

# Greek and Latin Hyphenation

Claudio Beccari

GUI *meeting* 2014

# Stato dell'arte

- 1 Fino al mese di giugno 2014 la sillabazione del greco e del latino presentava delle gravi insufficienze sia lavorando con pdfL<sup>A</sup>T<sub>E</sub>X e *babel*, sia con X<sub>3</sub>L<sup>A</sup>T<sub>E</sub>X o LuaL<sup>A</sup>T<sub>E</sub>X e *polyglossia*.

# Stato dell'arte

- 1 Fino al mese di giugno 2014 la sillabazione del greco e del latino presentava delle gravi insufficienze sia lavorando con pdfL<sup>A</sup>T<sub>E</sub>X e *babel*, sia con X<sub>Y</sub>L<sup>A</sup>T<sub>E</sub>X o LuaL<sup>A</sup>T<sub>E</sub>X e *polyglossia*.
- 2 In particolare non era prevista la sillabazione del latino classico, diversa da quella del latino moderno.

# Stato dell'arte

- 1 Fino al mese di giugno 2014 la sillabazione del greco e del latino presentava delle gravi insufficienze sia lavorando con pdfL<sup>A</sup>T<sub>E</sub>X e *babel*, sia con X<sub>3</sub>L<sup>A</sup>T<sub>E</sub>X o LuaL<sup>A</sup>T<sub>E</sub>X e *polyglossia*.
- 2 In particolare non era prevista la sillabazione del latino classico, diversa da quella del latino moderno.
- 3 Lavorando con pdfL<sup>A</sup>T<sub>E</sub>X e *babel* non era prevista la differenziazione fra le tre ortografie del greco, e non era usabile la sillabazione corretta per il greco moderno monotonico né per quello antico politonico; inoltre la sillabazione non era corretta se si scriveva nel file sorgente il testo greco in modo letterale.

## I file di descrizione delle lingue

Sia per *polyglossia* sia per *babel* la gestione delle lingue avviene attraverso opportuni file di descrizione di ciascuna lingua.

Questi file provvedono alla definizione delle parole fisse (per esempio Capitolo, Chapter, Chapître, . . . ), alla stampa delle date e alla gestione delle specificità tipografiche della lingua; fra queste, le regole di sillabazione già caricate nel formato usato dal motore di composizione.

## Descrizione del Latino

Fino a giugno 2014 i file di descrizione del latino *latin.ldf* e *gloss-Latin.ldf* prevedevano solo il latino moderno e il latino medievale.

*latin.ldf* prevedeva la possibilità di usare i segni prosodici; un pacchetto esterno, *ecclesiastic* consentiva la scrittura del latino moderno con l'accentazione usata nei testi devozionali della chiesa cattolica.

Ma nulla permetteva di selezionare la sillabazione del latino classico, anche perché non ne erano definite le regole né per *babel* né per *polyglossia*.

## Descrizione del greco

- 1 Con *polyglossia* erano definite sia le parole infisse sia la sillabazione per le tre forme di scrittura della lingua greca: quella moderna, monotonica, e quelle moderna e classica, entrambe politoniche.

## Descrizione del greco

- 1 Con *polyglossia* erano definite sia le parole infisse sia la sillabazione per le tre forme di scrittura della lingua greca: quella moderna, monotonica, e quelle moderna e classica, entrambe politoniche.
- 2 Con *babel* si poteva selezionare solo l'ortografia e la sillabazione del greco moderno politonico, benché fossero disponibili anche le regole di sillabazione per le altre ortografie.



## Descrizione del greco

- 1 Con *polyglossia* erano definite sia le parole infisse sia la sillabazione per le tre forme di scrittura della lingua greca: quella moderna, monotonica, e quelle moderna e classica, entrambe politoniche.
- 2 Con *babel* si poteva selezionare solo l'ortografia e la sillabazione del greco moderno politonico, benché fossero disponibili anche le regole di sillabazione per le altre ortografie.
- 3 Inoltre, poiché i font greci gestibili con pdfL<sup>A</sup>T<sub>E</sub>X sono solo *8-bit encoded*, originariamente si era fatto uso di una traslitterazione con le lettere ASCII a 7-bit, *in teoria* usabili con qualunque tastiera. Le regole di sillabazione erano legate solo a questa traslitterazione.

## Situazione decisamente insoddisfacente

È chiaro che la situazione era insoddisfacente. Si aggiunga il fatto che con l'ultima versione di *babel* è diventato possibile inserire direttamente il testo greco nel file sorgente, pur di disporre di una tastiera reale o virtuale che consenta l'immissione diretta dei caratteri greci codificati in *utf8*. La documentazione di *babel-greek* definisce questo metodo **literal Greek input**; attraverso un processo di mappatura i caratteri greci in ingresso vengono sostituiti con i caratteri nella seconda metà della tabella dei caratteri greci, non corrispondenti a nessuna delle posizioni dei caratteri ASCII. Con questa immissione letterale del testo greco la sillabazione non funzionava correttamente.

# I caratteri greci con codifica LGR a 8-bit

'000	—	0	~	1	⊠	2	⊡	3	⊢	4	⊣	5	⊤	6	⊥	7	⊦	8	Α	9	Η	10	Ω	11	Α	12	Ψ	13	α	14	ϖ	15	"00	
'020	/	16	\	17	ι	18	ϑ	19	ϒ	20	Ϛ	21	ϛ	22	λ	23	€	24	%	25	ə	26	λ	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	0	48	1	49	2	50	3	51	4	52	5	53	6	54	7	55	8	56	9	57	:	58	˙	59	˘	60	=	61	˙	62	˙	63	"30	
'100	˘	64	Α	65	Β	66	˘	67	Δ	68	Ε	69	Φ	70	Γ	71	Η	72	Ι	73	Θ	74	Κ	75	Λ	76	Μ	77	Ν	78	Ο	79	"40	
'120	Π	80	Χ	81	Ρ	82	Σ	83	Τ	84	Υ	85	˘	86	Ω	87	Ξ	88	Ψ	89	Ζ	90	[	91	˘	92	]	93	˘	94	˘	95	"50	
'140	˘	96	α	97	β	98	ς	99	δ	100	ε	101	φ	102	γ	103	η	104	ι	105	θ	106	κ	107	λ	108	μ	109	ν	110	ο	111	"60	
'160	π	112	χ	113	ρ	114	σ	115	τ	116	υ	117	118	ω	119	ξ	120	ψ	121	ζ	122	«	123	,	124	»	125	˘	126	—	127	"70		
'200	ά	128	ά	129	ά	130	ά	131	ά	132	ά	133	ά	134	ά	135	ά	136	ά	137	ά	138	ά	139	ά	140	ά	141	ά	142	ά	143	"80	
'220	ά	144	ά	145	ά	146	ά	147	ά	148	ά	149	ά	150	ά	151	ά	152	ά	153	ά	154	155	ά	156	ά	157	ά	158	159	"90			
'240	ή	160	ή	161	ή	162	ή	163	ή	164	ή	165	ή	166	ή	167	ή	168	ή	169	ή	170	ή	171	ή	172	ή	173	ή	174	ή	175	"A0	
'260	ώ	176	ώ	177	ώ	178	ώ	179	ώ	180	ώ	181	ώ	182	ώ	183	ώ	184	ώ	185	ώ	186	ώ	187	ώ	188	ώ	189	ώ	190	ώ	191	"B0	
'300	ω	192	ω	193	ω	194	ω	195	ω	196	ω	197	ω	198	199	ι	200	ι	201	ι	202	ι	203	ι	204	ι	205	ι	206	ι	207	"C0		
'320	ί	208	ί	209	ί	210	ί	211	ι	212	ι	213	ι	214	ι	215	ι	216	ι	217	ι	218	Ι	219	ι	220	ι	221	ι	222	Ψ	223	"D0	
'340	έ	224	έ	225	έ	226	έ	227	ο	228	ο	229	ο	230	ο	231	έ	232	έ	233	έ	234	έ	235	ο	236	ο	237	ο	238	ο	239	"E0	
'360	ϊ	240	ϊ	241	ϊ	242	ϊ	243	υ	244	υ	245	υ	246	υ	247	α	248	η	249	ω	250	ρ	251	ρ	252	253	’	254	’	255	"F0		
	"00	"01	"02	"03	"04	"05	"06	"07	"08	"09	"0A	"0B	"0C	"0D	"0E	"0F																		

Font: grmn1000.

## Modifica dei file di descrizione

Il primo problema da risolvere è quello di avere un solo file di descrizione di ogni lingua che disponga degli opportuni modificatori per selezionare regole di sillabazione diverse.

Per *polyglossia* i modificatori sono costituiti dalle varianti che possono essere specificate per ciascuna lingua con opzioni del tipo **chiave = valore**.

Con *babel* i modificatori sono appendici che si possono aggiungere al nome di ciascuna lingua.

### Esempi di modificatori per il latino

```
\setotherlanguage[variant=classic]{latin}  
\usepackage[latin.classic]{babel}
```

# Selezione delle regole di sillabazione

Con minime differenze la selezione delle regole di sillabazione diverse per ciascuna variante della lingua si ottiene dando un significato diverso alla macro `\l@<lingua>` che identifica il numero progressivo a cui le regole di sillabazione di quella `<lingua>` sono state associate durante la creazione del formato. `<lingua>` è il nome della lingua a cui si riferisce il file delle regole di sillabazione quando viene caricato nel file di formato. `<lingua>` non coincide necessariamente con il nome dell'opzione data a *babel* o della lingua specificata a *polyglossia*.

## Selezione delle regole di sillabazione

Perciò i modificatori di *babel* e di *polyglossia* devono provvedere non solo al cambiamento delle regole tipografiche per la composizione di una data lingua, ma devono anche provvedere a selezionare tramite il significato di `\l@<lingua>` anche alla scelta delle giuste regole di sillabazione.

# File di sillabazione per il latino classico

Per il latino moderno e il latino medievale basta un solo file di regole di sillabazione, ed è quello che tutte le distribuzioni del sistema T<sub>E</sub>X hanno usato fino a giugno 2014. È un file che ho predisposto una ventina di anni fa, e poi ho aggiornato via via.

Non rispetta le regole di sillabazione del latino classico perché questo necessita di altre regole che, suppongo, sono il frutto di accordi accademici fra i latinisti recenti; 20 secoli fa gli antichi romani non solo non dividevano in sillabe ma non separavano nemmeno le parole.

# File di sillabazione per il latino classico

Per scrivere i romani non avevano particolari difficoltà; la lettura, invece, era difficile, perché per prima cosa bisognava capire il senso del discorso separando mentalmente le parole e poi, volendo leggere ad alta voce, si potevano pronunciare le frasi lette con una intonazione diversa a seconda del loro significato e delle circostanze.

La sibilla cumana disse. . .

IBISREDIBISNONMORIERISINBELLO



# File di sillabazione per il latino classico

Nel primo secolo d.C. Quintiliano stesso, nelle suo *Institutiones oratoriae*, insegnava a inserire nello scritto un punto a mezza linea per separare le parole, e segni vari per separare le frasi di tipo *comma*, *colon* e *periodus*. Questi nomi sono rimasti ancora oggi in inglese per indicare i segni di separazione delle frasi. L'oratore così poteva leggere i suoi discorsi con facilità e con l'aiuto di altri segni poteva anche modulare l'intonazione delle varie frasi.

Ma di sillabe non si parlava ancora.

## File di sillabazione per il latino classico

I copisti che copiavano i codici, da quando il supporto era diventato il papiro prima e la pergamena poi, scrivevano il testo dentro una gabbia ideale, ma non si preoccupavano di giustificare il margine destro con la cesura in fin di riga, ma aggiustando il modo di scrivere, e all'occorrenza, usavano segni diacritici o simboli speciali che permettevano di accorciare la riga di testo. Per esempio scrivevano *librũ* invece di *librum*, oppure *populusq* invece di *populusque*.

# La sillabazione e la tipografia

Le regole per la sillabazione divennero necessarie quando la composizione meccanica di Gutenberg prese il sopravvento sulla costosissima riproduzione manuale dei testi.

I grammatici si diedero da fare; i tipografi ci misero del loro. Infatti le regole di cesura tipografica in fin di riga sono compatibili ma più stringenti delle regole grammaticali.

Con quasi tutte le lingue in un testo stampato non si va a capo dopo una sola lettera all'inizio della parola, o appena prima dell'ultima lettera della parola.

Per esempio la grammatica dice che la parola *idea* si divide in *i-de-a*, mentre per la tipografia la parola non si divide affatto.

# Regole di sillabazione per il latino classico

Il latino classico si divide in sillabe come l'italiano  
tranne nei seguenti casi.

- 1 Il latino classico non considera dittonghi multi  
gruppi di vocali che in latino moderno e in  
italiano sono considerati dittonghi.

# Regole di sillabazione per il latino classico

Il latino classico si divide in sillabe come l'italiano  
tranne nei seguenti casi.

- 1 Il latino classico non considera dittonghi molti gruppi di vocali che in latino moderno e in italiano sono considerati dittonghi.
- 2 Nel latino classico la *s impura* non esiste come, invece, esiste in italiano. Il gruppo *gn* quasi sempre si separa, ma talvolta resta unito.

# Regole di sillabazione per il latino classico

Il latino classico si divide in sillabe come l'italiano tranne nei seguenti casi.

- 1 Il latino classico non considera dittonghi molti gruppi di vocali che in latino moderno e in italiano sono considerati dittonghi.
- 2 Nel latino classico la *s impura* non esiste come, invece, esiste in italiano. Il gruppo *gn* quasi sempre si separa, ma talvolta resta unito.
- 3 Nel latino classico i prefissi e i suffissi si dividono etimologicamente.

# Regole di sillabazione per il latino classico

Il latino classico si divide in sillabe come l'italiano tranne nei seguenti casi.

- 1 Il latino classico non considera dittonghi molti gruppi di vocali che in latino moderno e in italiano sono considerati dittonghi.
- 2 Nel latino classico la *s impura* non esiste come, invece, esiste in italiano. Il gruppo *gn* quasi sempre si separa, ma talvolta resta unito.
- 3 Nel latino classico i prefissi e i suffissi si dividono etimologicamente.
- 4 Le parole di origine greca si dividono con le regole del greco (classico); in particolare i digrafi *th, rh, ph, ch*, corrispondenti a  $\vartheta, \rho, \phi, \chi$ , non si separano mai.

# Regole di sillabazione del latino classico

I motori di composizione del sistema T<sub>E</sub>X basano la sillabazione sulla grafia delle parole, non sulla loro pronuncia; perciò non sanno nulla sui dittonghi veri e quelli graficamente identici ma che non vengono pronunciati come tali.

Quei programmi non conoscono né l'etimologia delle parole, quindi non sanno in generale distinguere prefissi e suffissi, né se le parole sono veramente latine o di origine greca.

Essi non sanno quando *gn* è un digrafo o va spezzato fra sillabe diverse.



# Difficoltà oggettive nella sillabazione del Latino

Quanto sopra esposto spiega perché è difficilissimo programmare delle regole di sillabazione quando non c'è la certezza che la grafia di una parola non nasconda prefissi o suffissi, o non sia di origine greca o contenga dei dittonghi solo apparenti.

# Come ho risolto il problema

Ho seguito la regola che seguo sempre:

Regola precauzionale

Meglio una cesura mancata che una cesura sbagliata.

Con questa mia regola ho eliminato tutti i problemi relativi ai dittonghi.

## Come ho risolto il problema

Con prefissi e suffissi credo di aver predisposto un buon riconoscimento (non infallibile, ma decente) di quasi tutti i prefissi e i suffissi. Per tutti gli altri bisogna ricorrere alle eccezioni di sillabazione.

Bisogna ricorrere a queste liste di eccezioni anche per le parole di origine greca non riconoscibili come tali; sono riconoscibili quelle che contengono i digrafi *th*, *ph*, *rh*, *ch*, e le lettere *z* e *y* che propriamente non fanno parte dell'alfabeto latino in senso stretto.

Negli ultimi giorni ho ulteriormente aggiornato il file dei pattern arricchendolo in modo significativo rispetto alla prima versione descritta sommariamente nel mio articolo su *ArsT<sub>E</sub>Xnica*.

# Le eccezioni alle regole di sillabazione

Chi scrive in latino classico, quando compone la prima bozza non si preoccupi delle cesure errate in fin di riga.

Quando corregge la prima bozza ha tre possibilità.

- 1 Egli può usare il comando  $\text{\LaTeX}$  `\-` che inserisce nel punto dove è usato un punto di cesura, che però inibisce la cesura nel resto della parola.

# Le eccezioni alle regole di sillabazione

Chi scrive in latino classico, quando compone la prima bozza non si preoccupi delle cesure errate in fin di riga.

Quando corregge la prima bozza ha tre possibilità.

- 1 Egli può usare il comando `LATEX \-` che inserisce nel punto dove è usato un punto di cesura, che però inibisce la cesura nel resto della parola.
- 2 Può usare il carattere attivo `"` che inserisce un punto di cesura ma consente la cesura anche nel resto della parola.

# Le eccezioni alle regole di sillabazione

Chi scrive in latino classico, quando compone la prima bozza non si preoccupi delle cesure errate in fin di riga.

Quando corregge la prima bozza ha tre possibilità.

- 1 Egli può usare il comando `LATEX \-` che inserisce nel punto dove è usato un punto di cesura, che però inibisce la cesura nel resto della parola.
- 2 Può usare il carattere attivo `"` che inserisce un punto di cesura ma consente la cesura anche nel resto della parola.
- 3 Inserisce la parola divisa in sillabe nell'argomento del comando `\hyphenation` o del comando `\babelhyphenation`.

# Esempio

## Testo Latino

Flumen est Arar, quo per fines Haeduorum et Sequanorum in Rhodanum influit, incredibili lenitate, ita ut oculis in utram partem fluat iudicari non possit. Id Heluetii ratibus ac lintribus iunctis **trans-ibant**. Ubi per exploratores Caesar certior factus est tres iam partes Heluetios id flumen traduxisse, quartam fere partem citra flumen Ararim reliquam esse, de tertia uigilia cum legionibus tribus e castris profectus est ad eam partem peruenit quae nondum flumen transierat. Eos impeditos et inopinantes adgressus, magnam partem eorum concidit: reliqui sese fugae mandarunt atque in proximas siluas abdiderunt.

# I file per il latino

Ho già provveduto a rivedere i file di descrizione della lingua latina; ho aggiornato sia *latin.ldf* sia *gloss-latin.ldf* e li ho già caricati su CTAN. Per *babel* le cose sono già a posto; per *polyglossia* il curatore di quel pacchetto non ha ancora provveduto.

Lo stesso può dirsi per i nuovi file di sillabazione del latino classico; tutto a posto per *babel*; sono già installati per *polyglossia* ma *gloss-latin.ldf* ancora non vi accede.



# I file per il latino

Come curatore *pro tempore* dei due tipi di file non ho avuto problemi, ma non ho accesso a CTAN per provvedere alle cose ancora in sospeso.

Sarebbe opportuno però che di queste cose si occupasse un latinista vero e non un dilettante come me.

## Sillabazione greca

Come ho detto non ci sono problemi con *polyglossia* e la composizione in greco con Xe<sub>3</sub>L<sup>A</sup>T<sub>E</sub>X o LuaL<sup>A</sup>T<sub>E</sub>X.

I problemi esistono solo con pdfL<sup>A</sup>T<sub>E</sub>X e *babel*.

Si tratta di aggiustare il file *greek.ldf* e di creare i pattern giusti per l'immissione del testo sia con i caratteri greci sia con i comandi per i diacritici.

# Immissione del testo greco con pdf $\text{\LaTeX}$

Vale la pena di ricordare i tre metodi per immettere il testo greco; essi sono basati

- sulla traslitterazione latina e le legature; oppure
- sulla traslitterazione latina e i comandi per i diacritici; oppure
- sull'immissione letterale del greco e i comandi LICR

I comandi LICR ( [\$\text{\LaTeX}\$  Internal Character Representation](#)) sono speciali macro interne che fanno da separazione ideale fra l'input in qualunque codifica e l'output con font codificati LGR.

## Esempi di immissione del greco

### Immissione del greco

Traslitterazione e legature: >An'axandros

Traslitterazione e comandi: \`>An\`'axandros

Greco letterale: 'Ανάξανδρος

### Composizione della parola

In tutti e tre i casi il nome con l'iniziale maiuscola viene composto come 'Ανάξανδρος.

# Immissione del testo greco con pdf<sup>L</sup>A<sup>T</sup>E<sup>X</sup>

Fino a giugno 2014 la sillabazione greca funzionava solamente per l'immissione con la traslitterazione latina e le legature e solo per l'ortografia politonica moderna.

Bisognava poter accedere anche alle altre due ortografie del greco e predisporre i file di sillabazione per gli altri due metodi di immissione del testo greco.

# Soluzione per il file di descrizione della lingua

Ho contattato Günter Milde, il curatore dei file relativi alla lingua greca, non della sillabazione greca, e gli ho fornito una serie di patch per il file [greek.ldf](#). Milde ha rivisto e controllato i patch, ha aggiunto anche altre modifiche relative al problema in questione, e ha già caricato su CTAN la nuova versione di quel file.

Quindi oggi anche con pdfL<sup>A</sup>T<sub>E</sub>X e *babel* si possono specificare separatamente le tre ortografie greche invocando le regole di sillabazione “giuste” per ciascuna ortografia.

## Soluzione per i file di sillabazione

Oggi i file di sillabazione caricati nei file di formato sono ancora quelli che erano prima di giugno 2014, ma almeno sono selezionabili.

Si noti: il greco dispone di sei file di sillabazione, tre per essere usati nei formati che richiedono *polyglossia* (uno per ciascuna delle tre varianti della lingua) e sono basati sui caratteri UNICODE (o UTF-8); tre destinati ai formati che non sono *Unicode aware* e accettano solo font i cui glifi sono identificati con indirizzi a 8-bit.

I file di sillabazione per *babel* sono fortemente carenti se si usano due dei tre metodi per l'inserimento del testo greco nei file sorgente.

# Perché i file di sillabazione sono carenti

I tre file **oggi disponibili** per la sillabazione greca e da usare per generare i file di formato per i motori che sono solo **8-bit aware**, cioè *tex*, *etex* e *pdftex* sono adatti per introdurre il testo greco solo con **il metodo della traslitterazione latina con le legature**.



# Perché i file di sillabazione sono carenti

Per tornare al nostro esempio di Ἀνάξανδρος, la scrittura >**An'axandros** richiede solo che i pattern di sillabazione contengano caratteri ASCII e quella stringa di caratteri latini è totalmente ASCII. I pattern per sillabare sono scritti solo con caratteri ASCII dall'indirizzo 32 all'indirizzo 126; i pattern funzionano a dovere a la parola viene divisa in sillabe correttamente.

## Perché i file di sillabazione sono carenti

Dopo l'eventuale scomposizione in sillabe vengono inviati i singoli caratteri al font di uscita che è codificato LGR: i caratteri latini vengono mappati su quelli greci e le legature, oltre che le crenature, vengono eseguite; nel file di uscita appaiono i caratteri accentati, benché si trovino nella metà della tabella LGR con indirizzi più alti.

In particolare, lo spirito dolce non si lega con la alfa maiuscola, ma viene crenato; l'accento acuto si lega con la alfa minuscola e produce la  $\acute{\alpha}$ .

## Perché i file di sillabazione sono carenti

Invece se si usa il metodo della traslitterazione latina con i comandi per i diacritici, immettendo la stringa `\>An\'axandros` nel file sorgente, le due macro `\>` e `\'` seguiti dalle rispettive alfa, indirizzano direttamente i caratteri accentati (**indirizzi superiori a 126**) e crenati della codifica LGR; poi la stringa viene esaminata dall'algoritmo di sillabazione, che non trova **nessun pattern che contenga caratteri con indirizzi superiori a 126**, quello dell'ultimo carattere ASCII.

L'algoritmo di sillabazione, non trovando pattern adeguati, divide in sillabe la parola, solo fra la  $\nu$  e  $\delta$ .

# Perché i file di sillabazione sono carenti

- 1 Va da sé che se si usa l'immissione letterale del testo greco, le macro LICR mappano direttamente i caratteri greci codificati in entrata con `utf8` con le posizioni da 128 a 255 della codifica LGR, esattamente come succede con il metodo della traslitterazione latina con le macro dei diacritici.

# Perché i file di sillabazione sono carenti

- 1 Va da sé che se si usa l'immissione letterale del testo greco, le macro LICR mappano direttamente i caratteri greci codificati in entrata con `utf8` con le posizioni da 128 a 255 della codifica LGR, esattamente come succede con il metodo della traslitterazione latina con le macro dei diacritici.
- 2 Perciò l'elegante metodo dell'introduzione letterale del testo greco presenta oggi gli stessi inconvenienti del metodo della traslitterazione latina.

# Perché i file di sillabazione sono carenti

- 1 La “colpa” di questa carenza è dovuta ai tempi di esecuzione delle varie azioni.

# Perché i file di sillabazione sono carenti

- 1 La “colpa” di questa carenza è dovuta ai tempi di esecuzione delle varie azioni.
- 2 Con la traslitterazione latina e le legature **prima** si esegue l'algoritmo di sillabazione **poi** si manda la parola in uscita con il font codificato LGR che provvede alle sostituzioni mediante i comandi di legatura.

# Perché i file di sillabazione sono carenti

- 1 La “colpa” di questa carenza è dovuta ai tempi di esecuzione delle varie azioni.
- 2 Con la traslitterazione latina e le legature **prima** si esegue l'algoritmo di sillabazione **poi** si manda la parola in uscita con il font codificato LGR che provvede alle sostituzioni mediante i comandi di legatura.
- 3 Con gli altri due metodi **prima** si sostituiscono i caratteri e le macro in entrata con gli indirizzi dei glifi codificati LGR e **poi** si esegue l'algoritmo di sillabazione.



## Soluzione del problema

La soluzione del problema consiste dunque nel duplicare i pattern di sillabazione che contengano accenti leganti con gli indirizzi superiori a 126 dei rispettivi caratteri accentati nella codifica LGR. Non basta: se la parola inizia con una vocale maiuscola l'accento si conserva come tale ma implica solo una crenatura con l'iniziale maiuscola. Ne segue che ci vogliono altri pattern ibridi in caratteri ASCII e con gli indirizzi esadecimali per i valori superiori a 126. Per sillabare Ἀνάξανδρος sono perciò necessari almeno i seguenti tre pattern

### Esempio di pattern greci completi

>an'a2x1an **^^82n^^882x1an** >an**^^882x1an**

## Soluzione del problema

Ho già creato i tre file di sillabazione completi secondo la soluzione indicata sopra, e li ho inviati a Dimitrios Filppou, che è il curatore della sillabazione della lingua greca. Tocca a lui controllare la correttezza di quel che ho scritto io, e di caricare i tre nuovi file su CTAN.

Io non sono greco, quindi potrei avere fatto certi errori che l'occhio di un greco troverebbe immediatamente. D'altra parte ai file di sillabazione ho aggiunto fino a circa **1500 pattern**; fra tanti l'errore potrebbe essere scappato.

## Esempi di sillabazione

Ecco alcune parole greche con l'ortografia classica divise in sillabe.

ἀ-δυσ-ώ-πη-τος

ἀν-αί-δει-α

ἀν-αμ-φισ-βή-τη-τος

Ἄ-νάξ-αν-δρος

ἀ-νάξ-αν-δρος

ἄν-οι-κος

ἀ-πρό-σκο-πος

ἀ-πρό-σκο-πτος

εἰσ-ίν

ἐ-κλεί-σα-μεν

ἐ-κλη-ροῦ-τον

ἐ-κρυ-στάλ-λι-σα

Ἐλ-λήσ-πον-τος

ἐλ-λήσ-πον-τος

ἐν-ί-δρω-σα

εὐ-σύν-ε-τος

κα-λωσ-ό-ρι-σα

με-λα-νό-ρι-σα

με-λαν-όμ-μα-τος

μο-γισ-α-ψε-δά-φα

ξυν-α-γεί-ρα-το

πα-λιν-ά-γρε-τος

πα-νισ-δό-μην

προσ-εισ-πτάσ-σω

προ-ὑπεξ-ορ-μά-ω

ὑ-οσ-κύ-α-μος

ὑ-πεκ-λή-ψο-μαι

χα-ρίσ-αν-δρος

δισ-χί-λι-οι

## Lavori in corso

L'operazione greco e latino non è finita; io ho fatto tutto quel che potevo per sollevare dalla fatica materiale i curatori delle varie parti del sistema T<sub>E</sub>X.

I piccoli aggiornamenti del file latino *gloss-latin.ldf* sono già stati inviati al curatore di queste cose per *polyglossia* ma sembra che sia piuttosto lento ad aggiornare la sua parte, sebbene sia già tutto fatto e collaudato.

Il controllo della revisione dei file di sillabazione greca vengono rivisti con molta calma e speriamo che non manchi molto a quando verranno caricati su CTAN. Però i problemi sono stati individuati e le soluzioni sono state trovate.

Happy  
Latin and Greek  
L<sup>A</sup>T<sub>E</sub>Xing!