

# La composizione tipografica di alcune lingue orientali

Claudio Beccari

## Sommario

Scrivere nelle lingue orientali, usate dal Medio Oriente fino alla Cina e al Giappone, è una cosa che si può fare anche con i programmi del sistema  $\text{\TeX}$ . La difficoltà maggiore è la conoscenza di quelle lingue e dei loro sistemi di scrittura. Per brevità qui si parlerà dell'arabo, del giapponese e del coreano, con qualche accenno al cinese.

## Abstract

Typesetting documents with the languages written and spoken in the Near East to the Far East is possible with the  $\text{\TeX}$  system programs. The major difficulty, besides knowing those languages, consists in their writing systems. For conciseness here we shall deal with Arabic, Japanese and Korean with modest hints to Chinese.

## 1 Introduzione

Con i programmi del sistema  $\text{\TeX}$  è relativamente semplice comporre documenti con alfabeti diversi dai font latini, con direzioni di scrittura diverse e usando font molto legati. Con le lingue che fanno uso dei caratteri arabi la direzione di scrittura è da destra a sinistra e i font sono strettamente legati.

Col cinese e il giapponese la direzione di scrittura è generalmente da sinistra a destra, ma per alcuni testi formali non è esclusa la direzione dall'alto al basso. I font sono costituiti da ideogrammi. Altri sistemi di caratteri sono "letterali", con segni diversi da quelli latini, ma disposti in sillabe di forma pressoché quadrata come il coreano.

Si rende necessario usare font di tipo OpenType adatti alle lingue in questione; con l'arabo è necessario scrivere una specie di codice, un po' come si scrive un documento con  $\text{\LaTeX}$  o come si programma un disegno con  $\text{\TikZ}$ ; ma per chi conosce l'arabo e i suoi problemi di vocalizzazione o, per il Corano, i problemi di ritmica per marcare lo scritto con i segni prosodici necessari alla lettura rituale, non dovrebbe essere particolarmente complicato. In confronto all'arabo, l'ebraico è molto più semplice, anche se questa lingua ha i soliti problemi di vocalizzazione delle lingue semitiche<sup>1</sup>, i cui sistemi di caratteri rappresentano per lo più delle consonanti.

1. Il maltese è l'unica lingua semitica che si scrive con caratteri latini e le vocali sia pure con diversi diacritici, anche se le numerose dominazioni che l'isola ha subito hanno introdotto

Un altro problema per scrivere in queste lingue orientali è la tastiera. Va da sé che le tastiere dei calcolatori personali, che dispongono di un centinaio di tasti, benché siano usate correntemente anche in quei paesi hanno un numero di tasti di solito abbondantemente insufficiente per le lingue che scrivono con ideogrammi. Sembra che il totale dei caratteri cinesi, usati col nome di *kanji* anche in giapponese, ammontino a circa 35 000. Abituati come siamo con l'alfabeto latino di 26 caratteri (più altri 26 per le maiuscole), i numeri degli ideogrammi sono impressionanti. La presenza di accenti o altri segni diacritici non rende il nostro alfabeto sostanzialmente più complesso.

I giapponesi e i cinesi devono usare la loro tastiera anche per scrivere in caratteri latini; quindi devono disporre di tasti per commutare dalla tastiera latina (generalmente contente solo i 95 caratteri ASCII senza accenti) alla tastiera *kanji*. Per i giapponesi la cosa è ancora più complicata, perché devono usare anche i sillabari *katakana* e *hiragana*, che contengono ciascuno una sessantina di segni.

## 2 Le tastiere

### 2.1 La tastiera araba

I caratteri arabi sono relativamente pochi: nella documentazione del pacchetto *arabxetex* la tabella 1 riporta 28 righe, in una delle quali sono indicate le tre vocali che di tanto in tanto è necessario introdurre; queste non rappresentano una forma di vocalizzazione, ma sono "vocali di sostegno" che si usano in certe parole; esse corrispondono pressappoco ai nostri suoni 'u', 'a', 'i'. Ma il punto non è tanto questo, quanto il fatto che questi segni e quelli di tutte le consonanti hanno generalmente quattro forme per la lettera isolata, per l'iniziale, per la mediale e per la finale; ecco quindi che i segni diventano circa 120. Se si aggiungono i segni per la vocalizzazione, la moltitudine di segni aumenta molto. Magra consolazione: non ci sono maiuscole e minuscole.

Per questo motivo si è sviluppata una forma di scrittura codificata mediante lettere latine minuscole e maiuscole intercalate con segni alfabetici, in modo da caratterizzare ogni aspetto della lingua araba scritta, ma ci vuole un preprocessore o un programma comunque invocato dal programma di

molti forestierismi; quelli italiani sono facilmente riconoscibili sia per iscritto sia quando si parla; oggi il maltese è lingua ufficiale della Repubblica di Malta a pari livello dell'inglese.

composizione che trasformi l'input codificato, come si è detto, con caratteri ASCII in una sequenza di caratteri UNICODE, adatti per combinarsi nelle giuste forme e per legarsi fra di loro come richiede la scrittura araba. Questa transcodifica dell'arabo viene consigliata anche a coloro che parlano e scrivono in quella lingua, anche se devono usare la modalità ASCII della tastiera.

Di fatto, quindi, per scrivere in arabo non occorre una tastiera araba o “bilingue”, cosa che per noi europei torna particolarmente comoda.

## 2.2 La tastiera giapponese

Come ho detto, la tastiera che usano i giapponesi permette loro, mediante opportuni tasti, di cambiare la modalità di scrittura, dal *romaji* (caratteri latini), al katakana (sillabario giapponese usato specialmente per traslitterare i nomi stranieri, figura 1), al hiragana (usato specialmente per scrivere le pre- e post-posizioni della lingua giapponese, figura 2), e il *kanji* (per usare gli ideogrammi cinesi, troppo numerosi per poterli raffigurare in questo articolo). In teoria, quindi, con la tastiera giapponese si può scrivere anche in cinese.

Ma il problema è che gli ideogrammi cinesi sono delle singole parole di solito monosillabiche, pronunciate con i debiti toni. Gli stessi ideogrammi convogliano la stessa “idea” in giapponese, ma vengono pronunciati con parole generalmente polisillabiche, magari raggruppando in una sola parola giapponese l'idea complessiva del gruppo di ideogrammi. Il giapponese non ha toni, ma ha sillabe lunghe, medie e brevi. Quelle lunghe, scrivendo in romaji si rendono o raddoppiando la vocale della sillaba, o sovrapponendo un segno di lunga: *arigatoo*, oppure *arigatō*; talvolta in hiragana scrivono una sillaba ‘u’ alla fine della parola *ありがとう*, o appendono un segno di lunga dopo la sillaba ‘to’, *ありがとー*. Quelle brevi possono essere marcate con un segno di breve (ū) o vengono scritte in carattere di corpo minore.

Il sillabario hiragana è originale giapponese antico e risale al XI secolo; fu usato dalla dama di compagnia Murasāki Shikibu per scrivere *La leggenda di Genji* (ORSI, 2012), apparentemente il più antico romanzo della storia mondiale. Tutto il testo è scritto usando il sillabario hiragana; il testo è ancora usato oggi per rappresentare nel teatro tradizionale Kabuki l'interpretazione teatrale divisa in una sequenza interminabile di atti, che dura diversi mesi e che vengono rappresentati ciclicamente da alcuni secoli.

Il sillabario katakana, che comprende essenzialmente linee diritte, è stato “inventato” per scrivere telegrammi e permettere alle telescriventi civili e militari di trasmettere in giapponese. Gli ideogrammi cinesi sono entrati nell'uso diversi secoli fa quando per il Giappone l'Impero cinese rappresentava un modello di cultura e di stile.

Ovviamente il romaji è il più recente sistema di scrittura diffusosi anche con l'avvento dei calcolatori personali, in quanto prima i centri di calcolo erano sostanzialmente dedicati al trattamento di dati numerici; la gente scriveva con la penna, o meglio, esercitava la sua calligrafia con pennelli ed inchiostri speciali. Ancora oggi la calligrafia con il pennello è un'arte apprezzatissima.

Sul Mac si possono installare diversi driver di tastiera e sceglierli cliccando sull'icona apposita nella barra superiore. Come farlo dipende dal particolare sistema operativo, ma sostanzialmente bisogna cliccare l'icona “System Preferences”; nel dialogo che appare si clicca sull'icona della tastiera e si apre un altro dialogo; in questo si sceglie la voce di menu “Input Sources” dove bisogna marcare l'opzione “Show keyboard and emoji viewers in menu bar”, in modo che quando si vuole cambiare driver lo si sceglie dalla lista dei driver installati, ma si ha anche la possibilità di mostrare la tastiera virtuale; in questo dialogo si sceglie quali driver aggiungere e quali togliere. Quando si mostra la tastiera virtuale vi si può cliccare col mouse; in ogni caso è utilissima per sapere quali tasti hardware (con i simboli serigrafati sulla faccia superiore di ogni tasto della tastiera reale) corrispondano a quali tasti della tastiera virtuale.

Con il driver per il giapponese scegliere “Hiragana” fra i metodi di input del driver; nella tastiera virtuale compaiono i tasti romaji (latini), battendo la parola giapponese da scrivere appaiono via via le sillabe hiragana in una finestrina a parte, ma quando la parola è completamente scritta compaiono anche le versioni hiragana e kanji nella finestrina fra le varie parole si sceglie quella giusta cliccando sopra e la parola si trasferisce al documento che si sta scrivendo, mentre si chiude la finestrina, pronta a riaprirsi per la parola successiva.

Conviene definire nel preambolo un paio di comandi per scegliere un font OpenType che contenga i caratteri giapponesi e il cui nome sia `\japanesefont`; si definisca anche un comando per scrivere con quel font, e magari anche un ambiente; per esempio:

```
\newfontfamily{\japanesefont}{Hiragino Mincho Pro}
\newcommand*\textjapanese[1]{\{\japanesefont #1\}}
\newenvironment{japanese}[1] [] {\#1\japanesefont}{}
```

In questo modo il font scelto per scrivere in giapponese è l'Hiragino Mincho Pro<sup>2</sup>; il comando `\textjapanese` compone la stringa che gli si dà come argomento usando il font scelto. L'ambiente `japanese` accetta un argomento facoltativo che viene usato come prima cosa nell'apertura dell'ambiente; può servire per specificare comandi di corpo o qualunque altra cosa sia utile in un ambiente.

Per comporre quindi la parola ‘arigatoo’ in giapponese, basta inserire nel file sorgente il codice:

2. Questo font è uno dei vari disponibili sul Mac; altri sistemi operativi dispongono di font con altri nomi.

## カタカナ

ヒョ	チュ	キヤ	ン	ワ	ラ	ヤ	マ	ハ	ナ	タ	サ	カ	ア	ド	ズ	ガ
hyo	chu	kya	n	wa	ra	ya	ma	ha	na	ta	sa	ka	a	do	zu	ga
ミヤ	チヨ	キユ			リ		ミ	ヒ	ニ	チ	シ	キ	イ	バ	ゼ	ギ
mya	cho	kyu			ri		mi	hi	ni	chi	shi	ki	i	ba	ze	gi
ミュ	ニヤ	キヨ			ル	ユ	ム	フ	ヌ	ツ	ス	ク	ウ	ビ	ゾ	グ
myu	nya	kyo			ru	yu	mu	fu	nu	tsu	su	ku	u	bi	zo	gu
ミョ	ニユ	シャ			レ		メ	ヘ	ネ	テ	セ	ケ	エ	ブ	ダ	ゲ
myo	nyu	sha			re		me	he	ne	te	se	ke	e	bu	da	ge
リヤ	ニョ	シュ			ヲ	ロ	ヨ	モ	ホ	ノ	ト	ソ	コ	オ	ベ	ヂ
rya	nyo	shu			wo	ro	yo	mo	ho	no	to	so	ko	o	be	ji
リュ	ヒヤ	ショ	ピョ	ピャ	ビュ	ヂョ	ヂャ	ジュ	ギョ	ギャ		ペ	ピ	ボ	ズ	ザ
ryu	hya	sho	pyo	pya	byu	jyo	jya	jyu	gyo	gya		pe	pi	bo	zu	za
リョ	ヒュ	チャ			ピョ	ビョ	ビャ	ヂュ	ジュ	ジャ	ギュ		ポ	プ	パ	デ
ryo	hyu	cha			pyu	byo	bya	jyu	jyo	jya	gyu		po	pu	pa	de
																ジ

FIGURA 1: I 107 segni del sillabario giapponese katakana (カタカナ), dei quali 46 sono principali, e gli altri 61 sono repliche modificate mediante delle specie di diacritici.

\textjapanese{ありがとう}

e il gioco è fatto. Ovviamente per tratti più corposi di testo conviene usare l'ambiente.

### 2.3 Tastiera cinese

Come ho detto la tastiera giapponese consente di scrivere anche in cinese. Sono sicuro che in Cina usano tastiere adeguatamente configurate.

Sono altrettanto sicuro che per il cinese sia disponibile un metodo per immettere gli ideogrammi, tenendo conto che in questa lingua non esistono sillabari come in giapponese.

Non avendo la benché minima nozione di cinese, non mi sono cimentato nell'usare qualche driver del mio Mac, quindi non sono in grado di dare ulteriori informazioni. Più avanti indicherò un esempio di input in cinese; ma il testo l'ho semplicemente copiato da internet, perché la mia conoscenza di quella lingua è nulla.

### 3 La tastiera coreana

La lingua coreana si scrive con un alfabeto vero e proprio che contiene 14 consonanti e 10 vocali; in realtà queste cifre sono origine di controversie; sta di fatto che i tasti alfabetici della tastiera coreana sono 26; evidentemente la classificazione minimale di 24 lettere non è quella seguita da chi ha predisposto quella tastiera. I segni che rappresentano i vari suoni e le loro varianti si chiamano *jamo* e hanno forme completamente diverse da quelle latine, quindi, senza conoscere quell'alfabeto, la lingua coreana sembra scritta con strani ideogrammi. In realtà le singole lettere sono aggiustate dentro un ideale contorno quadrato che contiene un'intera

Combo Hiragana				Order from top to bottom, right to left			
にや	ちや	しや	きや	な	た	さ	か
nya	cha	sha	kya	na	ta	sa	ka
にゆ	ちゆ	しゆ	きゆ	に	ち	し	き
nyu	chu	shu	kyu	ni	chi	shi	ki
によ	ちよ	しよ	きよ				
nyo	cho	sho	kyo				
ぎや	りや	みや	ひや	ぬ	つ	す	く
gya	rya	mya	hya	nu	tsu	su	ku
ぎゆ	りゆ	みゆ	ひゆ				
gyu	ryu	myu	hyu				
ぎよ	りよ	みよ	ひよ	ね	て	せ	け
gyo	ryo	myo	hyo	ne	te	se	ke
びや	びや	ぢや	じや	の	と	そ	こ
pya	bya	dzya	jya	no	to	so	ko
びゆ	びゆ	ぢゆ	じゆ	ん	わ	ら	や
pyu	byu	dzyu	jyu	n	wa	ra	ya
びよ	びよ	ぢよ	じよ				
pyo	byo	dzyo	jyo				
Dakuten							
ば	ば	だ	ざ				
pa	ba	da	za				
び	び	ぢ	じ				
pi	bi	dzi	ji				
ぶ	ぶ	づ	ず				
pu	bu	dzu	zu				
べ	べ	で	ぜ				
pe	be	de	ze				
ぼ	ぼ	ど	ぞ				
po	bo	do	zo				

FIGURA 2: I 97 segni del sillabario giapponese hiragana (ひらがな) dei quali 41 sono principali e gli altri 56 sono repliche modificate mediante delle specie di diacritici.



FIGURA 3: La tastiera coreana sul Mac

sillaba; la sillaba può essere composta con uno, due o tre suoni elementari, quindi può contenere fino a tre lettere; se la sillaba è fatta di un solo suono, quindi un solo segno, gli viene aggiunto un segno speciale che non ha suono, di forma rotondeggiante con il significato di “aprire la bocca per parlare”. Questo modo di scrivere di chiama *hangul*. Ecco, questa parola si scrive con due sillabe ‘han’ e ‘gul’; essa dunque compare nel testo come due gruppi di singole lettere in questo modo 한글 È difficile per un occhio inesperto distinguere i tre segni in ciascuna delle due sillabe, ma guardando bene i segni ingranditi esse si riescono a distinguere (figura 4).

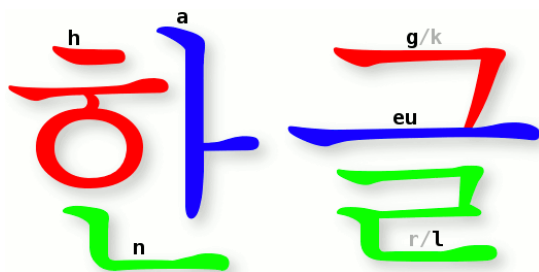


FIGURA 4: La parola coreana 한글 (hangül o hangeul) ingrandita e marcata con le lettere latine che corrispondono ad ogni lettera coreana (jamo).

Per comporre in coreano conviene di nuovo definire un comando `\textkorean` e un ambiente `korean` in modo del tutto simile a quanto detto per il giapponese. Se si specifica con `polyglossia` la lingua `korean`, questi comandi sono già definiti e, come per le altre lingue, il modulo di `polyglossia` per il coreano prevede le solite parole fisse, la data e qualche altro dettaglio di impostazione tipografica.

Per il Mac suggerisco di installare il driver 2-Set Korean; quando si imposta questo driver e si sceglie di mostrare la tastiera (figura 3), il Mac mostra il layout dei tasti con le lettere coreane disegnate al loro posto; digitando le lettere in sequenza, come si farebbe per scrivere con l’alfabeto latino, il driver forma i gruppi di lettere di ciascuna sillaba e nel testo sorgente esse appaiono già correttamente raggruppate negli schemi quadrati descritti sopra.

Per i font OpenType di solito quelli che contengono gli ideogrammi cinesi e giapponesi contengono anche le lettere coreane; volendo distinguere lo stile dei font coreani, si potrebbe definire nel preambolo il comando:

```
\newfontfamily{\koreanfont}{UnBatang}
```

Un esempio di più vasto respiro potrebbe essere il seguente:

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

#### 4 L’arabo nel file sorgente

Come già detto il modo preferibile per inserire il testo in arabo è quello di usarne una versione codificata con caratteri latini. Di traslitterazioni latine, che bisognerebbe chiamare *romanizzazione*, almeno secondo gli arabisti anglosassoni, ce ne sono diverse; le più comuni traslitterazioni/romanizzazioni sono le seguenti: **dmg**, **loc**, **arabica**; esse sono descritte in dettaglio poco più avanti. Il pacchetto **arabi** (JABRI, 2006) indica una traslitterazione; il pacchetto **arabluatex** (ALESSI, 2017) ne indica delle altre.

Benché apparentemente il pacchetto **arabi** abbia ricevuto recentemente un aggiornamento, personalmente ho incontrato delle difficoltà a farlo convivere con l’opzione per il greco. Questo, ovviamente non ha nulla a che vedere con la composizione tipografica della sola lingua araba. Infatti, per evitare conflitti ho composto in arabo con questo pacchetto quello che volevo comporre, poi ho ritagliato dalla pagina il testo composto e l’ho inserito come figura nell’altro documento che conteneva anche diffuse inserzioni in greco. Le indicazioni per rivolgersi al curatore del pacchetto **arabi** riportate nella prima pagina della documentazione sono errate o, comunque, non aggiornate nemmeno dopo quest’ultimo aggiornamento, quindi è impossibile segnalare il problema, perché il curatore non sembra più reperibile.

TABELLA 1: Le traslitterazioni principali in lettere latine dell'arabo come sono riportate nella documentazione del pacchetto `arabluatex`. Tratto integralmente dalla documentazione di `arabluatex`, compresi i richiami delle note, per le quali si rimanda direttamente a quella documentazione.

Letter	Transliteration <sup>9</sup>			ArabT <small>EX</small> notation
	dmg	loc	arabica	
أ <sup>10</sup>	'u	'u, 'a, 'i	'u, 'a, 'i	'u or 'a or 'i
ب	<i>b</i>	<i>b</i>	<i>b</i>	<i>b</i>
ت	<i>t</i>	<i>t</i>	<i>t</i>	<i>t</i>
ث	<i>ṭ</i>	<i>th</i>	<i>ṭ</i>	<i>_t</i>
ج	<i>ǧ</i>	<i>j</i>	<i>ǧ</i>	<i>^g</i> or <i>j</i>
ح	<i>ḥ</i>	<i>ḥ</i>	<i>ḥ</i>	<i>.h</i>
خ	<i>ḫ</i>	<i>kh</i>	<i>ḫ</i>	<i>_h</i> or <i>x</i>
د	<i>d</i>	<i>d</i>	<i>d</i>	<i>d</i>
ذ	<i>ḏ</i>	<i>dh</i>	<i>ḏ</i>	<i>_d</i>
ر	<i>r</i>	<i>r</i>	<i>r</i>	<i>r</i>
ز	<i>z</i>	<i>z</i>	<i>z</i>	<i>z</i>
س	<i>s</i>	<i>s</i>	<i>s</i>	<i>s</i>
ش	<i>š</i>	<i>sh</i>	<i>š</i>	<i>^s</i>
ص	<i>ṣ</i>	<i>ṣ</i>	<i>ṣ</i>	<i>.s</i>
ض	<i>ḍ</i>	<i>ḍ</i>	<i>ḍ</i>	<i>.d</i>
ط	<i>ṭ</i>	<i>ṭ</i>	<i>ṭ</i>	<i>.t</i>
ظ	<i>ẓ</i>	<i>ẓ</i>	<i>ẓ</i>	<i>.z</i>
ع	<i>ʿ</i>	<i>ʿ</i>	<i>ʿ</i>	<i>`</i>
غ	<i>ġ</i>	<i>gh</i>	<i>ġ</i>	<i>.g</i>
ف	<i>f</i>	<i>f</i>	<i>f</i>	<i>f</i>
ق	<i>q</i>	<i>q</i>	<i>q</i>	<i>q</i>
ك	<i>k</i>	<i>k</i>	<i>k</i>	<i>k</i>
ل	<i>l</i>	<i>l</i>	<i>l</i>	<i>l</i>
م	<i>m</i>	<i>m</i>	<i>m</i>	<i>m</i>
ن	<i>n</i>	<i>n</i>	<i>n</i>	<i>n</i>
ه	<i>h</i>	<i>h</i>	<i>h</i>	<i>h</i>
و	<i>w</i>	<i>w</i>	<i>w</i>	<i>w</i>
ي	<i>y</i>	<i>y</i>	<i>y</i>	<i>y</i> <sup>11</sup>
ة	<i>ah</i>	<i>ah</i>	<i>a</i>	<i>T</i>



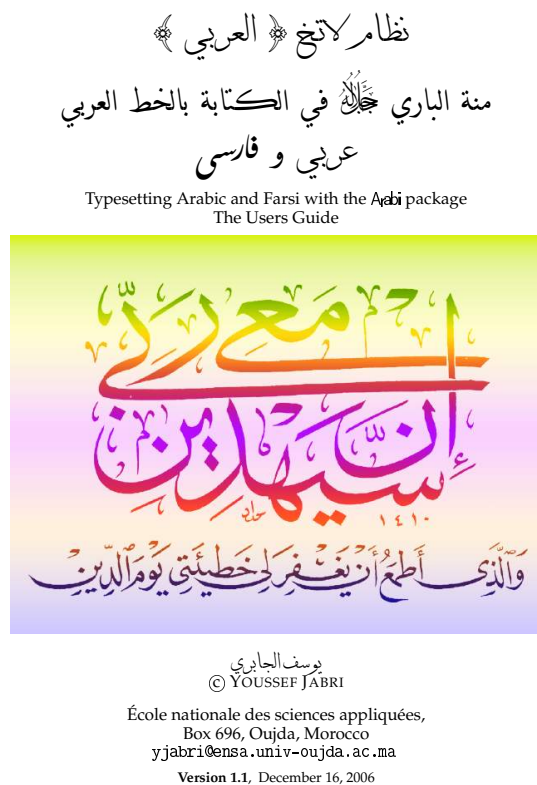


FIGURA 5: Frontespizio della documentazione del pacchetto *arabi*. Composizione di Youssef Jabri.

Il frontespizio della documentazione di *arabi* è particolarmente attraente (figura 5).

Devo segnalare anche che il pacchetto *bidi* (KHALIGHI, 2017), caricato da *arabi*, per comporre l'arabo in modo bidirezionale, da destra a sinistra, e da sinistra a destra la lingua italiana, dava problemi al contenuto di certi comandi; per esempio tre parole italiane composte con il comando `\text` in un contesto italiano e in ambiente matematico, a causa del pacchetto *bidi* risultavano scritte da destra a sinistra, o meglio, le singole parole erano composte correttamente da sinistra a destra, ma le parole erano allineate da destra a sinistra. Veramente strano, ma pur avendo segnalato questo bug, non ho ricevuto nessuna risposta.

C'era anche un conflitto con l'ebraico, che si scrive da destra a sinistra come l'arabo, ma i relativi pacchetti caricavano entrambi il pacchetto *bidi* senza eseguire nessun controllo se il pacchetto fosse già stato caricato. Questo era un piccolo difetto che si risolveva con un poco di astuzia TEX-nica; tuttavia è un peccato che i curatori dei rispettivi pacchetti non facciano attenzione a questi dettagli.

Questo modo di procedere, comporre a parte, ritagliare e incollare, non è molto efficiente se si deve comporre, per esempio, una edizione critica che contenga il testo arabo inframmezzato con qualche lingua composta in caratteri latini.

Ho la sensazione che il pacchetto *arabluatex*, più recente e assistito dal linguaggio Lua integrato con il motore di composizione, sia più efficiente ed efficace. All'occorrenza esiste anche il pacchetto *arabxetex* per comporre con *xelatex*.

Il pacchetto *arabluatex* può eseguire anche alcune traslitterazioni. Ma non bisogna confondere quest'operazione con la codifica d'entrata del file sorgente da comporre in arabo. Nella tabella 1 sono riportate sia le traslitterazioni che *arabluatex* è capace di eseguire, sia la codifica di entrata (nella colonna *Arab TEX notation*); si vede che in qualche modo questa codifica assomiglia alle romanizzazioni delle prime colonne, ma si tratta di cose molto diverse. Le romanizzazioni che *arabluatex* è finora in grado di fare sono descritte qui di seguito<sup>3</sup>.

**dmg** *Deutsche Morgenländische Gesellschaft* è la romanizzazione adottata dalla Convenzione Internazionale degli Orientalisti a Roma nel 1935; questa è la romanizzazione preimpostata dal pacchetto.

**loc** *Library of Congress* appartiene a un vasto insieme di standard per la romanizzazione delle scritture non romane adottato dall'American Library Association e dalla Library of Congress.

**arabica** *Journal of Arabic and Islamic Studies / Revue d'études arabes et islamiques*: è lo standard maggiormente usato dagli Arabisti

Ovviamente il pacchetto in esame consente anche la vocalizzazione parziale o totale del testo arabo. La cosa è troppo complessa da descrivere qui; per questo è necessario rivolgersi alla documentazione.

Naturalmente, per un testo completamente in arabo o per testi più lunghi di poche parole oltre al comando `\arb` si dispone anche dell'ambiente *arab*, come lo si vede usato nella figura 7.

Merita notare che tutti i pacchetti che consentono di inserire il testo arabo codificato mediante caratteri latini provvedono automaticamente a gestire i numeri, almeno gli interi che si trovano frequentemente nei testi, scrivendoli cominciando dalle unità; siccome la lettura avviene da destra a sinistra, le unità si trovano a destra nella stanga numerica, quindi per noi sembrano scritti nel verso normale; va da sé che se i numeri sono scritti con i numerali arabi, è necessario saperli leggere, visto che questi numerali possono avere forme diverse (ne esistono almeno due varietà in arabo) e comunque sono diversi da quelli che noi chiamiamo "cifre arabe".

Tutti i pacchetti che ho potuto esaminare provvedono all'inserimento della *kashida*, un allungamento verso sinistra delle lettere terminali delle parole, in modo che la giustificazione avviene riempiendo lo spazio interparola mediante la *kashida*,

3. Il pacchetto è ancora in fase di sviluppo ma, rimanendo retrocompatibile, amplierà questa lista.

1 From \textcite[i. 1 A]{Wright}:--- Arabic, like Hebrew and  
 2 Syriac, is written and read from right to left. The letters  
 3 of the alphabet (\arb{.hurUf-u 'l-hijA'-i}, \arb{.hurUf-u  
 4 'l-tahajjI}, \arb{al-.hurUf-u 'l-hijA'iyyaT-u}, or  
 5 \arb{.hurUf-u 'l-mu`jam-i}) are twenty-eight in number and  
 6 are all consonants, though three of them are also used as  
 7 vowels (see §3).

From Wright (1896, i. 1 A):— Arabic, like Hebrew and Syriac, is written and read from right to left. The letters of the alphabet (حُرُوفُ، حُرُوفُ الهجاءِ، الحُرُوفُ الهجائيةُ، التَّهَجِّيّ، حُرُوفُ الْمُعْجَمِ) are twenty-eight in number and are all consonants, though three of them are also used as vowels (see § 3).

FIGURA 6: Testo inglese con brevi inserti in arabo. Tratto dalla documentazione di arabluatex.

```
1 \begin{arab}
2 'at_A .sadIquN 'il_A ju.hA ya.tlubu min-hu .himAra-hu
3 li-yarkaba-hu fI safraTiN qa.sIraTiN fa-qAla la-hu:
4 \enquote{sawfa 'u`Idu-hu 'ilay-ka fI 'l-masA'-i
5 wa-'adfa`u la-ka 'ujraTaN.} fa-qAla ju.hA:
6 \enquote{'anA 'AsifuN jiddaN 'annI lA 'asta.tI`u 'an
7 'u.haqqiqa la-ka ra.gbata-ka fa-'l-.himAr-u laysa hunA
8 'l-yawm-a.} wa-qabla 'an yutimma ju.hA kalAma-hu bada'a
9 'l-.himAr-u yanhaqu fI 'i.s.tabli-hi. fa-qAla la-hu
10 .sadIqu-hu: \enquote{'innI 'asma`u .himAra-ka yA ju.hA
11 yanhaqu.} fa-qAla la-hu ju.hA: \enquote{.garIbuN
12 'amru-ka yA .sadIqI 'a-tu.saddiqu 'l-.himAr-a
13 wa-tuka_d_diba-nI?}
14 \end{arab}
```

أَتَى صَدِيقٌ إِلَى جُحَا يَطْلُبُ مِنْهُ حِمَارَهُ لِيَرْكَبَهُ فِي سَفَرَةٍ قَصِيرَةٍ فَقَالَ لَهُ: "سَوْفَ أُعِيدُهُ إِلَيْكَ فِي الْمَسَاءِ  
 وَأَدْفَعُ لَكَ أَجْرَهُ." فَقَالَ جُحَا: "أَنَا أَسَفٌ جَدًّا أَنِّي لَا أَسْتَطِيعُ أَنْ أُحَقِّقَ لَكَ رَغْبَتَكَ فَالْحِمَارُ لَيْسَ هُنَا  
 الْيَوْمَ." وَقَبْلَ أَنْ يَتِمَّ جُحَا كَلَامَهُ بَدَأَ الْحِمَارُ يَنْهَقُ فِي إِصْطَبْلِهِ. فَقَالَ لَهُ صَدِيقُهُ: "إِنِّي أَسْمَعُ حِمَارَكَ يَا جُحَا  
 يَنْهَقُ." فَقَالَ لَهُ جُحَا: "غَرِيبُ أَمْرِكَ يَا صَدِيقِي أَتَصَدِّقُ الْحِمَارَ وَتُكْذِّبُنِي؟"

FIGURA 7: Esempio d'uso dell'ambiente arab. Tratto dalla documentazione di arabluatex.

invece che con spazio bianco. Invece non sono al corrente della possibilità di eseguire la cesura delle parole nemmeno nelle traslitterazioni in caratteri latini. Quindi la giustificazione dei testi in lettere arabe è affidata esclusivamente alla kashida.

#### 4.1 Codifica del testo sorgente

Sempre prendendo integralmente dalla documentazione di `arabluatex`, si può vedere come si codifica un testo in arabo mescolato a un testo in lettere latine nella figura 6.

#### 4.2 L'uso del pacchetto `arabluatex`

Il pacchetto si carica nel solito modo, solo che bisogna anche scegliere un font per comporre in arabo.

```
\usepackage{opzioni}{arabluatex}
```

Fra le opzioni ci sono `voc`, `fullvoc`, `novoc`, `trans`. L'ultima opzione serve per produrre la traslitterazione secondo una delle tre scelte attualmente disponibili, `dmg`, `loc`, oppure `arabica`; senza specificare nulla viene usata la traslitterazione `dmg`. Le altre tre opzioni riguardano la vocalizzazione: dovrebbero essere auto esplicative; ma la cosa comoda è che, indipendentemente dall'opzione specificata al pacchetto, ciascuna di esse può essere specificata come opzione sia al comando `\arb` sia all'ambiente `arab`.

#### 4.3 I font OpenType che contengono anche l'arabo

Per quanto riguarda i font, questi devono essere OpenType. Non mi pare che attualmente il sistema TEX includa fra i suoi font OpenType anche font che contengano le pagine UNICODE relative all'arabo.

Invece fra i font disponibili con il sistema operativo ce ne sono molti che contengono l'arabo in diversi stili. Fra i primi che si incontrano scorrendo l'elenco dei font installati in un Mac, c'è la collezione di font Amiri; una loro particolarità sono i font specialmente adatti per la stampa del Corano, che prendono il nome di Amiri Quran e Amiri Quran colored, il secondo dei quali serve per inserire le lettere e le vocalizzazioni nel testo coranico in modo che i colori svolgano i compiti speciali loro riservati durante le preghiere collettive recitate nelle moschee. In ogni caso, se non si specifica nessun font, il pacchetto cerca di caricare il font Amiri, se lo trova nella particolare macchina su cui viene usato.

Se non ci si accontenta dei font installati nella propria macchina, se ne possono sempre installare altri; io ho trovato molto interessante il font Scheherazade completo anche dei simboli coranici.

Per terminare la descrizione della composizione in arabo, segnalo che prima dell'esistenza di `arabluatex` era già disponibile il pacchetto `arabxetex` per comporre con l'interprete `xetex` (quindi

lavorando a livello utente con `xelatex`) e senza le funzionalità del linguaggio Lua. Forse il precedente pacchetto `arabxetex` offre maggiori funzionalità per l'utente, ma soffre del fatto che il formato di default di `xelatex` è una variante estesa del formato DVI, che normalmente viene trasformata al volo nel formato PDF. Questo passaggio per il formato DVI presenta numerosi svantaggi sotto molti aspetti. Per questo preferisco usare il pacchetto `arabluatex` anche se è ancora sotto sviluppo e se funziona solo con `lualatex`.

## 5 Comporre in giapponese

Si è già detto molto su come configurare il preambolo e su come usare il driver della tastiera giapponese. Non lo si ripete. Qui parliamo di come comporre davvero in questa lingua.

Normalmente oggi il giapponese si scrive da sinistra a destra in righe giustificate ma seguendo una specie di *lectio continua*. Non si pongono problemi di divisione in sillabe, perché ogni ideogramma è una sillaba; senza alcun segno di "a capo" con un trattino o con un segno equivalente, ogni parola, quasi sempre polisillabica, si può dividere in fin di riga in una posizione qualunque. Questo è ciò che avviene quando si usano gli ideogrammi. Scrivendo in romaji, sarebbe desiderabile che esistesse un file contenente un sia pur minimo insieme di pattern, almeno il romaji potrebbe essere composto con parole staccate e correttamente divise in fin di riga. Tuttavia, ad oggi i colleghi giapponesi non hanno ancora sentito questa necessità, forse perché considerano il romaji come un ripiego a cui ricorrere solo in caso di bisogno estremo; per esempio vengono usati nei tabelloni delle stazioni dove i binari e le destinazioni dei treni sono presentati alternativamente in romaji, in katakana (figura 1, più adatto per i display con pochi pixel luminosi) e in kanji (comunque poco adatto a quei display).

L'ambiente `japanese` definito in precedenza viene ripreso:

```
\newenvironment{japanese}[1][\japanesefont]{}
```

Esso accetta un argomento facoltativo, che è predefinito come argomento vuoto; se lo si usa potrebbe servire per specificare la composizione in bandiera, così da non avere problemi di giustificazione. Per esempio, il codice

```
\begin{japanese}[\raggedright]
まだ何も無い時、
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}
```



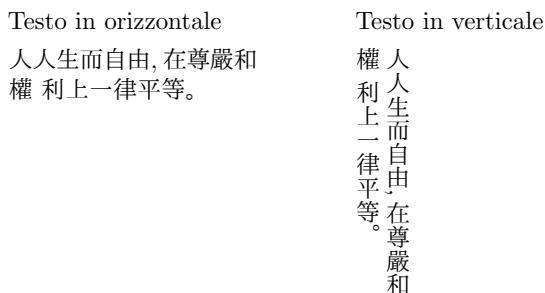


FIGURA 8: Testo composto dentro due `minipage` affiancate con le dichiarazioni `\yoko` e `\tate`

produce il testo

まだ何も無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。

Di fatto, la giustificazione si può ottenere scrivendo un testo sorgente molto ordinato e inserendo fra una parola e l'altra qualche comando di leggera spaziatura che permette al programma di inserire un po' di gomma fra le parole in modo da agevolare la lettura.

Tuttavia, in giapponese formale qualche volta si scrive in verticale con le colonne che si leggono ciascuna dall'alto al basso e le colonne intere da destra a sinistra. Convien fare riferimento al seguente esempio (scritto in cinese), dove lo stesso testo è scritto in righe composte da sinistra a destra che si susseguono dall'alto in basso, oltre che in colonne composte dall'alto in basso che si susseguono da destra a sinistra. Per comporre con lo stile formale con `xelatex` il testo viene composta normalmente in una `minipage` larga abbastanza per contenere senza andare a capo la riga più lunga. Poi bisogna mandare nell'output questa `minipage` ruotata di 90° in senso orario controruotando gli ideogrammi in essa contenuti di 90° in senso antiorario; con le opzioni di `fontspec` si può fare (figura 9).

Il risultato della figura 9 si ottiene con il codice seguente quando si usa `xelatex`:

```
\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*{-0.25em}Testo in verticale:\\
\rotatebox{-90}[origin=lt]{%
\parbox[t]{45mm}{%
\chinesefont
\addfontfeatures{Renderer=AAT,
Vertical=RotatedGlyphs}%
人人生而自由, 在尊嚴和權利上一律平等。}}
\end{minipage}
```

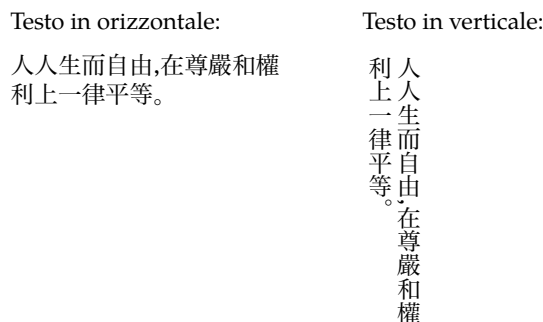


FIGURA 9: Composizione con `XεLaTEX` dello stesso testo cinese in orizzontale e in verticale

Con `lualatex` la procedura in un certo senso è più semplice, ma da solo non ci sarei arrivato; ringrazio Luigi Scarso che mi ha indicato il pacchetto e i comandi da usare, ma dei quali non ho trovato traccia nella documentazione né di `fontspec` né di `luatex`; forse ho guardato male, ma... Infatti, invece di caricare il pacchetto `fontspec`, bisogna caricare il pacchetto `luatexja-fontspec` (THE LUATEX-JA PROJECT TEAM, 2017) scritto apposta per comporre in giapponese con `lualatex`. Va notato che con questo pacchetto per caricare i font alfabetici come il latino, il greco, il cirillico si usano i soliti comandi, ma per caricare i font del set completo CJK (Chinese, Japanese, Korean) bisogna usare i comandi che contengono una 'j' nel loro nome, come, per esempio, `\newjfontfamily` invece di `\newfontfamily`.

Tra l'altro questo pacchetto `luatexja-fontspec` rende disponibili i comandi `\yoko` e `\tate`; il primo dichiara di voler comporre in orizzontale, il secondo in verticale. Ecco allora il codice

```
\usepackage{luatexja} % Nel preambolo...

% e nel corpo del documento
\begin{minipage}[t]{40mm}
Testo in orizzontale\par\medskip
\ vbox{\hsize=\textwidth
\yoko\noindent
人人生而自由, 在尊嚴和權利上一律平等。}
\end{minipage}
%
\hfill
%
\begin{minipage}[t]{40mm}
Testo in verticale\par\medskip
\ vbox{\hsize=\textwidth
\tate\noindent
人人生而自由, 在尊嚴和權利上一律平等。}
\end{minipage}
```

componere il suo testo nel modo mostrato nella figura 8. Si osservi che i comandi `\yoko` e `\tate` non possono essere preceduti da altro ed è per questo che sono stati introdotti dentro una scatola nativa dell'interprete, `\vbox`, con la sintassi nativa per

---

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

---

FIGURA 10: Lo stesso *haiku* in giapponese kanji verticale, in romaji e la sua traduzione in italiano

impostare la larghezza della scatola pari a quella della `minipage` che la contiene: infatti, in questo caso `\textwidth` si riferisce alla giustezza del testo dentro la `minipage`.

La figura 10 contiene un *haiku*, tratto da (STARACE, 2005), composto in tre modi diversi esposti da sinistra a destra: la versione dell’*haiku* in giapponese verticale, poi lo stesso *haiku* scritto in romaji, infine la traduzione italiana.

## 6 Conclusioni

I programmi `lualatex` e `xelatex` sono ottimi per gestire i font OpenType con i quali bisogna confrontarsi per comporre testi che contengano brani di testo oppure l’intero testo con caratteri diversi da quelli latini. Volendo trattare di lingue orientali in un’università, è decisamente necessario saper maneggiare questi font completi dei segni di quelle lingue. Questo articolo è solo un tutorial ma dovrebbe essere sufficiente per iniziare la strada dei programmi basati sul mark up  $\text{\LaTeX}$ ; la bibliografia dovrebbe contenere sufficienti informazioni per poter andare oltre questi primi rudimenti.

Nello stesso tempo, le lingue orientali non sono pane per il dente di tutti; un arabista, oltre a conoscere e a saper leggere e scrivere l’arabo o il farsi, deve imparare a scrivere il testo sorgente in arabo o in farsi servendosi della codifica in caratteri latini del sottoinsieme ASCII. Uno studioso di cinese o di giapponese, oltre a conoscere le lingue specifiche, deve conoscere migliaia, se non decine di migliaia, di ideogrammi. Gli studiosi di giapponese devono conoscere anche i sillabari hiragana e katakana, cosa che non dovrebbe essere loro particolarmente difficile, perché le due collezioni di segni ammontano ad un totale che non supera il numero di 150, un’inezia rispetto alle migliaia di kanji necessari per quella lingua.

Le difficoltà tecniche sono minime, rispetto al numero enorme di sillabe e di ideogrammi che bisogna saper maneggiare con destrezza.

Il coreano, a confronto è semplicissimo con un numero limitato segni per le consonanti e per le vocali o suoni vocalici che noi considereremmo dittonghi.

Non ho nemmeno accennato a due altre lingue e ai rispettivi alfabeti molto in uso in Oriente: il devanagari e il thai. Il devanagari ha una importanza enorme, anche grazie a tutta la letteratura relativa

all’induismo e al buddismo. Il thai è certamente meno importante del devanagari, ma si riferisce anch’esso non solo ai testi moderni, ma anche a quelli tradizionali del buddismo, che in Thailandia è religione di stato al punto che persino la data civile è calcolata dalla nascita del Buddha.

Spero che la lettura di questo testo introduttivo possa richiamare l’attenzione degli orientalisti, non solo per la loro personale utilità, ma anche per contribuire a migliorare il software esistente; giustamente quando gli utenti orientali predispongono qualcosa per il sistema  $\text{\TeX}$ , hanno l’uso quotidiano in mente, non le esigenze degli *scholars*, studiosi occidentali di queste discipline.

## 7 Ringraziamenti

Ringrazio di cuore Luigi Scarso; senza il suo aiuto non avrei saputo districarmi nelle finenze di  $\text{\LuaTeX}$  con la sua particolare gestione dei font, specialmente quelli orientali. Per giunta in questo periodo di tempo l’interprete `luatex` ha subito gli ultimi ritocchi prima di essere rilasciato in versione stabile, per cui è stato necessario ricorrere ad un vero esperto. Quindi la mia esperienza precedentemente acquisita con le versioni “beta” dell’interprete ha dovuto essere completamente riveduta e, specialmente, aggiornata.

## Riferimenti bibliografici

- ALESSI, R. (2017). «The arabluatex package». PDF document. Leggibile con `texdoc arabluatex` con  $\text{\TeX}$ Live.
- BECCARI, C. e GORDINI, T. (2016). «L’Arte di scrivere in diverse lingue con (Xe)LaTeX». PDF document. Scaricabile dal sito del `qJr`.
- CHARETTE, F. (2015). «ArabXeTeX – An Arab $\text{\TeX}$ -like interface for typesetting languages in Arabic script with XeLaTeX». PDF document. Leggibile con `texdoc arabxetex` con la distribuzione  $\text{\TeX}$ Live.
- GOOSSENS, M. (2011). *The XeTeX Companion –  $\text{\TeX}$  meets OpenType and Unicode*.  $\text{\LaTeX}$  Team. In <http://xml.web.cern.ch/XML/lgc2/xetexmain.pdf>.
- GOOSSENS, M., MITTELBAACH, F. e BRAAMS, J. (2004). *The  $\text{\LaTeX}$  Companion*. Addison-Wesley.

- GREGORIO, E. (2011). «Introduzione a XeLaTeX». PDF document. URL <http://profs.sci.univr.it/~gregorio/introxelatex.pdf>.
  - JABRI, Y. (2006). «Typesetting Arabic and Farsi with the Arabi package –The Users Guide». PDF document. Leggibile con `texdoc arabi` con la distribuzione TEXLive.
  - KHALIGHI, V. (2017). «The bidi Package». PDF document. Leggibile con `texdoc bidi` con la distribuzione TEXLive.
  - KNUTH, D. E. (1996). *The TEXbook*. Addison Wesley, Reading, Mass., 16<sup>a</sup> edizione.
  - LAMPORT, L. (1994). *LATEX. A Document Preparation System*. Addison-Wesley.
  - LUATEX DEVELOPMENT TEAM (21017). «LuaTeX Reference Manual». PDF document. Leggibile con `texdoc luatex` con TEXLive.
  - MITTELBAACH, F., GOOSENS, M. *et al.* (2004). *The LATEX companion*. Addison Wesley, Reading, Mass., 2<sup>a</sup> edizione.
  - (2007). *The LATEX graphics companion*. Addison Wesley, Reading, Mass., 2<sup>a</sup> edizione.
  - ORSI, M. T. (a cura di) (2012). *La storia di Genji*. Collana *I Millenni*. Einaudi, Torino. Prima traduzione dal giapponese antico (secolo XI) comprendente tutti i 54 capitoli. Titolo originale 源氏物語 (Genji monogatari).
  - ROBERTSON, W. (2011). «The XeTeX reference guide». PDF document. [\\$TEXMF/doc/xetex/xetexref/XeTeX-reference.pdf](#).
  - STARACE, I. (a cura di) (2005). *Il grande libro degli haiku*. Castelvechi.
  - THE LUATEX-JA PROJECT TEAM (2017). «The LuaTeX-ja package». PDF document. Leggibile con `texdoc luatexja` con TEXLive.
- ▷ Claudio Beccari  
claudio dot beccari at gmail  
dot com