco simile per suono o per forma; i diacritici sono tutti composti con caratteri ASCII preposti alla lettera da marcare, tranne lo iota sottoscritto, che viene posposto. Il modulo che realizza la conversione dal testo in ingresso al testo composto con i font LGR è già incluso in quello che gestisce il supporto della lingua greca per babel. Questo meccanismo si basa sulla LICR (*\LaTeX Internal Character Representation*), la quale, tenendo conto della codifica in entrata, a qualunque segno che compare nel file sorgente, greco o translitterato che sia, fa corrispondere un codice interno del tutto indipendente dalla codifica in entrata. Nel momento di comporre il documento, a ogni codice interno viene fatto corrispondere il singolo carattere con o senza diacritici tratto dalla polizza dei font LGR. Il procedimento è molto efficace e funziona bene se i diacritici singoli o a coppie sono inseriti come 'macro' non letterali. Si veda la tabella 2.2 a pagina 23.

Nella codifica Beta code, invece, tutti i diacritici sono posposti. Per le altre differenze si rimanda il lettore alle relative documentazioni. Con \LaTeX e Beta code il greco prende il nome di 'Ibycus' e bisogna specificare `ibycus` come opzione a babel. La codifica `ibycus` è usata specialmente negli archivi in rete, che riportano translitterati in quel modo brani e testi completi di greco classico. Il grecista, dunque, li può copiare e incollare nei propri file sorgente, per comporli con \LaTeX e babel a cui ha passato l'opzione `ibycus`. Per gli approfondimenti si rimanda il lettore alla documentazione di `ibycus` contenuta nel file `ibycus-babel.pdf` (consultabile come di consueto con il comando da terminale `texdoc`).

Tabella 2.1: Translitterazione latino-greca secondo la codifica LGR

a	b	g	d	e	z	h	j	i	k	l	m	n	x	o	p	r	s	c	t	u	f	q	y	w
α	β	γ	δ	ε	ζ	η	θ	ι	κ	λ	μ	ν	ξ	ο	π	ρ	σ	ς	τ	υ	φ	χ	ψ	ω

⁵Non si ricorderà mai abbastanza l'opportunità di salvare *sempre* i propri sorgenti nella codifica UTF-8, non solo componendo in italiano come lingua principale, ma anche in lingue prive di diacritici (come l'inglese): infatti, l'eventualità di dover scrivere parole in lingue che li richiedono è più frequente di quanto non si creda.

	'00	'01	'02	'03	'04	'05	'06	'07	'10	'11	'12	'13	'14	'15	'16	'17	
'000	—	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
'020	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	'10
'040	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	'20
'060	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	'30
'100	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	'40
'120	80	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	'50
'140	96	97	98	99	100	101	102	103	104	105	106	107	108	109	110	111	'60
'160	112	113	114	115	116	117	118	119	120	121	122	123	124	125	126	127	'70
'200	128	129	130	131	132	133	134	135	136	137	138	139	140	141	142	143	'80
'220	144	145	146	147	148	149	150	151	152	153	154	155	156	157	158	159	'90
'240	160	161	162	163	164	165	166	167	168	169	170	171	172	173	174	175	'A0
'260	176	177	178	179	180	181	182	183	184	185	186	187	188	189	190	191	'B0
'300	192	193	194	195	196	197	198	199	200	201	202	203	204	205	206	207	'C0
'320	208	209	210	211	212	213	214	215	216	217	218	219	220	221	222	223	'D0
'340	224	225	226	227	228	229	230	231	232	233	234	235	236	237	238	239	'E0
'360	240	241	242	243	244	245	246	247	248	249	250	251	252	253	254	255	'F0
	"00	"01	"02	"03	"04	"05	"06	"07	"08	"09	"0A	"0B	"0C	"0D	"0E	"0F	

Figura 2.1: I font greci con codifica LGR

2.4.1 CON PDFLATEX

Per scrivere in greco con `PDFLATEX` e `babel` bisogna specificare `greek` tra le opzioni del pacchetto e scegliere una variante ortografica indicando uno dei tre modificatori seguenti:

monotoniko per il greco moderno monotonico (essendo l'impostazione predefinita, può essere omissa);

polutoniko per il greco moderno politonico; e

ancient per il greco antico.

Ciascuno seleziona dal file di formato i relativi pattern per assicurare la corretta sillabazione del testo. Il programma userà i font codificati con LGR, la cui polizza completa è mostrata nella figura 2.1.

Nelle prime due righe della figura, dove normalmente cadono i caratteri ASCII relativi alle telescriventi del tempo andato, compaiono simboli speciali, i più importanti dei quali sono i numerali greci. Le righe dalla seconda all'ottava contengono le lettere greche non accentate maiuscole e minuscole in una precisa corrispondenza con le lettere latine mostrata nella tabella 2.1 a fronte. Nell'ultimo gruppo di otto righe ci sono tutte le lettere greche minuscole con i diacritici (per le maiuscole si veda poco più avanti).

COMPORRE IN GRECO MODERNO MONOTONICO

Per scrivere in greco monotonico basta passare l'opzione `greek` a `babel` senza modificatori né attributi. Un documento in greco monotonico come lingua principale si imposta in uno dei modi seguenti:

```
\documentclass[... ,english,...]{book}
\usepackage[... ,greek]{babel}
```

```
\documentclass[...]{book}
\usepackage[... ,english,... ,greek]{babel}
```

```
\documentclass[...]{book}
\usepackage[... ,main=greek,... ,english,...]{babel}
```

Il testo greco può essere immesso nel file sorgente direttamente dalla tastiera o tramite la traslitterazione mostrata nella tabella 2.1. I pochi diacritici previsti da questa varietà si ottengono facilmente premettendo a una vocale i seguenti segni:

- un apice per ottenere l'accento;
- un doppio apice per ottenere la dieresi; e
- un apice e un doppio apice (in un ordine qualsiasi) per ottenere dieresi e accentto.

Il font fa tutto da solo per sostituire la sequenza 'apice e/o doppio apice + vocale' con la vocale accentata. Motivi di crenatura possono richiedere una barra rovescia prima dei segni appena elencati: le scritture $\alpha''\upsilon\lambda\omicron\varsigma$, $\alpha'''\upsilon\lambda\omicron\varsigma$, $\alpha\'''\upsilon\lambda\omicron\varsigma$, $\alpha\''\upsilon\lambda\omicron\varsigma$ o $\acute{\alpha}\lambda\omicron\varsigma$, dunque, sono equivalenti, a meno di eventuali imperfezioni di crenatura con le prime due (ma solo con alcuni font). Si noti, infine, che la lettera latina *s* finale di parola produce lo stesso segno prodotto dalla lettera *c*: questa sostituzione è tanto forte che se si vuole scrivere un σ isolato bisogna nascondere la fine della parola e scrivere *sv*.

Ecco un esempio di testo greco monotonico con la sua immissione diretta.

<pre>\begin{otherlanguage*}{greek} Πρέπει να σκεφούμε μια συνάρτηση της οποίας γνωρίζουμε ότι υπάρχει το ολοκλήρωμα. \end{otherlanguage*}</pre>	<p>Πρέπει να σκεφούμε μια συνάρτηση της οποίας γνωρίζουμε ότι υπάρχει το ολοκλήρωμα.</p>
---	--

COMPORRE IN GRECO MODERNO POLITONICO

Per scrivere in greco moderno politonico vale quanto si è appena detto per la varietà monotonica, con le seguenti piccole differenze.

Innanzitutto bisogna specificare il modificatore/attributo *polutoniko* in uno dei modi seguenti, tutti equivalenti:

```
\documentclass[... ,english,...]{book}
\usepackage[... ,greek.polutoniko]{babel}
```

```
\documentclass[... ,english,...]{book}
\usepackage[... ,greek]{babel}
\languageattribute{greek}{polutoniko}
```

```
\documentclass[...]{book}
\usepackage[... ,english,greek]{babel}
\languageattribute{greek}{polutoniko}
```

```
\documentclass[...]{book}
\usepackage[... ,main=greek.polutoniko,english]{babel}
```

Poi bisogna usare tutti i diacritici o le macro per gli accenti richieste dalla codifica interna LICR mostrati nella tabella 2.2 nella pagina successiva: tranne lo iota sottoscritto, che nella traslitterazione latina va sempre postposto e non prevede una macro apposita, vanno sempre preposti alla lettera base.

Per scrivere direttamente in greco nel file sorgente, oltre a ricordarsi di passare l'opzione *utf8* al pacchetto *inputenc*, va tenuto presente che alcuni comandi interni usano le macro \langle e \rangle . Poiché esse confliggono o possono confliggere con omonime macro di altri pacchetti, e persino con una macro del

Tabella 2.2: Corrispondenza fra i caratteri latini e i diacritici greci. La terza riga mostra le macro per gli accenti: si possono indicare due accenti in coppia con una sola macro, tanto che le scritture `\>\'`, `\'\>`, `\>'` e `\'>` sono equivalenti. Si noti che è sconsigliabile usare `~` da solo, essendo il carattere attivo normalmente usato per lo spazio interparola indivisibile.

Diacritici greci		¨	´	´	´	ι
Traslitterazione latina	>	<	"	'	'	~
Macro per gli accenti	\>	\<	\"	\'	\'	\~

nucleo di L^AT_EX, dopo babel si ricordi di caricare anche il pacchetto textalpha per eliminare eventuali interferenze.⁶

Infine si ricorda il pacchetto begingreek, che simula per P_{DF}L^AT_EX il comportamento dell'omonimo ambiente greek disponibile con polyglossia. Ecco un esempio d'uso:

<pre>\begin{greek} Τη πάντα διδούση καὶ ἀπολαμβάνουσα φύσει ὁ παπαιδευμένος καὶ αἰδήμων λέγει; «δὸς, ὁ θέλεις, ἀπόλαβε, ὁ θέλεις». Λέγει δὲ τοῦτο οὐ καταθρασυνόμενος, ἀλλὰ πειθαρχῶν μόνον καὶ εὐνοῶν αὐτῇ. \end{greek}</pre>	<p>Τη πάντα διδούση καὶ ἀπολαμβάνουσα φύσει ὁ παπαιδευμένος καὶ αἰδήμων λέγει; «δὸς, ὁ θέλεις, ἀπόλαβε, ὁ θέ- λεις». Λέγει δὲ τοῦτο οὐ καταθρασυ- νόμενος, ἀλλὰ πειθαρχῶν μόνον καὶ εὐνοῶν αὐτῇ.</p>
--	--

COMPORRE IN GRECO ANTICO

Per comporre in greco antico valgono le stesse osservazioni fatte per il greco moderno politonico, con la differenza che anziché `politoniko`, va specificato il modificatore/attributo `ancient`. Le scritture seguenti, tutte equivalenti, mostrano come farlo, anche se va osservato che difficilmente si sceglierà questa varietà di lingua come quella principale del proprio documento:

```
\documentclass[... ,english,...]{book}
\usepackage[... ,greek.ancient]{babel}
```

```
\documentclass[...]{book}
\usepackage[... ,main=greek.ancient,english]{babel}
```

```
\documentclass[... ,english,...]{book}
\usepackage[... ,greek]{babel}
\languageattribute{greek}{ancient}
```

```
\documentclass[...]{book}
\usepackage[... ,english,greek]{babel}
\languageattribute{greek}{ancient}
```

⁶Il pacchetto textalpha non è precaricato, perché alcuni utenti potrebbero non averne bisogno e possono tranquillamente ignorare gli avvisi restituiti dal programma di composizione.

Ecco un esempio di composizione in versi: è l'incipit dell'*Iliade* di Omero.

```
\begin{greek}
Μῆνιν ἄειδε θεὰ Πηληϊάδεω Ἀχιλῆος\\
οὐλομένης, ἥ μυρὶ Ἀχαιοῖς ἄλγε'
ἔθηκε,\\
πολλὰς δ' ἰφθίμους ψυχὰς Ἄϊδι
προΐαψεν\\
ἡρώων, αὐτοῦς δὲ ἐλώρια τεῦχε
κύνεσσις\\
οἰωνοῖσί τε πᾶσι. Διὸς δ' ἐτελείτο
βουλή,\\
ἐξ οὗ δὴ τὰ πρῶτα διαστήτην ἐρίσαντε\\
Ἀτρεΐδης τε ἄναξ ἀνδρῶν καὶ δῖος
Ἀχιλλεύς.
\end{greek}
```

Μῆνιν ἄειδε θεὰ Πηληϊάδεω Ἀχιλῆος
οὐλομένης, ἥ μυρὶ Ἀχαιοῖς ἄλγε' ἔθηκε,
πολλὰς δ' ἰφθίμους ψυχὰς Ἄϊδι προΐαψεν
ἡρώων, αὐτοῦς δὲ ἐλώρια τεῦχε κύνεσσις
οἰωνοῖσί τε πᾶσι. Διὸς δ' ἐτελείτο βουλή,
ἐξ οὗ δὴ τὰ πρῶτα διαστήτην ἐρίσαντε
Ἀτρεΐδης τε ἄναξ ἀνδρῶν καὶ δῖος Ἀχιλλεύς.

2.4.2 CON XELATEX

Con polyglossia e Xe_{La}T_EX comporre in greco è molto più semplice. La varietà di greco scelta va passata come opzione al comando per impostare la lingua, come si mostra di seguito, a seconda che il greco sia una lingua secondaria:

```
\setotherlanguage[⟨opzioni⟩]{greek}
```

o la lingua principale del documento:

```
\setmainlanguage[⟨opzioni⟩]{greek}
```

Le tre possibili ⟨opzioni⟩ richiedono la solita sintassi ⟨chiave⟩=⟨valore⟩ (il valore predefinito è il primo).

variant = monotonic (o mono), polytonic (o poly), ancient.

numerals = greek, arabic.

attic = false, true.

La scelta della variante ortografica e della relativa sillabazione non richiedono commenti.

Qualche osservazione, invece, fa fatta per i numeri. La scelta preimpostata greek permette di usare i comandi `\greeknumeral` e `\Greeknatural` per scrivere i numeri con la notazione alfabetica, minuscola con il primo comando e maiuscola con il secondo:

```
\begin{greek}
\greeknumeral{1815}\\
\Greeknatural{1815}
\end{greek}
```

,αωιε'
,ΑΩΙΕ'

La chiave `attic` permette di scrivere i simboli acrofonici attici mediante i due comandi sinonimi `\atticnum` e `\atticnumeral` anche se, purtroppo, non tutti i font OpenType adatti a scrivere in greco ne contengono i glifi.⁷

⁷I font OpenType non contengono mai tutti i milioni di glifi che la codifica Unicode potrebbe identificare: numerali alfabetici 'normali' (maiuscoli e minuscoli), attici (rarissimi) e simboli paralfabetici usati in greco antico per indicare unità, decine e centinaia quasi mai sono compresi per intero nei glifi di un font. Se ne comprende facilmente la ragione considerando che i font OpenType sono stati creati e predisposti all'inizio del XXI secolo per gli usi moderni e non per gestire le edizioni critiche dei classici greci. La collezione di font CMU contiene in forma OpenType tutti i font precedentemente distribuiti con il sistema T_EX nella forma Type 1, e perciò sono

L'esempio seguente mostra un tipico inizio di file sorgente per un documento composto in inglese come lingua principale e greco antico come lingua secondaria. Il primo commento (una 'riga magica', in gergo) ricorda che il file deve essere memorizzato sul disco con la codifica UTF-8 ed è utile agli editor capaci di interpretarlo per autoconfigurarsi di conseguenza.

```
% !TEX encoding = UTF-8 Unicode

\documentclass{book}
\usepackage{fontspec}
\newfontfamily{\greekfont}{GFS Bodoni}
...
\usepackage{polyglossia}
\setmainlanguage{english}
\setotherlanguage[variant=ancient]{greek}
...
```

Con i comandi appena visti e senza bisogno di ulteriori pacchetti, polyglossia mette a disposizione l'ambiente greek, dentro il quale compone in greco con il font indicato nell'argomento di `\newfontfamily`. In questo caso si tratta del font OpenType Bodoni, prodotto e messo a disposizione dalla Greek Font Society e usato anche in tutti gli esempi greci di questa guida.

Come esempi di composizione con \XeTeX si potrebbero usare quelli mostrati nel paragrafo precedente per il greco composto con babel e \pdfTeX . Sintassi e risultati sono identici, perciò non li si ripete.

2.5 COMPOSIZIONE IN RUSSO

In questo paragrafo non si parla solo del russo, perché quello che si dirà, vale sostanzialmente per tutte le lingue scritte con l'alfabeto cirillico.

2.5.1 CON \pdfTeX

Lavorando con \pdfTeX , le codifiche dei font potrebbero essere diverse a seconda che il font contenga sia l'alfabeto latino sia quello cirillico con le varianti, o solo caratteri cirillici. Appartengono alla prima serie le tre codifiche seguenti, per ciascuna delle quali si dà l'elenco completo delle lingue relative.⁸

T2A Abaza, Avar, Agul, Adyghei, Azerbaijani, Altai, Balkar, Bashkir, Bulgarian, Buryat, Byelorussian, Gagauz, Dargin, Dungan, Ingush, Kabardino-Cherkess, Kazakh, Kalmyk, Karakalpak, Karachaeveskii, Karelian, Kirghiz, Komizyrian, Komipermyak, Kumyk, Lak, Lezghin, Macedonian, Mari-Mountain, Mari-Valley, Moldavian, Mongolian, Mordvin-Moksha, Mordvin-Erzya, Nogai, Oroch, Osetin, Russian, Rutul, Serbian, Tabasaran, Tadjik, Tatar, Tati, Teleut, Tofalar, Tuva, Turkmen, Udmurt, Uzbek, Ukrainian, Hanty-Obsskii, Hanty-Surgut, Gipsi, Chechen, Chuvash, Crimean-Tatar.

T2B Abaza, Avar, Agul, Adyghei, Meut, Altai, Balkar, Byelorussian, Bulgarian, Buryat, Gagauz, Dargin, Dolgan, Dungan, Ingush, Itelmen, Ka-

completi almeno quanto lo erano quelli originali. I font Type 1 CBgreek sono stati pensati per il greco classico, quindi anche i CMU lo sono per lo stesso motivo. In questo manuale non sono stati usati perché non si accoppiano bene con i font latini Palatino, che rappresentano una delle caratteristiche distintive della collezione delle guide *L'Arte di...*

⁸Gli elenchi sono tratti da *The L^AT_EX Companion*, seconda edizione. Naturalmente, né babel né polyglossia contengono file di definizione per ognuna di queste lingue.

bardino-Cherkess, Kalmyk, Karakalpak, Karachaevskii, Karelian, Ketskii, Kirghiz, Komi-Zyrian, Komi-Permyak, Koryak, Kumyk, Kurdian, Lak, Lezghin, Mansi, Mari-Valley, Moldavian, Mongolian, Mordvin-Moksha, Mordvin-Erzya, Nanai, Nganasan, Negidal, Nenets, Nivh, Nogai, Oroch, Russian, Rutul, Selkup, Tabasaran, Tadzhik, Tatar, Tati, Teleut, Tofalar, Tuva, Turkmen, Udyghei, Uigur, Ulch, Khakass, Hanty-Vahovskii, Hanty-Kazymiskii, Hanty-Obiskii, Hanty-Surgut, Hanty-Shurysharskii, Gipsi, Chechen, Chukcha, Shor, Evenk, Even, Enets, Eskimo, Ytlkagir, Crimean-Tatar, Yaltut.

T2C Abkhazian, Bulgarian, Gagauz, Karelian, Komi-Zyrian, Komi-Permyak, Kumyk, Mansi, Moldavian, Mordvin-Moksha, Mordvin-Erzya, Nanai, Orok (Ui1ta), Negidal, Nogai, Oroch, Russian, Saam, Old-Bulgarian, Old-Russian, Tati, Teleut, Hanty-Obiskii, Hanty-Surgut, Evenk, Crimean-Tatar.

Alla seconda serie appartiene la codifica X2, per tutte le lingue che scrivono in cirillico senza alcun carattere latino.

Si noti che alcune lingue compaiono in più di un elenco: ciò significa che il loro sottoinsieme di caratteri si trova in font cirillici codificati diversamente. Per sua fortuna l'utente non deve preoccuparsene, se non per scegliere la codifica dei font di uscita da passare al pacchetto fontenc nel preambolo:

```
\documentclass[...]{book}
\usepackage[T1,T2A]{fontenc}
\usepackage[utf8]{inputenc}
\usepackage[english,russian]{babel}
```

I font predefiniti per PDF_{LA}T_EX sono i CM-super, che contengono sia i caratteri latini sia tutti i glifi cirillici necessari.

Si noti che con PDF_{LA}T_EX è comodo definire un nuovo ambiente russian in questo modo:

```
\newenvironment{russian}{%
  \begin{otherlanguage*}{russian}}{%
  \end{otherlanguage*}}
```

per poterlo usare come si farebbe con X_{LA}T_EX:

<pre>\begin{russian} И сказал Бог: да произрастит земля зелень, траву, сеющую семя дерево плодovitoe, приносящее по роду своему плод, в котором семя его на земле. И стало так. \end{russian}</pre>	<p>И сказал Бог: да произрастит земля зелень, траву, сеющую семя дерево плодovitoe, приносящее по роду своему плод, в котором семя его на земле. И стало так.</p>
---	---

2.5.2 CON XELATEX

Con polyglossia e X_{LA}T_EX le cose sono più facili. Infatti, per poter comporre in tutte le lingue indicate sopra basta semplicemente scegliere nel preambolo un font che contenga i glifi cirillici. Ecco un esempio tipico per il russo come lingua principale.

```
\documentclass{book}
\usepackage{fontenc}
\newfontfamily{\cyrillicfont}{Linux Libertine 0}
...
```

```
\usepackage{polyglossia}
\setmainlanguage[...]{russian}
\setotherlanguage{english}
```

Quasi tutti i font OpenType distribuiti con il sistema TeX e quelli in dotazione al sistema operativo contengono l'alfabeto cirillico.

Come sempre, con polyglossia è disponibile l'ambiente `russian`, all'interno del quale il testo viene composto con i font della famiglia specificata nella definizione di `\cyrillicfont`, comando che vale per *tutte* le lingue che si scrivono in cirillico. Dunque, non serve che polyglossia associ il nome della lingua al nome del font come succede per le altre lingue. I puntini nell'argomento facoltativo di `\setmainlanguage` indicano le opzioni accettate nella consueta forma $\langle \text{chiave} \rangle = \langle \text{valore} \rangle$ e descritte di seguito (il $\langle \text{valore} \rangle$ predefinito è il primo).

babelshorhands = `false` (per attivare tramite scorciatoie le stesse funzionalità definite dal modulo `russian.ldf` di babel e descritte nel file di documentazione leggibile con il comando `texdoc babel-russian`), `true`.

spelling = `modern` (seleziona parole fisse e ortografia moderne), `old` (seleziona parole fisse e ortografia precedenti alla riforma del 1916).

L'opzione per il russo mette a disposizione anche i comandi `\asbuk` e `\Asbuk`, che permettono di comporre le enumerazioni con le etichette costituite da lettere russe minuscole o maiuscole rispettivamente.

2.6 COMPORRE IN SPAGNOLO E MESSICANO

Per comporre in spagnolo e in messicano non c'è niente di particolare da fare se non specificare l'opzione `spanish` a babel o a polyglossia; babel accetta anche il "dialetto" `mexican` ma le differenze sono minime. L'unica che rivela la vicinanza agli Stati Uniti è che il messicano usa sistematicamente il punto decimale, mentre lo spagnolo accetta sia la virgola sia il punto. Per la virgola prevede adeguate macro per usare in matematica la virgola come simbolo ordinario, invece che come segno di punteggiatura.

Per il resto spetta all'utente usare il lessico specifico di ciascuna delle due lingue che in verità differiscono per poche parole.

La documentazione del modulo `spanish` per babel dice che *probabilmente* lo spagnolo descritto da quel modulo potrebbe andare bene anche per la lingua scritta negli altri paesi dell'America centrale e meridionale.

2.7 COMPORRE IN PORTOGHESE E IN BRASILIANO

Il portoghese e il brasiliano parlati sono piuttosto diversi, ma le lingue scritte sono molto simili come lo sono le tradizioni tipografiche dei due paesi.

Per comporre in queste lingue basta specificare il nome della lingua `portugues` o `brazilian` a babel o a polyglossia con i modi propri di ciascun pacchetto; bisogna ricordarsi di specificare l'opzione `utf8` al pacchetto `inputenc` quando si usa `pdflatex`, mentre quando si usa `xelatex` basta ricordarsi di registrare il file sorgente con la codifica UTF-8.

Le due lingue, apparentemente, dispongono di file di pattern diversi, quindi ufficialmente potrebbero essere considerate lingue diverse, non varietà della stessa lingua. Di fatto il lessico differisce prevalentemente per un numero limitato di parole comuni, ma il brasiliano ha assimilato molte parole delle lingue africane introdotte in Brasile dai molti africani deportati in quella nazione nei tempi andati. Oggi poi la lingua sta assimilando anche le parole delle comunità di immigrati più numerose, comprese le comunità asiatiche.

Tabella 2.3: Le lingue che polyglossia può gestire (versione del 2015). I nomi di ogni lingua sono specificati con l'ortografia necessaria per selezionarla.

albanian	danish	icelandic	nko	slovenian
amharic	divehi	interlingua	norsk	spanish
arabic	dutch	irish	nynorsk	swedish
armenian	english	italian	occitan	syriac
asturian	esperanto	kannada	piedmontese	tamil
bahasai	estonian	khmer	polish	telugu
bahasam	farsi	korean	portuges	thai
basque	finnish	lao	romanian	tibetan
bengali	french	latin	romansh	turkish
brazilian	friulan	latvian	russian	turkmen
breton	galician	lithuanian	samin	ukrainian
bulgarian	german	lsorbian	sanskrit	urdu
catalan	greek	magyar	scottish	usorbian
coptic	hebrew	malaylam	serbian	vietnamese
croatian	hindi	marathi	slovak	welsh
czech				

Tuttavia le tradizioni tipografiche sono molto simili e le impostazioni predefinite al sistema \TeX tramite i suoi moduli di gestione delle lingue sono adeguati.

2.8 COMMENTI

Si potrebbe andare avanti a descrivere la composizione di tutte le lingue di cui il sistema \TeX è capace; nel prossimo capitolo si parlerà di alcune lingue che vengono composte con alfabeti che si scrivono da destra a sinistra o che si scrivono con collezioni di ideogrammi.

Ma non sembra utile insistere. Il lettore che dovesse scrivere in svedese, o in rumeno, o in islandese, o in maltese o in lingue meno diffuse o a diffusione regionale, potrà farlo trovando quella che gli serve fra l'ottantina di lingue per le quali il sistema \TeX è attrezzato. In ogni caso è sempre opportuno leggere la documentazione di polyglossia oppure quella di ciascuna lingua gestibile con babel; talvolta queste documentazioni specifiche per una data lingua contengono informazioni utili anche per lavorare con polyglossia.

La tabella 2.3 mostra l'elenco fornito con polyglossia.

Alcuni nomi sono direttamente i nomi locali; altri sono decisamente poco noti come le lingue *lsorbian* e *usorbian*; in italiano si chiamano lausazio inferiore e lausazio superiore. Sono lingue locali usate sia in territorio tedesco sia in quello boemo nelle zone di confine e non sorprende che siano una sorta di sintesi fra il tedesco e il boemo. Alcune lingue, come l'occitano, dispongono solo del file di descrizione della lingua con le parole fisse, la data e poco altro, ma non non eseguono la sillabazione.

In ogni caso, come si vede, le lingue gestibili con polyglossia sono davvero molte.

3

COMPORRE IN LINGUE ESOTICHE

3.1 INTRODUZIONE

In questa guida, con l'espressione 'lingue esotiche' si intendono le lingue che si scrivono da destra a sinistra (che etimologicamente si potrebbero chiamare 'retrograde') e quelle che si scrivono con ideogrammi e, a volte, in verticale anziché in orizzontale.

Per il primo gruppo si mostreranno l'ebraico e l'arabo. Quest'ultima lingua è particolarmente difficile da comporre, non solo perché è molto calligrafica e presenta lettere legate, ma anche perché la giustificazione del testo non si ottiene allargando gli spazi tra una parola e l'altra bensì allungando i tratti terminali delle lettere finali di parola.

Di seguito si descrivono le principali differenze che le lingue scritte da destra a sinistra (di qui in avanti *lingue RL* (*right-left*) in opposizione alle *lingue LR* (*left-right*), come vengono chiamate nella documentazione dei pacchetti del sistema $\text{T}_{\text{E}}\text{X}$) presentano rispetto alla scrittura tradizionale: l'utente non bilingue o poco abituato ad altri sistemi di scrittura, infatti, potrebbe trovarsi piuttosto disorientato. Si dà per caricato un driver di tastiera adatto alla scrittura di lingue RL.

1. Tutti i tasti che comportano movimenti del cursore verso destra producono un movimento verso sinistra, e viceversa.
2. L'effetto dei due tasti che servono a cancellare qualcosa ai due lati del cursore è invertito.
3. Lettere e simboli vengono inseriti a sinistra del cursore e non a destra.

Le lingue a ideogrammi come il giapponese pongono meno problemi, perché l'andamento normale della scrittura è LR su righe orizzontali. Si noti, però, che un testo composto in verticale richiede colonne con andamento dall'alto al basso e insieme RL. Ancora più difficile!

Di qui in avanti si considererà *unicamente* la composizione tramite $\text{X}_{\text{Y}}\text{L}_{\text{A}}\text{T}_{\text{E}}\text{X}$ e polyglossia. Non è impossibile comporre in queste lingue anche con $\text{P}_{\text{D}}\text{F}_{\text{L}}\text{A}}\text{T}_{\text{E}}\text{X}$, ma la gestione dei font è decisamente più complessa e il gioco non varrebbe la candela.

3.2 COMPORRE IN EBRAICO

Nonostante che polyglossia permetta di specificare lingue primarie e secondarie in un punto qualsiasi del documento, la composizione in ebraico richiede qualche attenzione in più. Essendo una lingua RL, infatti, si appoggia al pacchetto `bidi` che gestisce la scrittura con alfabeti bidirezionali. Poiché questo pacchetto va caricato *dopo* tutti gli altri, *compreso* `hyperref`, un buon punto per dichiararla è subito prima di `\begin{document}`, come mostra l'esempio seguente:

```

...
\usepackage{polyglossia}
\setmainlanguage{italian}
...
\usepackage[...]{hyperref}
...
\setotherlanguage{hebrew}
\newfontfamily{\hebrewfont}{Linux Libertine O}
\begin{document}

```

La scelta del font è caduta ancora una volta sul Linux Libertine O, che contiene anche i glifi per l'ebraico. Siccome non c'è bisogno di legature particolari tra i caratteri, lo si è dichiarato come nuova famiglia di font.

A questo punto, introdurre nel sorgente un brano in ebraico è semplice:

```

\begin{hebrew}
בראשית ברא אלהים את השמים ואת הארץ.
והארץ הייתה תהו ובהו וחשך על פני תהום;
אלהים מרחפת על פני המים.
ויאמר אלהים יהי אור; ויהי-אור.
\end{hebrew}

```

Si noti che anche il testo in ebraico nel file sorgente viene composto da destra a sinistra: alternando un alfabeto LR con uno RL, quindi, l'inversione del senso di scrittura potrebbe avvenire in punti scomodi come il fine riga, con non pochi grattacapi per l'utente.¹

3.3 COMPORRE IN ARABO

Per scrivere nelle varie lingue che usano l'alfabeto arabo servono essenzialmente tre strumenti. I primi due fanno parte del sistema \TeX : il consueto pacchetto `polyglossia`, che definisce lingue e ambienti necessari, e il pacchetto `arabxetex`, descritto nei prossimi capoversi. Il terzo strumento consiste in un font adatto, la cui scelta è compito dell'utente.

Tra i vari strumenti per scrivere in arabo, `arabxetex` è il pacchetto più completo e quello raccomandato a *tutti* gli utenti, specialmente a chi non parla e scrive abitualmente in questa lingua. Di seguito se ne descrivono le caratteristiche principali.

1. Gestisce numerose lingue fra quelle che si scrivono con l'alfabeto arabo. Eccone l'elenco con l'ortografia necessaria per selezionarle: arab (Arabic), farsi (persian), urdu, sindhi, pashto, ottoman (turk), kurdish, kashmiri, malay (jawi) e uighur.
2. Permette di usare particolari sequenze di caratteri ASCII (tutti accessibili da qualunque tastiera occidentale, tranne alcuni se si usa la tastiera italiana, elencate nella sua documentazione) che durante la composizione traslittera nelle corrispondenti lettere arabe in tutte le loro forme (isolata; iniziale, mediale e finale di parola; combinata con lettere particolari), realizzando automaticamente le varie legature e rispettando così la natura calligrafica di questa scrittura.
3. Permette di vocalizzare la scrittura (abitualmente costituita di sole 'consonanti' e pertanto di difficile lettura) con particolari segni diacritici da

¹Chi scrive non conosce l'ebraico, e anche la ripresa dei primi versetti della Bibbia da Internet e il loro adattamento alla composizione dell'esempio precedente ha richiesto non poca attenzione.

mettere sopra o sotto le consonanti. Già una vocalizzazione parziale rende il testo più leggibile e comprensibile; quella totale si usa di solito per scrivere il Corano, permettendo ai fedeli di leggerlo con precisione (tanto più che è scritto in un arabo antico contenente parole oggi obsolete) e contiene anche segni speciali per intonare il testo durante la liturgia.

4. Permette la traslitterazione in caratteri latini secondo le norme ISO.
5. Fornisce segni speciali per indicare il nome di Allah scritto in modo corretto e nelle varie forme necessarie a comporre brani coranici
6. Contiene anche ogni altra funzionalità specifica per le altre lingue che usano questo alfabeto.

A titolo d'esempio, si riporta un brano in arabo ripreso dalla documentazione di arabxetex. Nel file sorgente si è scritto quanto segue:

```
\begin{arab}
'amruN, 'ibiluN, 'u_htuN, '"u_ht"uN, '"Uql"Id"Is, ra'suN,
'ar'asu, sa'ala, qara'a, bu'suN, 'ab'usuN, ra'ufo, ru'asA'u,
bi'ruN, 'as'ilaTuN, ka'iba, qA'imuN, ri'AsaTuN, su'ila,
samA'uN, barI'uN, sU'uN, bad'uN, ^say'uN, ^say'iN, ^say'aN,
sa'ala, mas'alaTuN, saw'aTuN, _ha.tI'aTuN, jA'a, ridA'uN,
ridA'aN, jI'a, radI'iN, sU'uN, .daw'uN, qay'iN, .zim'aN
, yatasA'alUna, 'a'dA'akum, 'a'dA'ikum, 'a'dA'ukum maqrU'aT,
mU'ibAt, taw'am, yas'alu, 'a.sdiq^A$\'uh_u, ya^g^I'u, s^U'ila
\end{arab}
```

Come si vede, ogni parola è scritta in un codice composto di soli caratteri latini ASCII scritti da sinistra a destra. Il pacchetto arabxetex esegue le necessarie corrispondenze fra i caratteri latini maiuscoli e minuscoli e tiene conto del fatto che in arabo ciascuna parola ha tutte le lettere legate.

Ecco il risultato:

أمر، إبل، أخت، واخت، ءاوقل ايداييس، رأس، أرأس، سأل، قرأ، بؤس، أبؤس،
 رؤف، رؤساء، بئر، أسئلة، كئب، قائم، رأسة، سئل، سماء، بريء، سوء، بدء،
 شيء، شيء، شيئا، ساءل، مسألة، سوأة، خطيئة، جاء، رداء، رداء، جيئ،
 رديء، سوء، ضوء، قيء، ظمئا، يتساءلون، أعداءكم، أعدائكم، أعداؤكم
 مقروؤة، مؤنات، توأم، يسأل، أصدقاؤه، يجيء، سؤئل

Il font usato per comporre l'esempio è lo Scheherazade della SIL International. Poiché non è contenuto nel sistema T_EX né è in dotazione ai vari sistemi operativi, bisogna scaricarlo dalla Rete (è gratuito) e installarlo a mano. È lo stesso font usato dall'autore di arabxetex; in questa guida lo si è impostato con il comando

```
\newfontfamily{\arabfont}[Script=arabic]{Scheherazade}
```

Si noti che il nome della lingua (arab, in questo caso) è comune sia al nome della famiglia di font, di cui costituisce la prima parte, sia a quello dell'omonimo ambiente arab, come succede (quasi) sempre nelle associazioni lingua-font di polyglossia. In questo caso non si è definita la lingua arab con il comando `\setotherlanguage`, perché già impostata da arabxetex.

Per scrivere brani presi dal Santo Corano (القرآن الكريم), servono ulteriori simboli, quasi tutti già contenuti nel font Scheherazade. La documentazione di arabxetex spiega come fare e mostra molti esempi, ma in particolare documenta molto bene come usare allo scopo la traslitterazione ASCII.

Naturalmente, quanto detto per l'arabo vale anche per le altre lingue scritte con lo stesso alfabeto. Oltre all'ambiente `arab` e al comando `\textarab`, il pacchetto `arabxetex` definisce gli omologhi comandi per le altre lingue elencate precedentemente. La documentazione suggerisce anche scelte di font adatte e ne consiglia alcuni per definire i comandi analoghi a `\varabfont` in modo da associare il font giusto alla lingua giusta.

Si ricorda che con \LaTeX è possibile scrivere direttamente in arabo all'interno dell'ambiente `arab` e nell'argomento del comando `\textarab`. Quest'operazione, come già detto, è riservata a chi dispone di una tastiera araba e conosce la lingua; non è consigliabile invece a chi usa una tastiera per una lingua occidentale. In ogni caso, l'autore di `arabxetex` consiglia a *tutti* la codifica in caratteri latini, più dettagliata.²

3.4 COMPORRE IN GIAPPONESE E CINESE

3.4.1 IN ORIZZONTALE

Comporre con ideogrammi non è più difficile che farlo con un alfabeto, tuttavia la cosa richiede qualche attenzione in più.

La differenza più evidente con quanto detto finora è che il pacchetto `polyglossia` *non* dispone del file di descrizione di giapponese e cinese, perciò non si possono impostare queste due lingue con i comandi definiti dal pacchetto, non si può usare un font dichiarandolo con `\japanese font` come si farebbe con una lingua tradizionale, né, infine, si può contare su un corrispettivo ambiente `japanese`.

Certo: nulla vieta di definirsi una nuova famiglia di font con un nome a scelta, per comodità anche `\japanese font`, ma si ricordi che quel '`japanese`' è solo mnemonico per l'utente e non è legato alla lingua impostata con `polyglossia`. Per esempio:

```
\newfontfamily{\japanese font}{Hiragino Mincho Pro}
\let\chinese font\japanese font
```

Il font `Hiragino Mincho Pro` contiene sia i sillabari giapponesi `katakana` e `hiragana` sia i `kanji` utili per giapponese e cinese. Ecco perché si possono rendere equivalenti i nomi `\japanese font` e `\chinese font` per indicare la stessa famiglia di font. Analogamente si può fare con altri font.³

Volendo si possono definire anche un comando per immettere una breve serie di parole giapponesi:

```
\newcommand*\textjapanese[1]{\{\japanese font #1}}
```

e un ambiente simile a quelli che `polyglossia` mette a disposizione delle altre lingue:

```
\newenvironment{japanese}[1][\{\japanese font #1\japanese font\}]
```

L'argomento facoltativo permette di specificare il tipo di composizione, per esempio in bandiera con `\raggedright`. Analogamente si possono definire un comando `\textchinese` e un ambiente `chinese`.

²Chi volesse vedere un documento completamente in arabo con brevissimi inserti in inglese, può consultare il file `example-arabic.pdf`, il cui sorgente `example-arabic.tex`, moderatamente complesso e parzialmente in farsi, è scritto direttamente in arabo senza traslitterazione latina. È un esempio interessante, nonostante sia più d'aiuto agli arabisti che ai profani della lingua.

³Inoltre, si possono scegliere caratteri di foggia diversa per scrivere in ciascuna delle tre principali lingue a ideogrammi. Si tenga presente che, di solito, un font adatto al giapponese lo è anche per il cinese e contiene i glifi sillabici coreani.

Dopo queste dichiarazioni nel preambolo, si può comporre un testo come il seguente:

```
\begin{japanese}\raggedright
まだ何もない時、
神様は天と地をお造りになりました
地球はまだ形が定まらず
やみにおおわれた氷の上を
さらに神様の霊がおおっていました
「光よ、輝け」と神様が命じました。
すると光がさっと輝いたのです。
\end{japanese}
```

まだ何もない時、
 神様は天と地をお造りになりました
 地球はまだ形が定まらず
 やみにおおわれた氷の上を
 さらに神様の霊がおおっていました
 「光よ、輝け」と神様が命じました。
 すると光がさっと輝いたのです。

Le scritture giapponese e cinese presentano alcune caratteristiche:

- non prevedono spazi interparola;
- non conoscono regole di sillabazione; e
- usano in pratica caratteri monospaziati (ciascuno dei quali dovrebbe occupare lo stesso spazio orizzontale sulla riga).

Ciò richiede all'utente un certo lavoro manuale (meglio sarebbe che conoscesse anche un po' la lingua) per giustificare il testo, il che si può fare inserendo tra una parola e l'altra, dove serve, il comando `\XeTeXlinebreakskip`. In questo modo \XeTeX può fare un buon lavoro con la giustificazione. In alternativa, si può comporre `\raggedright`, con il margine destro non giustificato.

IMPOSTARE LA TASTIERA E SCRIVERE CON GLI IDEOGRAMMI

Con alcuni sistemi operativi (di sicuro (Mac) OS X, ma probabilmente anche Windows) si possono inserire i caratteri giapponesi direttamente nel sorgente tramite il metodo Kotoeri.

Innanzitutto bisogna impostare il driver di tastiera per il giapponese. Le istruzioni seguenti descrivono come farlo su (Mac) OS X.⁴

1. Si raggiunga `Preferenze di Sistema > Tastiera > Sorgenti di input`.
2. Nel pannello di dialogo visualizzato sullo schermo si preme il bottone `+` in basso a destra. Nel menù a discesa che compare si seleziona la voce `Giapponese` e si preme il pulsante `Aggiungi`. A questo punto, si dovrebbe vedere una finestra di dialogo più o meno come quella mostrata nella figura 3.1. Ci si assicuri che le voci `Hiragana`, `Katakana`, `Romaji` e `Mostra menu Tastiera nella barra dei menu` siano spuntate e si chiuda la finestra.

Ora, cliccando nella barra dei menù in alto a destra sulla bandierina che indica la sorgente di input predefinita scenderà un menù con dentro tre nuove voci: `A` Romaji, `あ` Hiragana e `ア` Katakana (i due ideogrammi rappresentano la sillaba *a* nei rispettivi sillabari, a segni arrotondati il primo e a segni spigolosi il secondo). La figura 3.2 mostra quanto si è appena descritto. A questo punto si può scrivere direttamente con gli ideogrammi nell'ambiente `japanese` o nell'argomento del comando `\textjapanese`.⁵

Un esempio chiarirà come fare. Dopo aver selezionato la tastiera Hiragana nel menù `Tastiera`, si scriva con la tastiera latina la sequenza di lettere che

⁴Chi scrive ha installato nella propria macchina El Capitan aggiornato alla versione 10.11.2, ma la procedura indicata dovrebbe valere anche per le versioni precedenti.

⁵Per farlo, è appena il caso di ricordare che o l'utente conosce la lingua, o copia e incolla del testo in giapponese da altre fonti (il che rende del tutto inutile selezionare la tastiera giapponese).

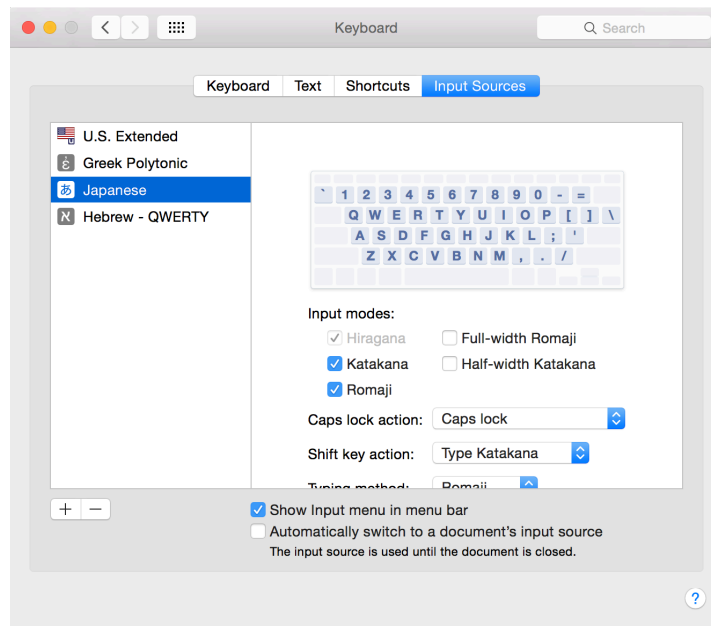


Figura 3.1: Finestra di dialogo per impostare la tastiera giapponese

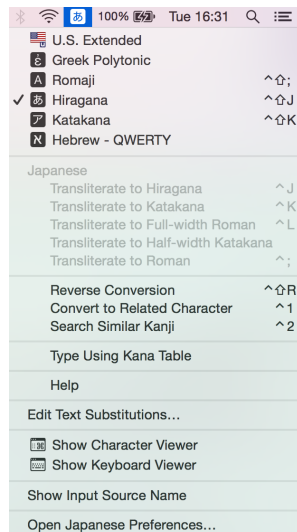


Figura 3.2: Il menu Tastiera in cui selezionare la tastiera Hiragana

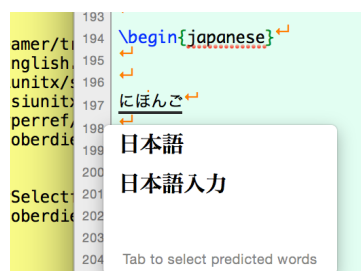


Figura 3.3: La parola *nihongo* ('lingua giapponese') scritta in caratteri hiragana nel sorgente e con due grafie in kanji nella finestrella contestuale

formano la parola *arigato* ('grazie'): si vedrà che ogni sillaba scritta in lettere latine (*a-ri-ga-to*) viene sostituita contestualmente con la corrispondente sillaba hiragana. I casi sono due, descritti di seguito.

1. Parola in hiragana.
 - a) La parola ha una sola grafia hiragana. Basta premere il tasto spazio e la parola comparirà nel file sorgente scritta in giapponese nell'unica grafia esistente.
 - b) La parole ha più grafie hiragana. Basta selezionare la grafia giusta tra quelle proposte nella finestrella contestuale con un doppio clic.⁶
2. Parola in hiragana e kanji.
 - a) La parola ha anche una sola grafia kanji.
 - b) La parola ha anche più grafie kanji.

Nei casi 2.a e 2.b valgono regole simili, salvo che compaiono sia la scrittura in hiragana la o le grafie con gli ideogrammi kanji; se questa è unica basta eseguire un doppio click sull'unica versione kanji; se di ortografie kanji corrette ne esistono più di una, bisogna conoscere il giapponese piuttosto bene (o disporre di un dizionario) per sapere quale grafia sia quella giusta facendovi sopra un doppio click per farla apparire nel testo sorgente.

La lingua giapponese si chiama *nihongo*. Scrivendo questa parola nel sorgente in caratteri latini, nella finestrella contestuale compaiono due scritture kanji. L'ortografia giusta è 日本語 e bisogna selezionarla con un doppio clic. La figura 3.3 mostra la finestrella con le due grafie in kanji e sopra la grafia hiragana.

Un'avvertenza superflua per il lettore attento. Quando si è finito di comporre in giapponese bisogna ricordarsi di selezionare nuovamente la tastiera preimpostata: l'ambiente *japanese*, infatti, non ha nulla a che vedere con il sistema operativo.

Sulle macchine Windows si possono installare i font per le lingue orientali (non presenti per impostazione predefinita) dai dischi di installazione del sistema operativo o con le modalità specifiche di ciascuna versione. Installando poi il driver per la tastiera orientale che interessa, si potrà usare una tastiera virtuale come quella descritta sopra per il giapponese.

⁶Va da sé che bisogna conoscere la lingua per decidere. Chi scrive non conosce che poche parole giapponesi e viste solo in scrittura romaji. Possedendo, però, un piccolo vocabolario bilingue italiano-giapponese, sa che l'ortografia giusta di *arigato* fra le due proposte nella finestrella è ありがとう, formata dalle sillabe a-ri-ga-to seguite dal segno orizzontale che marca l'ultima sillaba come lunga (la pronuncia corretta, infatti, è 'arigatoo').

Si segnala che fra i programmi di composizione del sistema \TeX c'è anche \LaTeX , esplicitamente configurato per gestire i font giapponesi tramite il motore \ptex , una variazione di \tex o di \pdftex che vale solo per il Giappone e il giapponese.

3.4.2 IN VERTICALE

Merita invece segnalare che nei testi 'molto' classici ed eleganti si possono trovare tanto il giapponese quanto il cinese (e forse anche il coreano) composti in verticale. In questo caso, le colonne di testo sono composte dall'alto in basso e si susseguono sulla pagina con andamento RL. Evidentemente si tratta di una modalità di composizione (e di lettura) molto diversa dal solito.

Con le accortezze descritte di seguito e un font adatto, \XeLaTeX \fontenc e \polyglossia permettono di riprodurre anche questo tipo di scrittura.

1. I capoversi vanno composti dentro scatole (ottenibili facilmente con un ambiente `minipage`) di giustezza pari all'altezza delle colonne inserendo gli ideogrammi nel consueto verso orizzontale.
2. A questo punto, bisogna operare una doppia rotazione: quella delle scatole prima, in modo da metterle in verticale; e quella dei glifi al loro interno poi (mediante le funzionalità di `fontspec`), in modo da orientarli correttamente.

L'esempio seguente, tratto da [Goossens, 2011] con alcune semplificazioni e usando un font diverso, è scritto in cinese (anche se si può usare lo stesso codice per il giapponese):

```

\begin{minipage}[t]{45mm}
Testo in orizzontale:\\[1.5ex]
\chinesefont 人人生而自由,在尊嚴和權 利上一律平等。
\end{minipage}
\quad
\begin{minipage}[t]{45mm}
\hspace*{-0.5em}Testo in verticale:\\

\rotatebox{-90}{%
\parbox[b][25mm][t]{45mm}{%
\chinesefont\addfontfeatures{Vertical=RotatedGlyphs}
人人生而自由,在尊嚴和權 利上一律平等。}}
\end{minipage}

```

Ecco il risultato:

Testo in orizzontale:

人人生而自由,在尊嚴和權
利上一律平等。

Testo in verticale:

利 人
上 人
一 生
律 而
平 自
等 由
。 在
在 尊
尊 嚴
嚴 和
和 權
權

Come si vede, rotazione delle ‘righe’ e controrotazione dei caratteri permettono di ottenere il risultato voluto.

Questo modo di comporre è tipicamente giapponese e si ritrova soprattutto negli *haiku*, brevissimi poemi tradizionali di tre versi sempre di 5, 7 e 5 sillabe. Oltre che nei libri di poesia, la composizione verticale si può trovare in testi di altro tipo che si rifanno alla tradizione e spesso anche in testi moderni, dove sono composte in questo modo epigrafi o dediche.

Ecco, infine, un haiku tratto (liberamente⁷) da [Starace, 2005]. Da sinistra a destra, l’esempio riporta il testo originale in giapponese verticale, la sua traslitterazione in romaji e la traduzione in italiano, lingua che si presta particolarmente bene alla composizione degli haiku.

風 か お る	不 味 く が 原 の	祇 園 絵 や	Gion e ya Makazu ga hara no kaze kaoru	Immagine di Gion. Della piana di Makazu profuma il vento.
------------------	----------------------------	------------------	--	---

3.5 COMPORRE IN COREANO

A differenza di cinese e giapponese, per il coreano esiste il file di descrizione `gloss-korean.ldf` da usare con `polyglossia`. Va da sé che sono definite la lingua `korean` e l’omonimo ambiente `korean`, dentro il quale viene usato automaticamente il font `koreanfont` se precedentemente impostato. L’alfabeto coreano si chiama *hangul*, ed è questo nome che bisogna dare come valore alla chiave `Script` nella specificazione eseguita con i comandi del pacchetto `fontspec`. Per esempio:

```
\newfontfamily{\koreanfont}{UnBatang}[Script=Hangul]
```

Il file di descrizione definisce le solite parole fisse, la data e i contatori alfabetici con i corrispettivi coreani. Non definisce, invece, un insieme di pattern di sillabazione, perché il coreano si scrive separando le parole, di solito formate da pochi glifi, tanto che di solito per giustificare le righe di testo basta lo spazio interparola. Tuttavia, definisce un insieme di segni per i quali è vietato eseguire un fine riga prima o dopo.

Di seguito si elencano le caratteristiche principali della scrittura coreana.

1. Ogni glifo rappresenta una sillaba ed è composto di due o tre segni elementari che rappresentano ciascuno una consonante o una vocale.
2. Ci sono 14 segni per le consonanti e 10 per le vocali.
3. In un glifo di due segni, la consonante sta sopra e la vocale sotto; un glifo di tre segni contiene due consonanti e una vocale.
4. I segni sono raggruppati dentro un quadrato ideale per cui i glifi del coreano sono naturalmente equispaziati.

Il nome della scrittura coreana *hangul* è formato da due glifi, ciascuno contenente tre segni: `h-a-n` e `g-u-l`. In coreano si scrive 한글. Ecco un esempio di testo coreano copiato e incollato dal file `example-korean.tex`:

⁷*Liberamente* significa che chi scrive (CB) non è riuscito a replicare esattamente *tutti* i kanji che appaiono nel testo originale. Inoltre, nella traduzione in italiano due versi sono stati scambiati perché così il testo originale sembra più fedelmente rispettato e più vicino alla sequenza di 5-7-5 sillabe, anche se i versi italiani contano un numero di sillabe diverso.

모든 인간은 태어날 때부터 자유로우며 그 존엄과 권리에 있어
동등하다. 인간은 천부적으로 이성과 양심을 부여받았으며 서로
형제애의 정신으로 행동하여야 한다.

Comporre in coreano richiede una tastiera adatta, selezionabile tramite un driver di tastiera opportuno (si consiglia la variante 2-Set Korean) come si è indicato per il giapponese. Lo stesso driver accosterà i singoli segni nelle giuste posizioni dentro il 'quadrato' della sillaba come si vede nell'esempio appena mostrato.

4

LE TRADIZIONI TIPOGRAFICHE

In questo capitolo si riassumeranno le tradizioni tipografiche in vigore in diversi paesi; in alcuni esse sono codificate tramite norme ufficiali predisposte dall'Ente Normativo dello Stato, in altre sono solamente delle consuetudini.

Spesso le indicazioni tipografiche sono una caratteristica stilistica delle varie case editrici; non è quindi facile dire che cosa sia corretto o scorretto fare quando si compone in prima persona un testo che abbia qualità tipografiche valide per la maggioranza dei paesi che usano la lingua principale del documento.

Nel seguire o nel darsi certe regole il compositore deve sempre avere come metro di giudizio il fatto che la buona tipografia non serve solo per produrre degli stampati accattivanti, ma deve servire per rendere la lettura facile e scorrevole, che non stanchi il lettore con cose che lo distraggano, o con righe mal composte, o con un uso eccessivo di cambiamenti di font; non stiamo parlando del contenuto, che comunque deve essere composto avendo il lettore in mente, ma di come il contenuto è presentato.

Ci sono in commercio diversi libri sullo stile tipografico; noi consideriamo il testo [Bringhurst, 1992] particolarmente utile. Ne esiste anche una versione in italiano che oggi sembra essere esaurita e non ristampata, ma consiglieremmo comunque la versione originale in inglese, non solo perché è scritta in un modo elegante e scorrevole, ma anche perché l'impaginazione è quella originale disegnata da Bringhurst stesso e rende particolarmente evidente che cosa significa "bella tipografia".

Si può dissentire da Bringhurst sotto molti aspetti; secondo noi le sue raccomandazioni sono adatte più a testi umanistici che a testi tecnico-scientifici; la parte del testo dedicata alla composizione di formule matematiche e fisiche è ridotta come succede spesso in questo tipo di pubblicazioni; ma, almeno per le scienze che fanno uso di grandezze misurabili e non solamente di variabili matematiche, esistono delle precise norme ISO valide per tutto il mondo, non solo per questo o quel paese: [ISO 31/XI, 1978; UNI-ISO 5966, 1989].

In questo capitolo non parleremo della composizione della matematica se non in quei pochi casi in cui le norme o le regole nazionali se ne discostano.

Parleremo anche degli insiemi di caratteri usati, talvolta in modi diversi, nelle varie tradizioni tipografiche per comporre in lingue diverse. Mostreremo anche qualche esempio di composizione per le lingue che non abbiamo trattato nel corpo di questa guida.

Vedremo quindi in pratica come vengono usati certi caratteri nazionali e come vengono composte certe strutture tipografiche.

Per quanto riguarda gli insiemi di caratteri, vale la pena di premettere le seguenti osservazioni.

La prima è che l'unica codifica che accetta tutti i set di caratteri è quella denominata UNICODE. Questa codifica prevede centinaia di migliaia di codici per caratteri diversi di diverse lingue e per diversi scopi. Normalmente oc-

corrono indirizzi di sei cifre esadecimali per accedere a qualunque carattere di un ipotetico font che contenga tutti i caratteri codificati da UNICODE.

Per molte lingue i segni usati per scriverle sono in numero molto maggiore delle 26 lettere latine minuscole con le corrispondenti 26 lettere maiuscole; inoltre ci sono i caratteri analfabetici, cifre, segni di interpunzione, simboli speciali ed altro che sono necessari per comunicare per iscritto.

Alcune lingue usano ideogrammi; il cinese, il giapponese, in un certo modo il coreano (anche se i suoi glifi in realtà sono alfabetici), sono le lingue più note che usano ideogrammi; questi insiemi ammontano a diverse decine di migliaia di glifi.

Alcune lingue usano segni sillabici, dove ogni segno rappresenta un'intera sillaba; i più noti sono i due sillabari giapponesi katakana e hiragana, ognuno dei quali contiene un centinaio di glifi; tra l'altro questi sillabari sono quelli che vengono insegnati nelle prime classi delle scuole elementari; successivamente vengono insegnati gli ideogrammi kanji; sembra che una persona debba conoscere un minimo di 2000 kanji, ma il numero totale di questi ideogrammi potrebbe ammontare a 35 000.

In altre lingue si usano alfabeti che cercano di rappresentare i suoni della lingua nel miglior modo possibile; mi risulta che l'alfabeto per scrivere in thailandese ammonti a circa 450 segni.

Ci sono alfabeti o insiemi di simboli per le lingue antiche, dal geroglifico, al sanscrito, dal lineare A al cuneiforme; ci sono segni speciali moderni che non possono descriversi come gli altri alfabeti, come i segni braille destinati agli ipovedenti che leggono passando i polpastrelli delle dita sui segni in rilievo.

Tutto ciò è codificato da UNICODE.

Per quel che riguarda le tipografie delle lingue dove si usano alfabeti o collezioni di segni diversi da quelli latini estesi, non scenderemo nei dettagli; in particolare non ci occuperemo dei caratteri cirillici e greci. È noto che queste lingue si scrivono con caratteri diversi da quelli latini ma le varianti cirilliche sono troppo estese per la moltitudine delle lingue che usano quell'alfabeto, che avrebbe la pretesa di rappresentare ogni fonema con un grafema unico e specifico, cosa che non è vera nemmeno per il russo, dove almeno la 'o' serve per indicare due suoni distinti.

Per il greco l'esposizione dei segni sarebbe molto semplice, ma ci sono molte varianti stilistiche nel testo e semantiche in matematica. Nel testo è noto che le lettere σ e ς indicano entrambe l'equivalente della lettera latina 's', ma il secondo glifo si usa sempre e soltanto in fine di parola, mentre il primo non si usa mai in quella posizione. Certi autori usano due forme della lettera beta: β e β , la seconda solo e soltanto all'inizio di parola. Il discorso si farebbe molto lungo e quindi lo si evita del tutto.

Tuttavia qui premettiamo un cenno su quella particolare lingua che si chiama *matematica*. È noto che in matematica si usano virtualmente tutti i segni dell'alfabeto greco, ma spesso alcuni segni hanno varianti grafiche che in matematica possono avere significati diversi; chi scrive, per esempio, ha usato φ per indicare la sezione aurea e ϕ per indicare il numero aureo; per cui si possono distinguere le due cose spesso indicate entrambe con uno solo dei due nomi; essi invece indicano cose diverse legate dalle relazioni:

$$\phi = \frac{1}{\varphi} \quad \text{e} \quad \phi = 1 + \varphi$$

In tutto quel che segue si indicheranno le lettere minuscole; generalmente per le maiuscole cambia il disegno della lettera base, ma si conservano i diacritici. In greco questo avviene solo per le iniziali maiuscole, ma quando una parola viene composta in tutte maiuscole si omettono tutti i diacritici tranne la dieresi.

4.1 TIPOGRAFIA ITALIANA

In Italia, che pur ha una tradizione tipografica molto lunga a partire dalla famosa tipografia di Aldo Manuzio a Venezia nella seconda metà del XV secolo, nell'arco dei secoli lo stile tipografico si è andato adattando allo stile francese e solo nel XX secolo ha ripreso una sua fisionomia, sebbene non molto originale rispetto a quella della tipografia anglo-americana.

Il testo di Cevolani, 2006 riassume le regole tradizionali; tuttavia va specificato che non esistono norme specifiche emesse dall'Ente Italiano di Unificazione (UNI), se non per alcuni dettagli come l'uso del segnapunto obbligatorio, quello che si usa per le parole tronche: sempre grave, tranne sulla 'e' dove talvolta deve essere acuto; la norma consente ma non impone l'uso dell'accento acuto sulla 'i' e sulla 'u'. Consiglia l'uso del segnapunto grave o acuto a seconda dei casi per le parole con l'accento sulla terzultima sillaba o ancora più arretrato, lo vieta sulle parole piane (accento sulla penultima sillaba) tranne nei casi che la lettera base sia 'e' oppure 'o' e il differente accento abbia un valore semantico, per esempio pèsca e péscia oppure còlto e cólto. Tuttavia, tradizionalmente, questi consigli sono spesso trascurati e si vede spesso scritto seguito (come la norma vieta di fare) per distinguerlo da séguito (che spesso non è accentato come la norma raccomanda di fare).

Le norme UNI danno anche le regole per la divisione in sillabe, ma riguardano solo le parole strettamente italiane e che non derivano da radici straniere (L^AT_EX e i tre programmi discussi in questa guida fanno molto meglio).

Ci sono poi le regole per la matematica, per i disegni tecnici e per la struttura dei rapporti tecnici, ma si tratta di poche cose e molto specializzate.

Direi che le tre cose che più differenziano la tipografia italiana da quella di altri paesi e da quella anglo-americana in particolare sono le seguenti.

1. In italiano si usano sia le virgolette alte inglesi (" "), sia i caporali o virgolette uncinate (« ») e le si usa sia doppie sia semplici (' ', e < >) sebbene le uncinate semplici siano piuttosto rare. A differenza di altre tradizioni tipografiche *non* si lasciano spazi dopo le virgolette aperte e prima delle virgolette chiuse.
2. La prima riga del capoverso dovrebbe essere sempre rientrata; non sarebbe necessario rientrare la prima riga del primo capoverso di una qualsiasi sezione, ma è una pratica diffusa. Il rientro serve appunto per marcare l'inizio di un nuovo capoverso anche nei casi in cui il capoverso precedente termini a filo del margine destro; ma il primo capoverso di ogni sezione è preceduto dal titolo della sezione, non da un altro capoverso, quindi il rientro non è necessario; chi lo imposta lo fa per uniformità; per impostazione predefinita il rientro del primo capoverso manca ma, quando si compone con pdf_latex, lo si può impostare richiamando il pacchetto indentfirst; con xel_latex e con lua_latex il rientro del primo capoverso è preimpostato; per disimpostarlo bisogna dare il comando `\PolyglossiaSetup{italian}{indentfirst=false}`.
3. Le lineette in italiano sono tre come in altre lingue, ma
 - il trattino breve '-' viene usato sia come segno di cesura in fin di riga sia come trattino copulativo, cioè che unisce due parole;
 - il trattino medio '—' viene usato per separare gli estremi di un intervallo numerico in un contesto non matematico (per esempio per indicare un intervallo di pagine); il trattino medio si usa anche per delimitare un inciso;
 - il trattino lungo '—' non si usa quasi mai; lo si vede usato quando si vuole omettere un nome (per esempio 'il signor A— . . .'); talvolta lo si vede usato per iniziare una battuta di un dialogo.

Ovviamente il segno ‘meno’ è un segno matematico e non si confonde con nessuno dei trattini; piuttosto, se bisogna inserire un segno ‘meno’ nel testo, bisogna scriverlo in modo matematico.

4. In italiano i vari segni di interpunzione del testo non vanno mai separati con nessuno spazio dal testo che li precede.

Per impostazione predefinita con `pdflatex` gli spazi *dopo* i segni di interpunzione sono, come in inglese, maggiori dopo il punto fermo e dopo alcuni altri segni; se li si vuole uguali, è sufficiente specificarlo all’inizio del documento con il comando `\frenchspacing`. Con `xelatex` e `lualatex` l’impostazione ‘alla francese’ è predefinita.

Gli incisi sono una questione delicata, possono essere racchiusi fra virgole, fra trattini o fra parentesi tonde; di solito i trattini sono separati dall’inciso mediante uno spazio non separabile, se non altro per distinguere un trattino che marca un intervallo da un trattino che delimita un inciso. Se l’inciso conclude un periodo che non cade alla fine di un capoverso, il punto fermo va messo dopo il trattino di chiusura; se invece l’inciso conclude un capoverso, il ‘punto e a capo’ permette di omettere il trattino di chiusura.

Per la matematica la situazione è più fluida; c’è chi sostiene che la matematica richieda i segni di interpunzione come il testo anche quando le espressioni sono fuori testo, e che siano necessari quei segni sia al loro interno sia al loro esterno; c’è chi ritiene che i segni di interpunzione interni alle espressioni ne facciano parte integrante, mentre i segni esterni siano adeguatamente sostituiti con gli spazi presenti nelle espressioni fuori testo. Vengono usati entrambi questi metodi ma la buona tipografia richiede coerenza; scelto un metodo i segni esterni o si mettono sempre o non si mettono mai; non bisogna insomma imbarazzare il lettore che non capirebbe perché talvolta ci siano dei segni esterni e talvolta non ci siano.

Anche i dialoghi sono una questione delicata. Le battute, per esempio possono essere dette o solo ‘pensate’; una battuta ne può contenere un’altra. Insomma c’è una gerarchia fra le battute che quindi vanno marcate differenzialmente. Poiché l’italiano usa sia le virgolette alte sia le virgolette uncinate doppie o semplici c’è la possibilità di distinguere le varie battute almeno a quattro livelli. Bisogna decidere quali virgolette usare per ogni livello, e poi essere coerenti.

Le virgolette, specialmente quelle alte, servono per evidenziare alcune parole, vuoi perché vengono usate come oggetti, non per il loro significato, sia perché la parola viene usata in senso ironico o metaforico. Talvolta le definizioni vengono evidenziate in corsivo. In certi testi didattici si usa anche il nero. Le parole straniere non assimilate in italiano solitamente vengono composte in corsivo, almeno la prima volta che vengono usate. Bisogna però stare attenti a non esagerare con i metodi di enfasi: se ne usa uno solo alla volta. La sottolineatura è vietata in tipografia: era il massimo che si poteva fare quando si usavano le macchine da scrivere; oggi ci sono metodi tipograficamente più idonei. Sarebbe errato mettere in corsivo nero una locuzione di cui si dà la definizione; sarebbe errato mettere fra virgolette una parola in corsivo. Ma queste considerazioni valgono qualunque sia la lingua usata per comporre un testo.

In italiano bisogna fare attenzione ad usare le lettere maiuscole con molta parsimonia; solo i nomi propri vanno scritti con la lettera iniziale maiuscola, oltre le parole che iniziano un periodo. Le maiuscole di cortesia (per esempio ‘Lei’, ‘Vostro’, ‘Signore’, eccetera) fanno parte del linguaggio burocratico, non del buon italiano; mesi, giorni, nazionalità, punti cardinali, e simili sono nomi comuni. Figura, tabella, equazione, eccetera, sono nomi comuni; inoltre se sono accompagnati da un numero o da altre espressioni qualificative sono

oggetti determinati e quindi richiedono gli articoli determinativi e le corrispondenti preposizioni articolate; omettere gli articoli o usare le preposizioni semplici vuol dire imitare lo stile anglo-americano.

Anche nei titoli dei libri, delle opere, dei palazzi e simili, a parte l'iniziale maiuscola della prima parola e il corsivo, vanno maiuscole solo le iniziali dei nomi propri. In inglese va invece usata l'iniziale maiuscola per quasi tutte le parole, ma non si usano quando si cita un titolo inglese in un contesto italiano; in tedesco, oltre ai nomi propri, richiedono l'iniziale maiuscola i sostantivi e queste iniziali maiuscole vanno conservate anche in un contesto italiano.

Infine un'avvertenza tecnica: per comporre in italiano, oltre a specificare la lingua a babel o a polyglossia, conviene scrivere il testo sorgente con la codifica UTF-8; con pdf_{latex} bisogna specificare l'opzione `utf8` al pacchetto `inputenc`, è necessario specificare l'opzione `T1` al pacchetto `fontenc` e, corrispondentemente, bisogna scegliere font che contengano i caratteri codificati con quella codifica. Con x_{el}at_{ex} e lua_lat_{ex} bisogna comporre il file sorgente con la codifica UTF-8, ma non bisogna invocare il pacchetto `inputenc`; bisogna invece usare `fontspec` per specificare font OpenType sufficientemente completi da contenere tutti i caratteri necessari.

Le vocali italiane possono assumere gli accenti grave o acuto, ma in certi contesti possono prendere anche la dieresi e il circonflesso; la dieresi indica lo scioglimento di un dittongo in uno iato, quindi può apparire solo sulla 'i' e sulla 'u'; il secondo si usa in poesia quando un'apocope fa cadere la parte finale di una parola sdrucchiola per ragioni metriche o per la rima.

Ecco un esempio tratto da una traduzione di VINCENZO MONTI dell'Iliade in endecasillabi senza rima¹.

O Atridi, ei disse, o coturnati Achei,
gl'immortali del ciel abitatori
concedanvi espugnar la *Prïameia*
cittade, e salvi al patrio suol tornarvi.
Deh! mi sciogliete la diletta figlia,
ricevetene il prezzo, e il saettante
figlio di Giove rispettate. — Al prego
tutti *acclamâr*: doversi il sacerdote
riverire, e accettar le ricche offerte.

In definitiva le lettere italiane semplici o con diacritici sono le seguenti:

a à â b c d e è é ê f g h i ì í î ï j k l
m n o ò ó ô p q r s t u ù ú û ü v w x y z

ma è del tutto ovvio che la dieresi e il circonflesso sono usati solo in contesti speciali e che solo sulla 'e' e sulla 'o' sono richiesti gli accenti grave o acuto secondo il loro suono; sulla 'i' e sulla 'u' l'accento acuto, stando alle norme UNI, è 'tollerato' ma non ne è prescritto l'uso.

4.2 TIPOGRAFIA FRANCESE

La tipografia francese, fra quelle europee, è quella che ha una personalità maggiormente definita, che non dipende solo dalla lingua usata ma da tradizioni secolari e dalla tendenza francese di non farsi imporre norme internazionali che non abbiano stabilito loro stessi; non c'è nessuna critica in questa frase, è semplicemente una constatazione naturale.

Per esempio solo in francese tutti i segni di interpunzione 'alti' (quindi tutti tranne il punto e la virgola) sono preceduti da uno spazio; il francese

¹Le parole contenenti dieresi o circonflesso sono evidenziate in corsivo.

usa i caporali, ma solo in francese i caporali aperti sono seguiti da uno spazio inseparabile, e quelli chiusi sono preceduti da uno spazio inseparabile.

Per esempio, le norme ISO per la matematica tecnica impongono l'uso del corsivo per tutte le variabili che indicano delle grandezze misurabili, che siano scritte in lettere latine o greche non importa. Ma solo in francese le lettere greche in matematica vanno scritte non inclinate.

Altro esempio; in quasi tutte le lingue quando vengono usati i numeri romani, questi hanno il significato di ordinali, non di cardinali. Quindi mentre in italiano si scrive 'XX secolo' e si legge 'ventesimo secolo', in francese bisogna scrivere 'XX^e siècle' e si legge corrispondentemente 'vingtième siècle' e, attenzione, per certi ordinali l'apice va scritto distinguendo il maschile dal femminile, per esempio I^{er} e I^{ère}.

Ma tolte queste piccole diversità, la tipografia francese non differisce molto da quella italiana descritta sopra. La caratteristica maggiore è che in francese il primo capoverso di ogni sezione viene sempre obbligatoriamente rientrato e che gli spazi *dopo* i segni di interpunzione sono sempre costanti.

Queste particolarità del francese sono già tutte inglobate dentro le impostazioni che si ottengono specificando l'opzione `french` a `babel` o specificando la lingua francese a `polyglossia`.

Piuttosto in francese, che usa moltissimi accenti fonici oltre alla dièresi e al circonflesso a ai segni æ, Æ, œ e Œ, è essenziale usare la codifica UTF8 per il testo sorgente e bisogna selezionare font adeguatamente completi sia per comporre con `pdflatex` sia con `xelatex` o `lualatex`. Come è noto il francese usa una miriade di accenti fonici, e nessuno tonico, visto che l'accento tonico cade sempre sull'ultima vocale non muta o sull'ultimo dittongo; in definitiva il set di caratteri francesi è il seguente:

a à â æ b c ç d e è é ë ê f g h i ï î j k l
m n o ö ô œ p q r s t u ù ü û v w x y ÿ z

L'accento sulla 'a' serve per distinguere parole altrimenti omografe; sulla 'e' gli accenti grave e acuto ne segnano il suono aperto o chiuso; la dièresi sulle vocali 'e', 'i', 'o' e, specialmente, 'u' segna la separazione di un dittongo in uno iato, ma per la 'u' in particolare ne determina la funzione vocalica che altrimenti la lettera non avrebbe perché servirebbe per modificare il suono della lettera precedente. L'accento circonflesso viene usato in molti casi per allungare il suono di una vocale; spesso sostituisce una 's' dell'antico francese, come nel moderno 'même' dall'antico 'mesme'. Il francese ortograficamente usa consonanti doppie, ma le pronuncia senza rafforzamento, mentre, a differenza dell'italiano, distingue il suono lungo o breve delle vocali; 'patte' e "pâte" si pronuncerebbero nello stesso modo se nella seconda parola mancasse l'accento circonflesso che allunga il suono della 'a' e, analogamente, la doppia 't' nella prima parola accorcia il suono della 'a'.

Queste caratteristiche della lingua francese si riflettono sull'ortografia; peccato che l'ortografia resti un grosso problema anche per i francesi che c'è solo una modesta corrispondenza biunivoca fra lo scritto e il parlato.

4.3 TIPOGRAFIA ANGLO-AMERICANA

Le consuetudini anglo-americane sono molto simili grazie alla comune tradizione, ma differiscono in alcuni dettagli, che non riguardano solo l'ortografia delle parole simili; per esempio *theatre* e *theater*. Questo ha rilievo anche sulla sillabazione e sull'uso delle parole infisse. Ma queste cose sono già tenute in conto da `babel` e `polyglossia`; si ricorda che il sistema `TEX` è nato negli USA, quindi la specificazione `english` per la lingua vuol dire scegliere la sillabazione americana; l'opzione `english` è predefinita, quindi quando si compone in

qualunque lingua non specificata fra le opzioni per babel o non impostate per polyglossia, o senza i relativi pattern inseriti nel file di formato, l'eventuale censura in fine di riga viene eseguita con la lingua english; è facile immaginare il disastro compositivo che ne scaturisce.

Volendo comporre in inglese britannico si può specificare l'opzione per la lingua british a babel; meglio ancora, con l'una o l'altra variante è possibile specificare le lingue USenglish o UKenglish che dispongono di file di pattern più completi e di liste di eccezioni molto più ampie rispetto a quanto si può ottenere con i pattern originali predisposti per il Plain T_EX usato da Knuth stesso per comporre il T_EX-book nel 1984.

Ovviamente non basta selezionare la variante giusta per la sillabazione, bisogna anche scrivere il testo sorgente in vero inglese americano o in vero inglese britannico. Quindi è opportuno conoscere i piccoli dettagli in cui le due tradizioni tipografiche differiscono.

Va ricordato anche che per l'inglese americano esiste l'ottimo *Chicago manual of style* [Chicago University, 1994]; probabilmente esiste una pubblicazione simile anche per l'inglese britannico.

Le differenze maggiori fra le due tradizioni tipografiche sono minuzie ma si notano.

1. Le virgolette (alte) semplici e doppie sono usate in modo opposto: in inglese britannico una citazione è normalmente evidenziata con le virgolette semplici; se a sua volta essa ne contenesse una seconda essa verrebbe evidenziata con le virgolette doppie; in americano la prima citazione verrebbe racchiusa fra virgolette doppie e la citazione interna fra virgolette semplici.
2. Gli appellativi abbreviati sono chiusi con il punto fermo in americano e sono senza punto in britannico: Mr. Smith e Mr Smith; Dr. Johnes e Dr Johnes.
3. In inglese il numero del giorno precede il nome del mese e richiede la terminazione dell'ordinale, mentre l'americano richiede solo il numero che però segue il nome del mese; la versione australiana è una via di mezzo: in inglese britannico: 1st January 2015; in inglese australiano: 1 January 2015; in inglese americano: January 1, 2015.

Normalmente in americano gli spazi che vengono lasciati dopo i segni di interpunzione sono diversi a seconda del livello del segno: la virgola è seguita da un normale spazio interparola; il punto e virgola da uno spazio maggiorato del 50%; i due punti e i segni che terminano un periodo sono seguiti da uno spazio doppio o triplo. In inglese britannico si segue una via di mezzo fra la spaziatura francese e quella americana.

È per questo che un punto che segna la fine di una abbreviazione va seguito da un normale spazio interparola perché non funziona da segno di interpunzione; ed è per questo che i motori di composizione del sistema T_EX distinguono un punto che segue una lettera minuscola da quello che segue una lettera maiuscola; nel primo caso esso viene considerato punto fermo, mentre nel secondo esso viene considerato punto di abbreviazione. Allora per fare sì che un punto dopo una lettera maiuscola sia considerato come punto fermo, bisogna farlo precedere dal comando \@. Certo, scegliendo la spaziatura francese con `\frenchspacing` il problema non si pone nemmeno.

Sia l'inglese americano sia quello britannico non usano accenti, quindi in teoria i font originali di Knuth, i Computer Modern con codifica OT1 che contengono solo 128 caratteri nessuno dei quali accentato, potrebbero andare bene. In fondo il sistema T_EX è stato usato per almeno 10 anni senza lettere veramente accentate, ma simulate mediante due glifi, l'accento e la lettera di

base, sovrapposti mediante l'uso di opportune macro. Gli europei continentali e ogni altro utente di $\text{T}_{\text{E}}\text{X}$ e $\text{L}^{\text{A}}\text{T}_{\text{E}}\text{X}$ non anglofono si arrangiavano come potevano, ma era evidentemente una situazione insostenibile. Lo è rimasta fino al 1990, quando Knuth mise a punto $\text{T}_{\text{E}}\text{X}$ 3.0 che poteva gestire font con 256 caratteri e poco dopo comparvero le prime polizze di caratteri con codifica T1, finalmente in grado di soddisfare le necessità di tutti gli utenti sparsi per il mondo. Ma tutto ciò spiega l'evoluzione del sistema $\text{T}_{\text{E}}\text{X}$; non spiega perché sia opportuno *non* usare i font predefiniti Computer Modern nemmeno in inglese quando si usa `pdflatex`.

La spiegazione è semplice: oggi più di ieri si citano nomi propri scritti in lingue diverse dalla propria; si citano brani di testo più o meno lunghi in lingue diverse dalla propria; si compongono bibliografie che contengono anche riferimenti in lingue diverse dalla propria. Succede anche agli inglesi e agli americani. Bene le altre lingue usano diacritici; usare font senza caratteri completi dei diacritici necessari vuol dire comporre male e impedire all'algoritmo di sillabazione di funzionare al massimo delle sue capacità. Senza glifi contenenti lettere già accentate, ma ottenendo gli accenti per sovrapposizione di glifi diversi, una parola come 'elettricità', che in italiano potrebbe dividersi in 'elet-tri-ci-tà',² si divide solo in 'elet-tri-cit'; la corrispondente parola francese 'électricité', che potrebbe dividersi in 'élec-tri-ci-té',³ non si divide per niente a causa della prima 'é'. Ecco quindi l'opportunità di scegliere font completi di lettere accentate impostandone la codifica T1.

In ogni variante di inglese moderno non si usano accenti, quindi le 26 lettere `ASCII` sono sufficienti per scrivere in ogni variante di questa lingua:

a b c d e f g h i j k l m
n o p q r s t u v w x y z

Come si è già sottolineato, sebbene le 26 lettere latine siano sufficienti, si consiglia sempre l'uso della codifica UTF-8, perché anche nei testi inglesi compaiono nomi in lingue diverse che possono richiedere segni diacritici di vario genere. Si consiglia anche per tutte le varietà di inglese di usare font OpenType o Type 1 completi di lettere accentate e di evitare assolutamente i font Computer Modern e altri font codificati con la codifica OT1 che contiene solo l'alfabeto `ASCII` privo di accenti.

Si ricorda che l'opzione `english` per `babel` è sinonimo di `american`; se si vogliono la sillabazione britannica e le altre impostazioni relative all'inglese della Gran Bretagna bisogna specificare l'opzione `british`. Per polyglossia si possono impostare diverse varianti mediante le solite opzioni espresse con la sintassi `variant = <lingua>`; la variante americana è quella preimpostata, ma si possono specificare anche `usmax`, `british`, `ukmax`, `australian` e `newzealand`. Le varianti `usmax` e `ukmax`, rispettivamente per la variante americana e britannica, invocano file di pattern più completi e liste di eccezioni molto estese. A seconda della variante scelta, queste due impostazioni sono quelle più consigliabili.

4.4 TIPOGRAFIA TEDESCA, AUSTRIACA, SVIZZERA

La tipografia tedesca differisce un poco da quella delle altre nazioni dell'Europa occidentale per alcune particolarità specifiche della lingua e per alcune tradizioni tipografiche. Esistono anche delle varianti ortografiche differenti

²Si tenga presente la lunghezza minima delle sillabe iniziale e finale, che in italiano sono entrambe di due lettere.

³In francese le lunghezze minime preimpostate delle sillabe iniziale e finale sono rispettivamente di due e tre lettere, ma nessuno vieta di impostare la lunghezza della sillaba finale a due lettere.

da nazione a nazione; per esempio ‘gennaio’ si chiama ‘Januar’ in Germania e ‘Jänner’ in Austria.

Inoltre verso al fine del XX secolo è stata definita in Germania la “nuova ortografia” che nelle intenzioni vorrebbe semplificare un poco l’ortografia di un certo numero di parole; sembra che abbia avuto poco successo nell’uso pratico, ma è comprensibile visto che ci vogliono tempi lunghi per modificare abitudini di decenni. In ogni caso babel e polyglossia possono seguire le regole della nuova o della vecchia ortografia a seconda di quale nome venga usato per specificare la lingua a babel o quale variante si specifichi impostando la lingua a polyglossia. Ovviamente i due pacchetti si occupano delle parole infisse e della cesura in fine di riga, ma è compito dell’utente immettere il testo sorgente con l’ortografia corrispondente a quella specificata all’uno o all’altro pacchetto.

Il problema della cesura in fine di riga in tedesco è complesso; i pattern cercano di dividere secondo l’etimologia delle parole sovente composte agglutinando più di due termini. I pattern per svolgere questo lavoro sono numerosissimi e sono senza dubbio fra i più complessi di cui il sistema T_EX sia dotato. Ma mentre i pattern riescono a dividere lunghe parole composte in modo sostanzialmente corretto nella maggior parte dei casi, esistono delle situazioni in cui la scomposizione di certe parole composte implica un cambio di ortografia vuoi per l’aggiunta di un’ulteriore lettera, vuoi per la sostituzione di una lettera con un’altra.

In questa situazione l’utente è chiamato ad intervenire con l’uso del doppio apice ASCII ” attivo. Nella pagina 16 sono illustrate le funzionalità di questo carattere attivo per il tedesco; ci si rende conto, quindi che anche con polyglossia è assolutamente necessario impostare l’opzione babelshorthands, altrimenti la sillabazione tedesca potrebbe contenere errori significativi.

Il tedesco usa la spaziatura fissa dopo ogni segno di interpunzione e il comando `\frenchspacing` per ottenere questo scopo è già incluso nelle impostazioni per il tedesco sia con babel sia con polyglossia. Per queste lingue bisogna usare molto il segno diacritico chiamato *umlaut* o *trema*; a stretto rigore le due parole non sono sinonimi di ‘dieresi’; esse infatti sostituiscono una piccola ‘e’ che nei manoscritti medievali scritti con la calligrafia *fraktur* si scriveva sopra certe vocali ma di fatto si riduceva a due trattini o due puntini; questa ‘e’ riappare oggi quando si debbano usare font che non contengono i glifi con i due puntini, e quindi si scrive ‘Universitaet’ invece di ‘Universität’, ‘Guenter’ invece di ‘Günter’, eccetera. Nel comporre testi non sarebbe uno stile elegante quello di sostituire la umlaut con la ‘e’ posposta. In definitiva l’alfabeto tedesco diventa:

a ä b c d e f g h i j k l m n
o ö p q r s ß t u ü v w x y z

La legatura della ‘s lunga’ con una ‘s corta’, ß, si usa solo scrivendo con lettere minuscole; in maiuscolo diventa una doppia ‘S’: ‘SS’.

Come già detto le virgolette tedesche sono particolari e si usano solo in Germania e in Austria, mentre nella Svizzera tedesca si usano i caporali.

Infatti una particolarità del tedesco sono le virgolette aperte e chiuse. Quelle aperte si ottengono con “ e quelle chiuse con ”; non sembra nulla di strano, ma quelle aperte sono al piede delle lettere e con la concavità a sinistra, quelle chiuse sono in alto ma con la concavità a destra: „deutsche Sprache“. Per scrivere in svizzero-tedesco sono disponibili anche i comandi “< e ”> che producono i caporali non spaziati: «Schweizerische Sprache», ma sta all’utente non lasciare spazi spuri (diversamente dall’italiano dove eventuali spazi spuri sono eliminati quando si usino gli stessi comandi).

4.5 TIPOGRAFIA LATINA

La tipografia latina propriamente detta non esiste; oggi, però, si possono scrivere edizioni critiche di testi latini; si possono scrivere testi didattici di latino comprensivi di brani antologici sufficientemente lunghi; si possono comporre dizionari monolingui e multilingui che comprendano anche il latino; si possono scrivere libri devozionali anche completamente in latino. Dire quindi che non esiste la tipografia latina è una affermazione un po' forte. Non dimentichiamo poi che il latino moderno è la lingua ufficiale dello Stato della Città del Vaticano, piccolo stato i cui testi vengono però letti in tutto il mondo.

I programmi di composizione del sistema \TeX permettono di scegliere fra quattro forme diverse di composizione; inoltre esistono pacchetti per estenderne ancora le funzionalità compositive.

Le quattro forme sono quella moderna, quella medievale, quella classica e quella ecclesiastica; con babel si scelgono impostandone l'opzione `latin` e ulteriormente specificando dei modificatori e degli attributi come spiegato nel paragrafo 2.3 a pagina 17. In quel paragrafo si è mostrato anche come impostare le varianti per impostare la lingua latina per polyglossia. I modificatori e gli attributi e le varianti hanno gli stessi nomi; essi sono i seguenti.

`modern` è l'impostazione predefinita; l'ortografia prevede l'uso della 'u' e della 'v' minuscole e maiuscole come nelle altre lingue moderne dell'Europa occidentale; non usa i dittonghi legati \ae e \oe né minuscoli né maiuscoli; la divisione in sillabe è fonetica come in italiano. Le parole infisse sono coerenti con l'ortografia appena descritta, ma l'ortografia del testo da comporre spetta all'utente.

`medieval` serve per comporre testi medievali; l'ortografia prevede l'uso della sola 'u' minuscola e della sola 'V' maiuscola, quindi la 'u' ha valore sia di vocale sia di consonante; nello stesso tempo la 'V' è impostata come maiuscola della 'u'; quindi quando le parole sono trasformate in tutte maiuscole (per esempio nelle testatine prodotte da certe classi) viene eseguita la giusta trasformazione. Usa i dittonghi legati \ae e \oe , anche se in molti testi sono contratti in 'e'; spesso la 'I' maiuscola iniziale è sostituita con 'Y'; la divisione in sillabe è fonetica e cerca di distinguere la funzione consonantica della 'u' dalla sua funzione vocalica. Le parole infisse sono coerenti con l'ortografia descritta, ma l'ortografia del testo da comporre è lasciata all'utente.

`classic` serve per comporre testi classici; l'ortografia prevede l'uso della 'u' minuscola e della corrispondente 'V' maiuscola, ma non usa i dittonghi legati; i latinisti generalmente pronunciano questi dittonghi sciolti e pronunciano la 'u' consonantica come una 'u' semiconsonantica italiana. La divisione in sillabe è etimologica e tiene anche conto delle desinenze della declinazione e della coniugazione; molti gruppi consonantici sono divisi fra le sillabe adiacenti in modo decisamente diverso dal latino moderno e da quello medievale. Cerca in tutti i modi di non considerare dittonghi quelli che in italiano sarebbero considerati tali. Le parole infisse sono coerenti con l'ortografia descritta, ma è compito dell'utente scrivere correttamente il testo da comporre.

`ecclesiastic` serve per comporre il latino ecclesiastico quale viene usato in alcuni testi di teologia e nei testi devozionali della Chiesa Cattolica. Questa scrittura è sostanzialmente quella del latino moderno con l'aggiunta dei dittonghi legati \ae e \oe ; inoltre le parole che presentano problemi di accentazione da parte di lettori di diverse culture e lingue madri sono accentate, se occorre anche sui dittonghi legati.

Si noti che *œ* non esiste in nessuna codifica, nemmeno con la codifica UNICODE. Con `pdflatex` e `babel` non ci sono problemi perché basta usare la normale sintassi del sistema `TEX`: `\'oe o`, con la codifica d'entrata `utf8`, `\'æ`. Con `polyglossia` e i font `OpenType` la sintassi tradizionale, per esempio `\'a` funziona regolarmente, perché quella macro `\'` seguita dalla *'a'*, produce direttamente il segno *'á'* che esiste nella codifica `utf8`; invece `\'æ` non funziona perché la macro seguita da *'œ'* non può produrre direttamente il segno *œ* che non esiste nella codifica `utf8`. Poco più avanti si vedrà come il problema è stato risolto.

Il latino ecclesiastico prevede una spaziatura alla francese per i segni di interpunzione alti, ma con spaziature più ridotte; prevede anche la composizione delle note al piede in stile francese con il richiamo numerico sulla linea di base, non ad esponente.

Con `pdflatex` è possibile usare il pacchetto `liturg` per comporre i testi devozionali e i testi per la liturgia, messali e breviari e provvede anche a quei segni speciali che nei testi liturgici marcano le parti lette dal celebrante e quelle che costituiscono le risposte dei fedeli. Per compatibilità con il passato invece del modificatore `ecclesiastic`, che è disponibile solo dalla versione 3.5 del modulo di `babel` per il latino, si può caricare il pacchetto `ecclesiastic`, ma si sconsiglia questo uso con versioni aggiornate dell'installazione del sistema `TEX`. Con `XlATEX` e `LuaATEX` il modulo per la lingua latina che usa `polyglossia` è stato aggiornato e prevede la variante `ecclesiastic`, ma forse sarà disponibile solo a partire dagli aggiornamenti del 2016. Nel frattempo bisogna fare di necessità virtù e accontentarsi di quello che si può fare con i font `OpenType` *senza caricare il pacchetto* `ecclesiastic` che non funziona con quei programmi perché non è costruito per lavorare con i font `OpenType`.

Per il resto lo stile tipografico del latino ecclesiastico non differisce granché da quello del latino moderno; solo i messali da comporre con il pacchetto `liturg` richiedono delle particolarità tipografiche di cui si occupa il pacchetto citato. Si consiglia la lettura della documentazione di questo pacchetto per prendere familiarità con i comandi specifici necessari per la composizione di quei testi.

Riassumendo, la situazione per il latino è la seguente.

1. In latino moderno usato normalmente non si usano accenti né dittonghi legati; se in latino moderno si usa la notazione ecclesiastica compaiono i dittonghi legati e possono venire usati gli accenti acuti su qualsiasi lettera e dittongo. Per comporre grammatiche, dizionari, edizioni critiche, testi filologici, in latino moderno possono apparire anche i segni di *'lunga'* (o *'macron'*) e *'breve'*.
2. In latino medievale si usano i dittonghi legati, talvolta anche la *'y'* iniziale anche in parole non di origine greca; non compare la *'v'* minuscola né la *'U'* maiuscola. Talvolta appaiono accenti su alcune preposizioni e sull'ablativo singolare di alcune parole.
3. In latino classico non si usano i dittonghi legati né la *'v'* minuscola e la *'U'* maiuscola; nelle edizioni critiche potrebbero apparire i segni prosodici di *'lunga'* e *'breve'*; il latino classico non usava la punteggiatura ma usava la *lectio continua*; Quintiliano stesso si lamentava di questo fatto e suggeriva ai futuri oratori di dividere la parti del discorso da leggere con i segni di comma, colon e periodus, rappresentati rispettivamente con i tre segni `. . .`. Oggi anche le edizioni critiche probabilmente non verrebbero composte in *lectio continua* e verrebbe usata la punteggiatura moderna.

In totale l'alfabeto latino conterrebbe molte lettere:

a á â ã ä å æ b c d e é ē ě f g h i (j) k l m
n o ó ô õ ö œ p q r s t u ú û ü v (w) x y z

di cui la 'j' e la 'w' verrebbero usate solo per certi nomi propri di origine diversa dal latino. Va notato che il carattere œ non è presente in nessuna codifica, nemmeno nella codifica UNICODE, quindi va ottenuto con una macro che sovrappone l'accento acuto al dittongo legato oe.

La macro usata in questo testo è la seguente:

```
\DeclareRobustCommand\eccla@lift@acute[2]%
[\char\string"00B4]%          default acute accent
{\leavevmode
{\edef\slant@{\strip@pt\fontdimen1\font}%
\setbox\z@\hbox{#1}%          accent
\setbox\tw@\hbox{#2}%          base character
\@tempdima=\dimexpr\ht\tw@-0.9ex\relax
\dimen@=\dimexpr0.6\wd\tw@+\slant@\@tempdima-0.5\wd\z@\relax
\raise\@tempdima\hbox to\z@{\kern\dimen@\box\z@
\relax\hss}\box\tw@}}

\newcommand\oeacute{\eccla@lift@acute{\oe}}
\newcommand\OEacute{\eccla@lift@acute{\OE}}
```

e vale per comporre con xelatex e luatex; le modifiche per creare una macro valida per comporre con pdflatex sono molto più semplici perché anche il dittongo legato æ va trattato con la macro per gli accenti, cioè '\ae e '\oe.

Il fatto che alcune poche lettere accentate debbano essere trattate con le macro per gli accenti implica che la sillabazione eventualmente necessaria ne risulta un po' impedita e all'occorrenza richiede interventi manuali.

4.6 TIPOGRAFIA GRECA

La tipografia greca moderna ricorre quasi esclusivamente alla scrittura monotonica.

GRECO MODERNO MONOTONICO In questa scrittura le parole polisillabiche contengono un solo segnaccento che generalmente ha la forma di un accento acuto e omette tutte quella varietà di segni diacritici usati nel greco moderno politonico e nel greco antico. Conserva solo la dièresi che rimane anche quando si trasforma una parola in tutte maiuscole, mentre il 'tonos' (l'accento monotonico) in questa trasformazione scompare. La divisione in sillabe è relativamente semplice, nel senso che vale la regola, come in italiano, che un gruppo di consonanti si divide fra due sillabe adiacenti lasciando a destra della cesura il maggior numero di lettere che possano trovarsi all'inizio di una parola greca.

GRECO MODERNO POLITONICO Il greco moderno politonico conserva gli spiriti dolce e aspro, i tre segnaccenti acuto, grave e circonflesso, e la dièresi, oltre alla iota sottoscritta che diventa adscritta (cioè scritta come un pedice) quando la vocale che la contiene viene trasformata in maiuscola. Tutti gli accenti e la iota spariscono quando una parola viene trasformata in tutte maiuscole; solo la dièresi rimane per marcare uno iato anche perché la pronuncia dello iato è generalmente diversa da quella del dittongo, specialmente per quel che riguarda la upsilon che si pronuncia 'i' nello iato ma 'v' nel dittongo. Il greco moderno politonico, così come il monotonico usa una moltitudine di digrafi strani per rendere i suoni

delle parole straniere entrate nell'uso moderno ma che non sono riproducibili con la pronuncia moderna delle lettere greche. Si pensi soltanto alla lettera 'β' che modernamente si pronuncia sempre come una 'v' italiana; per rendere il suono della 'b' italiana in greco moderno bisogna ricorrere al digrafo 'μπ' che però talvolta si pronuncia 'b' (ed è indivisibile fra sillabe adiacenti) e talvolta 'mp' o 'mb' (ed è divisibile fra sillabe adiacenti). La divisione in sillabe è piuttosto complicata non tanto per le regole, quanto per la mobilità degli accenti nella declinazione e nella coniugazione.

GRECO ANTICO Il greco antico nella scrittura assomiglia molto al greco moderno politonico, ma le parole sono generalmente diverse e non viene eseguito nessun tentativo di riprodurre i suoni moderni delle parole straniere entrate nell'uso quotidiano e che mancano nel greco antico vero e proprio. La divisione in sillabe è complessa non solo a causa della mobilità degli accenti ma anche perché bisogna seguire l'etimologia delle parole

Ecco quindi che sia con babel sia con polyglossia bisogna essere in grado di distinguere le tre forme linguistiche con i modificatori mono o monotoniko, poly o polutoniko, e ancient; le varianti hanno gli stessi nomi.

Per il resto sia babel sia polyglossia si occupano di tutto quel che occorre per la tipografia greca.

Ma con pdf_{latex} c'è un altro problema; i linguisti sparsi per il mondo, che verosimilmente non usano una tastiera greca, possono immettere il testo greco mediante macro scritte con caratteri latini, oppure sono semplicemente corrispondenze fra le 26 lettere latine e le 25 lettere greche. In realtà disponendo di un sistema operativo che consenta di disporre di una tastiera virtuale, l'immissione diretta del testo greco consiste solamente nell'immettere le lettere greche semplici o con diacritici mediante la tastiera virtuale.

Nel paragrafo 2.4 a pagina 19 si sono già mostrate le corrispondenze fra le lettere latine e quelle greche, nonché le macro per inserire i diacritici. Qui merita sottolineare che ci sono due pacchetti molto utili, che in teoria si potrebbero usare anche con x_{el}at_{ex} e polyglossia, e che permettono di scrivere in greco senza disporre di una tastiera greca vera o virtuale.

I pacchetti sono textalpha e alphabeta. Il primo rende usabili tutti i diacritici inseriti mediante macro e definisce il comando `\ensuregreek` il cui argomento viene composto in greco anche se non è stata specificata la lingua greca, ma è stata solo specificata l'opzione LGR al pacchetto fontenc prima di specificare l'opzione T1. Tutti i comandi letterali per gli accenti diventano quindi accessibili con le macro `\<`, `\>`, `\'`, `\`, `\~`, eccetera. I primi quattro appena elencati hanno anche altri significati in alcuni ambienti, quindi vengono ridefiniti in modo da non collidere con le altre definizioni. Il secondo consente di immettere sia in modo testuale sia in modo matematico le lettere greche; usa gli stessi comandi che L^AT_EX definisce solo per la matematica, `\alpha`, `\beta`, ..., `\omega`, ma il pacchetto alphabeta ne estende l'uso anche al testo; nel testo compone questi simboli nello stesso stile dei font latini circostanti, diritti o inclinati, di nerezza media o decisamente neri; in matematica li disegna sempre diritti. Questa funzionalità può essere molto utile anche nella tipografia francese per usare i simboli greci diritti secondo la tradizione francese.

Il pacchetto textalpha è essenziale quando si scrive in greco politonico o antico senza disporre di una tastiera greca. Il secondo pacchetto alphabeta è molto utile per risparmiare battute quando si vogliono inserire lettere greche in modo testo, ma non si vuole specificare la lingua greca; è utile anche per i bookmark creati con il pacchetto hyperref, perché di per sé questo pacchetto non accetterebbe i caratteri greci specificati con la codifica UTF-8. Si consi-

glia di esaminare con attenzione la documentazione di questi due pacchetti mediante i comandi:

```
texdoc textalpha
texdoc alphabeta
```

immessi un una finestra ‘terminale’, ‘prompt dei comandi’, ‘xterm’, comunque si chiami con il sistema operativo in uso.

Per il greco abbiamo già detto che non scenderemo nei dettagli perché i caratteri necessari sono già tutti disponibili nei font OpenType e lo sono quasi tutti anche quando si compone con i caratteri Type 1 con codifica LGR. Si tenga presente che questi ultimi caratteri sono stati creati per l’uso con pdf_latex avendo i filologi in mente; non è escluso che in futuro possano essere riveduti e corretti per migliorare le loro prestazioni, ma certamente non sarà possibile avere un un solo font con tutti i circa 300 glifi che sarebbero necessari.

Va segnalato che la tipografia greca non ha grosse differenze rispetto alle tradizioni italiane o inglesi; usa i caporali come virgolette; nello scrivere in greco monotonico usa gli stessi segni di interpunzione che si usano per le lingue occidentali scritte in lettere latine. L’unica cosa diversa è che ammette la cesura in fin di riga lasciando prima della cesura anche una sillaba di una sola lettera, così come lo consente per l’ultima sillaba dopo la cesura.

Ovviamente, però, le impostazioni per babel e le funzionalità dei font OpenType quando si usa polyglossia, prevedono l’immissione diretta del testo greco nel file sorgente. È il solito problema della tastiera già discusso e che con i sistemi moderni Mac, Windows e Linux che dispongono di un visualizzatore interattivo della tastiera permette anche a chi non dispone di una vera tastiera greca politonica di scrivere direttamente in greco nel file sorgente.

4.7 TIPOGRAFIA SPAGNOLA E MESSICANA

La tipografia spagnola differisce poco da quella italiana, a parte la lingua; ma per impostazione predefinita è sempre in vigore la spaziatura costante dopo i segni di interpunzione.

Altre piccole differenze si hanno per le liste numerate dove diversamente dall’inglese, si usa il numero puntato per il primo livello, una lettera minuscola corsiva seguita da parentesi per il secondo; un numero seguito da parentesi per il terzo; una minuscola corsiva con un apice seguito da parentesi per il quarto: 1. a) 1) a’). La numerazione alfabetica delle liste include 27 segni, perché oltre ai normali 26 segni dell’alfabeto latino viene usata anche la lettera ñ.

Analogamente per le liste puntate, i simboli predefiniti sono diversi da quelli preimpostati per l’inglese; per lo spagnolo essi sono: ■ • ◦ ◇.

I richiami di nota non numerici o letterali sono ottenuti con una serie di asterischi, uno per la prima nota, due per la seconda, eccetera. Si suppone che note marcate in questo modo non siano numerose.

L’opzione per la lingua spagnola di babel offre, fra gli altri, i modificatori `mexico` e `mexico-com`, che permettono di adattare lo stile alle consuetudini usate in Messico; il secondo modificatore viene descritto come probabilmente più adatto in altri paesi dell’America centrale e meridionale.

In spagnolo non si usano i numeri romani minuscoli prodotti normalmente dal comando `\roman`; li si compone in maiuscoletto e a tale scopo il comando `\roman` è ridefinito allo scopo. Tuttavia sembra che questa ridefinizione sia poco gradita al programma `makeindex`, il programma che predispone l’indice analitico a partire da una raccolta grezza di dati. Per cui la documentazione dell’opzione `spanish` per babel descrive un modo di aggirare questo problema. Chi scrive ha invece trovato diversi altri problemi con l’uso del pacchetto

hyperref specialmente se viene simultaneamente invocata la lingua greca. La documentazione dell'opzione `spanish` non riporta invece nessun problema in merito ai collegamenti ipertestuali.

La documentazione dell'opzione `spanish` indica molte altre funzionalità offerte dai pacchetti `babel` e `polyglossia`; sono cose che agevolano l'utente, ma non portano sostanziali modifiche allo stile tipografico.

Le tradizioni tipografiche spagnole e messicane, oltre che del resto dell'America latina, non sono tanto diverse da quelle europee in generale. Certo l'insieme di lettere dell'alfabeto comprende anche la 'ñ' intercalata alfabeticamente dopo la 'n'; si usa sempre l'accento tonico per le parole diverse dalle parole piane; nei numeri in Messico e negli altri paesi dell'America centrale e meridionale si usa più frequentemente il punto rispetto alla virgola, grazie alla vicinanza e all'influenza del mondo anglofono nordamericano, molto di più da quando il Bureau International des Poids et Mesures non obbliga più ad usare solo la virgola per tutto il mondo non anglofono, e il punto solo per il mondo anglofono. Componendo in spagnolo si abbia l'accortezza di usare solo e sempre lo stesso segno, possibilmente la virgola, che va trattata intelligentemente visto che anche in matematica essa è preimpostata come segno di interpunzione, non come separatore decimale. I moduli per gestire la lingua spagnola sia per `babel` sia per `polyglossia` provvedono alla gestione intelligente della virgola in una maniera simile a quella che viene usata per l'italiano. Inoltre lo spagnolo usa sistematicamente i punti esclamativo e interrogativo capovolti per introdurre un'esclamazione o una domanda

L'insieme di caratteri dello spagnolo è il seguente:

a á b c d e é f g h i í ï j k l m n
ñ o ó p q r s t u ú ü v w x y z ¿ ¡

e certamente l'uso della codifica UTF-8 per immettere i caratteri si rivela opportuno oltre che necessario quando si usino i programmi `xelatex` e `luaLaTeX`.

A titolo d'esempio si riportano alcuni articoli della costituzione messicana.

ARTÍCULO 10.-

En los Estados Unidos Mexicanos todo individuo gozará de las garantías que otorga esta Constitución, las cuales no podrán restringirse ni suspenderse, sino en los casos y con las condiciones que ella misma establece.

Está prohibida la esclavitud en los Estados Unidos Mexicanos. Los esclavos del extranjero que entren al territorio nacional alcanzarán, por este solo hecho, su libertad y la protección de las leyes.

Queda prohibida toda discriminación motivada por origen étnico o nacional, el género, la edad, las capacidades diferentes, la condición social, las condiciones de salud, la religión, las opiniones, las preferencias, el estado civil o cualquier otra que atente contra la dignidad humana y tenga por objeto anular o menoscabar los derechos y libertades de las personas.

ARTÍCULO 20.- LA NACIÓN MEXICANA ES ÚNICA E INDIVISIBLE.

La Nación tiene una composición pluricultural sustentada originalmente en sus pueblos indígenas que son aquellos que descienden de poblaciones que habitaban en el territorio actual del país al iniciarse la colonización y que conservan sus propias instituciones sociales, económicas, culturales y políticas, o parte de ellas.

La conciencia de su identidad indígena deberá ser criterio fundamental para determinar a quiénes se aplican las disposiciones sobre pueblos indígenas.

Son comunidades integrantes de un pueblo indígena, aquellas que formen una unidad social, económica y cultural, asentadas en un territorio y que reconocen autoridades propias de acuerdo con sus usos y costumbres.

El derecho de los pueblos indígenas a la libre determinación se ejercerá en un marco constitucional de autonomía que asegure la unidad nacional. El reconocimiento de los pueblos y comunidades indígenas se hará en las constituciones y leyes de las entidades federativas, las que deberán tomar en cuenta, además de los principios generales establecidos en los párrafos anteriores de este artículo, criterios etnolingüísticos y de asentamiento físico.

4.8 TIPOGRAFIA CATALANA

Il catalano come lingua è decisamente diversa dal castigliano, come gli spagnoli chiamano la loro lingua ‘pura’. Ovviamente ne è diversa anche l’ortografia. Ma le tradizioni tipografiche catalane non differiscono sostanzialmente da quelle del resto della Spagna.

Una particolarità del catalano è che usa il digrafo ‘ny’ per indicare quello che in castigliano si rappresenta con la ‘ñ’; ma la caratteristica specifica è che il catalano prevede sia il suono e l’ortografia identiche con il castigliano per quel che riguarda il suono rappresentato dal digrafo ‘ll’, ma prevede anche la ‘l’ rafforzata, quella che in italiano è rappresentata dalla ‘ll’. Per separare questi due usi della doppia ‘l’ e non fare confusione, la ‘l’ rafforzata si rappresenta con ‘l·l’. Il modulo usato da babel per rappresentare questo speciale ‘trigrafo’ dispone di ben tre macro per comporre questo particolare segno: `\l·l`, `\lgem`, `"l`, replicate in altrettante varianti con le ‘L’ maiuscole: `\L·L`, `\Lgem`, `"L`. Con polyglossia sono disponibili solo le macro `"l`, `"L`, `\lgem` e `\Lgem`. Si noti che il trigrafo speciale ‘l·l’ non esiste nemmeno nella codifica UNICODE, quindi va sempre costruito con una macro. In caso di cesura in corrispondenza della ‘l geminata’ i moduli della lingua catalana provvedono automaticamente mediante un comando `\discretionary` adeguatamente configurato per lasciare il punto a mezza riga quando non c’è cesura, ma sostituendolo con il trattino breve quando avviene la cesura; invece il digrafo ‘ll’ non viene diviso in fine di riga.

In definitiva i segni che compaiono nell’alfabeto usato in catalano sono i seguenti:

a à á b c ç d e è é f g h i í ï j k l
ll l·l m n o ò ó p q r t u ú ü v w x y z

La codifica UTF-8 continua ad essere opportuna con `pdflatex` e necessaria con `xelatex` e `lualatex`.

A titolo di esempio di un testo catalano si riporta l’articolo 6 dello *Estatut d’autonomia*. Il testo di questo articolo mostra praticamente tutti i segni usati per comporre in catalano.

ARTICLE 6. LA LLENGUA PRÒPIA I LES LLENGÜES OFICIALS

1. La llengua pròpia de Catalunya és el català. Com a tal, el català és la llengua d’ús normal i preferent de les administracions públiques i dels mitjans de comunicació públics de Catalunya, i és també la llengua normalment emprada com a vehicular i d’aprenentatge en l’ensenyament.

2. El català és la llengua oficial de Catalunya. També ho és el castellà, que és la llengua oficial de l'Estat espanyol. Totes les persones tenen el dret d'utilitzar les dues llengües oficials i els ciutadans de Catalunya tenen el dret i el deure de conèixer-les. Els poders públics de Catalunya han d'establir les mesures necessàries per a facilitar l'exercici d'aquests drets i el compliment d'aquest deure. D'acord amb el que disposa l'article 32, no hi pot haver discriminació per l'ús de qualsevol de les dues llengües.
3. La Generalitat i l'Estat han d'emprendre les accions necessàries per al reconeixement de l'oficialitat del català a la Unió Europea i la presència i la utilització del català en els organismes internacionals i en els tractats internacionals de contingut cultural o lingüístic.
4. La Generalitat ha de promoure la comunicació i la cooperació amb les altres comunitats i els altres territoris que comparteixen patrimoni lingüístic amb Catalunya. A aquests efectes, la Generalitat i l'Estat, segons que correspongui, poden subscriure convenis, tractats i altres mecanismes de col·laboració per a la promoció i la difusió exterior del català.
5. La llengua occitana, denominada aranès a l'Aran, és la llengua pròpia d'aquest territori i és oficial a Catalunya, d'acord amb el que estableixen aquest Estatut i les lleis de normalització lingüística.

4.9 TIPOGRAFIA PORTOGHESE E BRASILIANA

I file di descrizione della lingua distinguono lo stile portoghese da quello brasiliano; le differenze sono minime in poche parole infisse, mentre sono visibili nei nomi dei mesi che in portoghese sono con l'iniziale maiuscola, mentre in brasiliano sono in tutte minuscole.

Con il carattere attivo `"`, vi sono le macro `"< e ">`, per inserire i caporali, ma si tratta solamente di "abbreviazioni" per non scrivere `\guillemotleft` e `\guillemotright` forniti dal pacchetto `textcomp` quando si compone con `pdflatex`. In questo caso, però, le legature della codifica T1, `<< e >>`, sarebbero ancora più semplici e facili da leggere nel file sorgente.

La cosa curiosa, invece, è che il file che produce le macro per scrivere in portoghese fornisce anche un comando per disattivare le legature; infatti sembra che le legature `'ff'`, `'fi'` e `'fl'`, e quelle associate, siano da comporre slegate; siccome la legatura dipende dai font usati, in quanto in molti font sono assenti, questa macro per il portoghese è particolarmente adatta quando si usa `pdflatex`. Usando `xelatex` o `lualatex` e i font OpenType, esistono comandi specifici per disabilitare le legature; si rinvia alla documentazione del pacchetto `fontspec` che permette di gestire correttamente quei tipi di font.

Una cosa che non è menzionata né nella documentazione dell'opzione portoghese per `babel` né in quella di `polyglossia` è che nelle citazioni in linea con il testo, la citazione stessa è racchiusa fra virgolette caporali, ma se la citazione dura più righe, allora le virgolette caporali aperte sono ripetute all'inizio di ogni riga successiva alla prima fino alla fine della citazione. Questa funzionalità non sembra essere presente né con `babel` né con `polyglossia`.

Il portoghese e il brasiliano condividono la scrittura e le tradizioni tipografiche. L'Accademia Portoghese per la Lingua si occupa di mantenere stabile la lingua portoghese nella madrepatria e nelle ex colonie dove il portoghese è lingua ufficiale; il Brasile è in testa alla classifica, ovviamente, anche se ha ottenuto l'indipendenza dal Portogallo nel 1822 e oggi fa un certo senso chiamarlo ex colonia, anche perché fu il principe reggente della Corona del Portogallo a dichiarare l'indipendenza prendendo il titolo di Pedro I del Brasile.

A parte la storia, la lingua brasiliana, pur essendo “gestita” da un’unica Accademia della Lingua, si pronuncia in modo diverso da come si pronuncia in Portogallo e ha un lessico ricchissimo di parole assimilate dalle lingue delle popolazioni indigene e degli africani deportati in Brasile nei secoli in cui la tratta degli schiavi era diffusa in tutta l’America del nord e del sud.

Il portoghese/brasiliiano usa molti accenti oltre la ç; il suo alfabeto è il seguente:

a à á â ã b c ç d e é ê f g h i í j k l
m n o ò ó ô õ p q r s t u ü v w x y z

È quindi evidente che registrare i file sorgente con la codifica UTF-8 è molto importante; con pdf_latex è importante richiamare il pacchetto inputenc con l’opzione utf8.

Ecco a titolo di esempio alcuni articoli della costituzione portoghese dove sono usate molte vocali accentate. Vi si notano anche alcune particolarità come gli ordinali indicati sia dal punto fermo, sia dall’apice maschile (o, in altri contesti, femminile).

ARTIGO 1.º (REPÚBLICA PORTUGUESA) Portugal é uma República soberana, baseada na dignidade da pessoa humana e na vontade popular e empenhada na construção de uma sociedade livre, justa e solidária.

ARTIGO 2.º (ESTADO DE DIREITO DEMOCRÁTICO) A República Portuguesa é um Estado de direito democrático, baseado na soberania popular, no pluralismo de expressão e organização política democráticas, no respeito e na garantia de efectivação dos direitos e liberdades fundamentais e na separação e interdependência de poderes, visando a realização da democracia económica, social e cultural e o aprofundamento da democracia participativa.

ARTIGO 3.º (SOBERANIA E LEGALIDADE)

1. A soberania, una e indivisível, reside no povo, que a exerce segundo as formas previstas na Constituição.
2. O Estado subordina-se à Constituição e funda-se na legalidade democrática.
3. A validade das leis e dos demais actos do Estado, das regiões autónomas, do poder local e de quaisquer outras entidades públicas depende da sua conformidade com a Constituição.

4.10 TIPOGRAFIA RUMENA

L’ortografia rumena fa uso di pochi segni diacritici che però usa frequentemente; tanto per citare una parola che vuol dire ‘insegnamento’, *învăţământ*, ha tutte le sue sillabe con almeno un segno diacritico: 4 sillabe, 5 diacritici. Ma i diacritici distinti sono solo tre: il circonflesso, la breve, e la ‘virgola sotto’ la ‘t’; non si tratta di una cediglia, ma di una virgola ribassata che viene posta solo sotto le lettere ‘t’ ed ‘s’ minuscole e maiuscole; il segno della breve si usa solo sulla ‘a’ minuscola e maiuscola, il segno del circonflesso si usa solo sulla ‘a’ e sulla ‘i’, sempre minuscole e maiuscole. Marcano il modo di pronunciare la lettera di base, ma non indicano nessuna posizione di accento tonico.

Per il resto la sillabazione del rumeno non pone difficoltà particolari, visto che tutto sommato si tratta di una lingua neolatina, ma gli adstrati slavi e turchi sono numerosi, per cui non è difficile trovare parole con gruppi di consonanti insolite.

I pacchetti babel e polyglossia trovano nei file di formato dei programmi di composizione i pattern adeguati sia per l'uso con font a 256 caratteri sia per i font OpenType. Quindi non ci dovrebbero essere problemi particolari nelle tradizioni tipografiche rumene da realizzare mediante i questi programmi.

La documentazione per l'opzione `romanian` per babel risale al 2005, quindi è un po' vecchiotta e non è stata nemmeno aggiornata per richiamare il pacchetto `combelow` disponibile dal 2010, con il quale è possibile creare i segni della 's' e della 't' con la 'virgola sotto', come è prescritto per la lingua rumena. Il pacchetto che contiene le macro dell'opzione `romanian` contengono il minimo necessario, senza nessuna impostazione stilistica ma con la cediglia al posto della virgola sotto.

Il file che contiene le macro per la composizione del rumeno con polyglossia contiene lo stesso il minimo necessario ma sembra che contenga almeno i diacritici giusti per la 's' e la 't'.

Questa situazione ampiamente insoddisfacente per una lingua che si scrive con l'alfabeto latino esteso, mostra che non ci sono tutte rose e fiori con il sistema \TeX ; tuttavia mostra anche che l'uso di `xelatex` o `lualatex` con il pacchetto polyglossia e i font OpenType consente almeno una composizione decorosa. Per il resto bisogna che chi usa queste indicazioni per citare brani di testo rumeno si serva correttamente dei programmi e dei font giusti; per la composizione di un brano citato basta conservare quella del documento 'contenitore'. Per scrivere un intero documento in rumeno è necessario procurarsi dei testi composti in Romania e pubblicati da buone case editrici per cercare di impostare la composizione imitando quegli esempi.

Per le citazioni brevi o lunghe in lingua, conviene sfruttare la capacità di polyglossia di mettere a disposizione dell'utente un ambiente col nome della lingua; nel nostro caso:

```
\begin{romanian}
<testo in rumeno>
\end{romanian}
```

e la citazione apparirà in linea con il testo se l'ambiente è compreso dentro un capoverso, mentre apparirà in display se l'ambiente costituisce un capoverso a se stante.

La lingua rumena ha subito molte variazioni da quando i Romani colonizzarono la Dacia e la lingua latina si evolse sul substrato della lingua dei Daci diventando l'ossatura della lingua moderna. Nell'arco dei secoli la Romania ha subito invasioni da parte di slavi di varie etnie, di ungheresi, di tedeschi e di ottomani; parte della Transilvania ancora oggi è abitata da popolazioni di lingua madre ungherese. Hanno tutte lasciato nella lingua rumena parole e elementi sintattici che ne hanno modificato in modo notevole la sua origine neolatina. Fra le lingue romanze è l'unica che ha conservato la declinazione dei nomi⁴; la sua particolarità specifica è quella di agglutinare alla fine delle parole il rispettivo articolo che viene declinato come la parola che lo precede: *motor* (motore, senza articolo), *motorul* (il motore), *motorului* (del motore), *motorile* (i motori), *motorilor* (dei motori), eccetera.

Anche la scrittura ha subito modifiche di diversi tipi e fino a non molto tempo fa era basata sull'alfabeto cirillico. Anche se l'alfabeto latino ha cominciato ad essere preso in considerazione a metà del XIX secolo, dopo la Grande Guerra è stato dichiarato ufficiale, ma l'ultima riforma ortografica risale all'ultimo decennio del XX secolo. Questo spiega in parte l'uso di diversi diacritici e la sussistenza in Romania di alcune distinte ortografie. L'alfabeto rumeno contiene i seguenti segni:

⁴Sembra che anche l'occitano abbia conservato tracce di declinazione.

a ă â b c d e f g h i î j k l m
n o p (q) r s ș t ț u v (w) x (y) z

dove le lettere tra parentesi vengono usate per riportare parole straniere. Questo alfabeto latino con i suoi diacritici vale per scrivere con le diverse ortografie. Credo però che i pattern di sillabazione vadano bene sia per l'ultima sia per la penultima ortografia ufficiale. A suo tempo, precedentemente alla pubblicazione di T_EX v.3.0, con la collaborazione di alcuni amici rumeni avevo creato (CB) una prima versione di pattern per il rumeno, ma andavano bene per la penultima ortografia soltanto. Poi, giustamente, i miei pattern sono stati sostituiti da altri predisposti da rumeni esperti della loro lingua e della programmazione dei pattern. Ora non dispongo più di file di prova per controllare le due ultime ortografie ufficiali. Ma sono piuttosto convinto che i pattern oggi disponibili siano adeguati per l'ultima ortografia ufficiale.

I segni diacritici servono solo per cambiare la pronuncia ai segni di base; il rumeno non dispone di accenti tonici, quindi per uno straniero risulta difficile leggere correttamente certe parole; in linea di massima l'accento tonico cade sulla penultima vocale, tenuto conto di possibili dittonghi che in rumeno sono più numerosi che in italiano.

Le tradizioni tipografiche rumene si sono sviluppate negli ultimi tempi, e i moduli per la configurazione di questa lingua per babel e per polyglossia non impostano nulla di particolare in merito a virgolette, elenchi, rientri dei capoversi, e simili.

A titolo di esempio, quanto segue è il primo articolo della Costituzione rumena:

ARTICOLUL 1

1. România este stat național, suveran și independent, unitar și indivizibil.
2. Forma de guvernământ a statului român este republica.
3. România este stat de drept, democratic și social, în care demnitatea omului, drepturile și libertățile cetățenilor, libera dezvoltare a personalității umane, dreptatea și pluralismul politic reprezintă valori supreme, în spiritul tradițiilor democratice ale poporului român și idealurilor Revoluției din decembrie 1989, și sunt garantate.
4. Statul se organizează potrivit principiului separației și echilibrului puterilor - legislativă, executivă și judecătorească - în cadrul democrației constituționale.
5. În România, respectarea Constituției, a supremației sale și a legilor este obligatorie.

Il testo precedente è scritto con l'ultima ortografia decretata nel 1993; le poche cesure presenti sono corrette, ma non si può certo dedurre che i pattern rumeni siano tutti corretti da solo tre esempi.

4.11 COMMENTI

Non si presentano ulteriori esempi; come si è detto i pattern disponibili servono per una ottantina di lingue; quasi tutti i file di descrizione delle lingue si servono di almeno un insieme di pattern di sillabazione; sono poche le lingue che non ne dispongono. Come si è detto l'arabo non divide in sillabe ma giustifica i testi ricorrendo alla kashida, il prolungamento del tratto terminale delle lettere finali. Le lingue scritte con ideogrammi si scrivono senza spazi fra le parole. Non abbiamo fatto esperimenti con lingue simili di cui non

sappiamo nulla, come, per esempio, il sanscrito, il thailandese, o altre lingue scritte con alfabeti che ci sono ignoti.

Ci sono alcune lingue, come l'occitano, che sono scritte in caratteri latini, ma non dispongono di file di sillabazione, quindi sono composti come se fossero un dialetto della 'lingua' nohyphenation; ovviamente la composizione in queste lingue può risultare a posto se si compone con una giustezza piuttosto grande, dove si riesce a giustificare le righe mediante l'elasticità dello spazio interparola, ma la composizione presenta tutti i difetti che si manifestano con alcuni word processor quando non si imposta la cesura in fin di riga. Siamo al corrente che si sta procedendo alla sistemazione dei pattern occitani, ma è una magra consolazione, specialmente per gli umanisti specializzati in filologia romanza, oltre che, evidentemente, per gli occitani stessi.

ELENCO DEI SITI INTERNET

CTAN

Sito ufficiale di CTAN.
<http://www.ctan.org/>

EMACS

Pagina Web di Emacs.
<http://www.gnu.org/software/emacs/>

ESCHER

Sito ufficiale di M. C. Escher.
<http://www.mcescher.com/>

FONTS

Catalogo danese dei font per L^AT_EX.
<http://www.tug.dk/FontCatalogue/>

GUI

Sito ufficiale del G_UI_T.
<http://www.guitex.org/>

KILE

Pagina Web di Kile.
<http://kile.sourceforge.net/>

KNUTH

Pagina Web di Donald E. Knuth.
<http://www-cs-faculty.stanford.edu/~knuth/>

LAMPORT

Pagina Web di Leslie Lamport.
<http://research.microsoft.com/en-us/um/people/lamport/>

MACTEX

Pagina Web di MacT_EX.
<http://www.tug.org/mactex/>

MIKTEX

Sito ufficiale di MiK_T_EX.
<http://miktex.org/>

SAROVAR

Catalogo L^AT_EX online.
<http://texcatalogue.sarovar.org/>

TEXDOC

Versione online di texdoc.
<http://texdoc.net/>

TEXLIVE

Pagina Web di T_EX Live.
<http://www.tug.org/texlive/>

TEXMAKER

Pagina Web di Texmaker.
<http://www.xmlmath.net/texmaker/>

TEXSTUDIO

Pagina Web di TeXstudio.
<http://sourceforge.net/projects/texstudio/>

TEXSHOP

Pagina Web di TeXShop.
<http://www.uoregon.edu/~koch/texshop/>

TEXWORKS

Pagina Web di T_EXworks.
<http://www.tug.org/texworks/>

BIBLIOGRAFIA

Beccari, Claudio

- 2012 *Introduzione alla definizione della geometria della pagina*, Guide tematiche [G_UIT](http://www.guitex.org/home/images/doc/GuideGuIT/intropagedesign.pdf), <http://www.guitex.org/home/images/doc/GuideGuIT/intropagedesign.pdf>.
- 2013a *Font e tipografia. PDF_{La}TeX, X_YLaTeX e i font*, Guide tematiche [G_UIT](http://www.guitex.org/home/images/doc/GuideGuIT/guidafont.pdf), <http://www.guitex.org/home/images/doc/GuideGuIT/guidafont.pdf>.
- 2013b *Introduzione alla creazione di file di classe*, Guide tematiche [G_UIT](http://www.guitex.org/home/images/doc/GuideGuIT/introcreaclassi.pdf), <http://www.guitex.org/home/images/doc/GuideGuIT/introcreaclassi.pdf>.
- 2014a *Introduzione all'arte della composizione tipografica con LaTeX*, <http://www.guitex.org/home/images/doc/GuidaGuIT-B5.pdf>.
- 2014b *Introduzione all'arte della composizione tipografica con LaTeX*, <http://www.guitex.org/home/images/doc/guidaguit.pdf>.
- 2015 *Il LaTeX Reference Manual commentato*, Guide tematiche [G_UIT](http://www.guitex.org/home/images/doc/GuideGuIT/latexhandbookcommentato.pdf), <http://www.guitex.org/home/images/doc/GuideGuIT/latexhandbookcommentato.pdf>.

Beccari, Claudio e Tommaso Gordini

- 2012 *Codifiche in T_EX e L_ATeX. Dal sorgente al PDF. Guida pratica per lavorare con successo*, Guide tematiche [G_UIT](http://www.guitex.org/home/images/doc/GuideGuIT/introcodifiche.pdf), <http://www.guitex.org/home/images/doc/GuideGuIT/introcodifiche.pdf>.

Bringinghurst, Robert

- 1992 *The Elements of Typographic Style*, Hartley & Marks, Vancouver, (edizione italiana a cura di Lucio Passerini, *Gli Elementi dello Stile Tipografico*, Sylvestre Bonnard, Milano 2001). (Citato a p. 39.)

Cevolani, Gustavo

- 2006 «Norme tipografiche per l'italiano in LaTeX», *ArsTeXnica*, 1, <http://www.guitex.org/home/images/ArsTeXnica/AT001/arstexnica01.pdf>. (Citato a p. 41.)

Chicago University

- 1994 *The Chicago Manual of Style*, 14^a ed., The University of Chicago Press. (Citato a p. 45.)

De Marco, Agostino

- 2014 «Scrivere la tesi di laurea in LaTeX», Guida Tematica disponibile nel forum del sito [G_UIT](http://www.guitex.org/home/images/doc/GuideGuIT/IntroTesi.pdf), <http://www.guitex.org/home/images/doc/GuideGuIT/IntroTesi.pdf>.

De Marco, Agostino e Roberto Giacomelli

- 2011 «Creare grafici con pgfplots», *ArsTeXnica*, 12.

Eco, Umberto

- 1977 *Come si fa una tesi di laurea. Le materie umanistiche*, Bompiani, Milano.

Fairbairns, Robin

- 2013 *The \TeX FAQ*, <http://www.tex.ac.uk/tex-archive/help/uk-tex-faq/newfaq.pdf>.

Fiandrino, Claudio

- 2011 *Introduzione alla personalizzazione di beamer*, Guide tematiche \LaTeX , <http://www.guitex.org/home/images/doc/GuideGuIT/intropersbeamer.pdf>.
- 2012a *Introduzione all'uso della macro \tikzmark* , Guide tematiche \LaTeX , <http://www.guitex.org/home/images/doc/GuideGuIT/introtikzmark.pdf>.
- 2012b *Introduzione all'uso di TikZ in ingegneria*, Guide tematiche \LaTeX , <http://www.guitex.org/home/images/doc/GuideGuIT/introingtikz.pdf>.

Giacomelli, Roberto

- 2013 *Guida tematica alla riga di comando*, Guide tematiche \LaTeX , <http://www.guitex.org/home/images/doc/GuideGuIT/guidaconsole.pdf>.

Giordano, Mosè, Orlando Iovino e Matteo Leccardi

- 2013 *Guida pratica all'uso di GNU Emacs e \LaTeX* , Guide tematiche \LaTeX , <http://www.guitex.org/home/images/doc/GuideGuIT/guidaemacsauctex.pdf>.

Goossens, Michel

- 2011 *The \XeTeX Companion – \TeX meets OpenType and Unicode*, In <http://xml.web.cern.ch/XML/lgc2/xetexmain.pdf>, \LaTeX Team. (Citato a p. 36.)

Goossens, Michel, Frank Mittelbach e Johannes Braams

- 2004 *The \LaTeX Companion*, Addison-Wesley, Reading (Massachusetts).

Gregorio, Enrico

- 2009 *Appunti di programmazione in \LaTeX e \TeX* , 2^a ed., <http://profs.sci.univr.it/~gregorio/introtex.pdf>.
- 2010a *\LaTeX . Breve guida ai pacchetti di uso più comune*, <http://profs.sci.univr.it/~gregorio/breveguida.pdf>.
- 2010b «L'arte esoterica di scrivere in cirillico con \LaTeX », *ArsTeXnica*, aprile, 9 (apr. 2010), p. 57-73.
- 2011a *Introduzione a \XeLaTeX* , In <http://profs.sci.univr.it/~gregorio/introxelatex.pdf>. (Citato a p. xi.)
- 2011b *Introduzione a \XeLaTeX* , <http://profs.sci.univr.it/~gregorio/introxelatex.pdf>.
- 2011c *Introduzione a \TeX works*, <http://profs.sci.univr.it/~gregorio/introtexworks.pdf>.
- 2013a *Installare \TeX Live 2012 su Ubuntu*, <http://profs.sci.univr.it/~gregorio/texlive-YEAR-ubuntu.pdf>.
- 2013b *Introduzione a [arara](#)*, <http://profs.sci.univr.it/~gregorio/introarara.pdf>.

Guiggiani, Massimo e Lapo Filippo Mori

- 2008 «Consigli su come *non* maltrattare le formule matematiche», *ArsTeXnica*, 5, <http://www.guitex.org/home/images/ArsTeXnica/AT005/Consigli%20su%20come%20non%20maltrattare%20le%20formule%20matematiche.pdf>.

ISO 31/XI

- 1978 *Mathematical signs and symbols for use in the physical sciences and technology*, Norma ISO 31/XI/-1978, International Organization for Standardization, Ginevra. (Citato a p. 39.)

Knuth, Donald Ervin

- 1973 *Computer Programming as an Art*, Addison-Wesley, Reading (Massachusetts), vol. 3.
1984 *The T_EXbook*, Addison-Wesley, Reading (Massachusetts).

Lamport, Leslie

- 1994 *L^AT_EX. A Document Preparation System*, Addison-Wesley, Reading (Massachusetts).

Lesina, Roberto

- 2004 *Il Nuovo Manuale di Stile*, edizione 2.0, con la collaborazione di Federico Boggio Merlo, Zanichelli, Bologna.

Milde, Günter

- 2014 *Greek diacritics with standard accent macors*, PDF document, <http://www.ctan.org/tex-archive/language/greek/greek-fontenc/diacritics.pdf>.
2015 *The textalpha package*, PDF document, <http://www.ctan.org/tex-archive/language/greek/greek-fontenc/textalpha.pdf>.

Mittelbach, Frank, Michel Goossens et al.

- 2004 *The L^AT_EX companion*, 2^a ed., Addison Wesley, Reading, Mass.
2007 *The L^AT_EX graphics companion*, 2^a ed., Addison Wesley, Reading, Mass.

Mittelbach, Frank, Gianluca Pignalberi e Dave Walden

- 2007 «Intervista a Frank Mittelbach», *ArsTeXnica*, 3, <http://www.guitex.org/home/images/ArsTeXnica/AT003/arstexnica03.pdf>.

Mori, Lapo Filippo

- 2006 «Tabelle su L^AT_EX 2_ε: pacchetti e metodi da utilizzare», *ArsTeXnica*, 2, <http://www.guitex.org/home/images/ArsTeXnica/AT002/arstexnica02.pdf>.
2007 «Scrivere la tesi di laurea con L^AT_EX 2_ε», *ArsTeXnica*, 3, <http://www.guitex.org/home/images/ArsTeXnica/AT003/arstexnica03.pdf>.

Mori, Lapo Filippo e Maurizio Himmelmann

- 2007 «Scrivere il curriculum vitae», *ArsTeXnica*, 4, <http://www.guitex.org/home/images/ArsTeXnica/AT004/arstexnica04.pdf>.

Oetiker, Tobias, Hubert Partl, Irene Hyna e Elisabeth Schlegl

- 2011 *The Not So Short Introduction to L^AT_EX 2_ε*, <http://www.ctan.org/tex-archive/info/lshort/english/lshort.pdf>.

Pakin, Scott

- 2009 *The Comprehensive L^AT_EX Symbol List*, <http://www.ctan.org/tex-archive/info/symbols/comprehensive/symbols-a4.pdf>.

Pantieri, Lorenzo

- 2007 «Introduzione allo stile ClassicThesis», *ArsTeXnica*, 6, <http://www.guitex.org/home/images/ArsTeXnica/AT006/arstexnica06.pdf>.
- 2009 «L'arte di gestire la bibliografia con biblatex», *ArsTeXnica*, 8, <http://www.guitex.org/home/images/ArsTeXnica/AT008/arstexnica08.pdf>.

Pantieri, Lorenzo e Tommaso Gordini

- 2009 *L'arte di fare una presentazione con beamer*, http://www.lorenzopantieri.net/LaTeX_files/Presentazioni.pdf.
- 2015 *L'arte di scrivere con L^AT_EX*, In: http://www.lorenzopantieri.net/LaTeX_files/ArteLaTeX.pdf.

Robertson, Will

- 2011 *The X_YL^AT_EX reference guide*, \$TEXMF/doc/xetex/xetexref/XeTeX-reference.pdf.

Sabatini, Francesco e Vittorio Coletti

- 1997 *Il Sabatini Coletti. Dizionario della Lingua Italiana*, Giunti Gruppo Editoriale, Firenze.

Serianni, Luca

- 1989 *Grammatica italiana. Italiano comune e lingua letteraria*, con la collaborazione di Alberto Castelvetti, UTET, Torino.

Starace, Irene

- 2005 (a cura di), *Il grande libro degli haiku*, Castelvetti. (Citato a p. 37.)

UNI-ISO 5966

- 1989 *Presentazione dei rapporti scientifici e tecnici*, Ente Italiano di Unificazione, Milano. (Citato a p. 39.)

Wilson, Peter

- 2004 *The memoir class for configurable typesetting — User guide*, In \$TEXMF/doc/latex/memoir/memman.pdf, The Herries Press, Normandy Park, WA.
- 2009 *A few notes on book design*, In \protect\char"0024\relaxTEXMFdoc/fonts/memdesign/memdesign.pdf, The Herries Press, Normandy Park, WA.

INDICE ANALITICO

COMANDI PARTICOLARI

\', 49, 51

\<, 22, 51

<<, 55

\>, 22, 51

>>, 55

", 47, 55

\', 51

\~, 51

A

\alpha, 51

Ambiente

arab, 31, 32

checkhyphens, 7

chinese, 32

greek, 23, 25

japanese, 32, 33, 35

korean, 37

minipage, 36

otherlanguage, 3

otherlanguage*, 3

russian, 26, 27

\ap, 15

\arabfont, 32

\Asbuk, 27

\asbuk, 27

\atticnum, 24

\atticnumeral, 24

B

\begin, 3, 57

\beta, 51

C

\cdot, 15

\chinesefont, 32

Classe

beamer, 64, 66

ClasseArteLaTeX, III

ClasseArteLingue, III

\cyrillicfont, 27

D

\discretionary, 54

E

\end, 3, 57

\ensuregreek, 51

F

\foreignlanguage, 3

\frenchspacing, 42, 45, 47

G

\Greeknumeral, 24

\greeknumeral, 24

\guillemotleft, 55

\guillemotright, 55

H

\hyphenation, 3, 5, 7

I

\IntelligentComma, 15

J

\japanese font, 32

L

\lefthyphenmin, 14

\Lgem, 54

\lgem, 54

"L, 54

"l, 54

\L.L, 54

\l.l, 54

\lq, 14

N

\newfontfamily, 25

\NoIntelligentComma, 15

\num, 15

O

\omega, 51

P**Pacchetto**

- alphabeta, 51
- arabxetex, 30–32
- ArsClassica, xiii
- babel, 1–3, 7, 8, 13–21, 23, 25, 27, 28, 43–49, 51–55, 57, 58
- begingreek, 23
- biblatex, 66
- bidi, 29
- ClassicThesis, xiii
- combelow, 57
- ecclesiastic, 18, 49
- fontenc, 26, 36, 43, 51
- fontspec, 36, 37, 43, 55
- hyperref, 29, 51, 53
- indentfirst, 41
- inputenc, 1, 22, 43, 56
- listings, xi
- liturg, 49
- liturgic, 18
- microtype, xv, xvi
- pgfplots, 63
- polyglossia, 18
- poliglossia, 19
- polyglossia, 1–3, 7, 8, 13, 16–19, 23–32, 36, 37, 43–49, 51–55, 57, 58
- siunitx, 15
- testthyphens, 7
- textalpha, 23, 51, 65
- textcomp, 55

TikZ, 64

- \ped, 15
- \PolyglossiaSetup, 41
- \ProsodicMarksOff, 18
- \ProsodicMarksOn, 18

R

- \raggedright, 32, 33
- \roman, 52
- \rq, 14

S

- \selectlanguage, 3
- \setactivedoublequote, 14
- \setISOcompliance, 15
- \setmainlanguage, 2, 16, 24, 27
- \setotherlanguage, 16, 24, 31
- \shorthandsoff, 14
- \shorthandson, 14
- \slash, 14

T

- \testthyphens, 7
- \textarab, 32
- \textchinese, 32
- \textjapanese, 33
- \tikzmark, 64

U

- \unit, 15

X

- \XeTeXlinebreakskip, 33