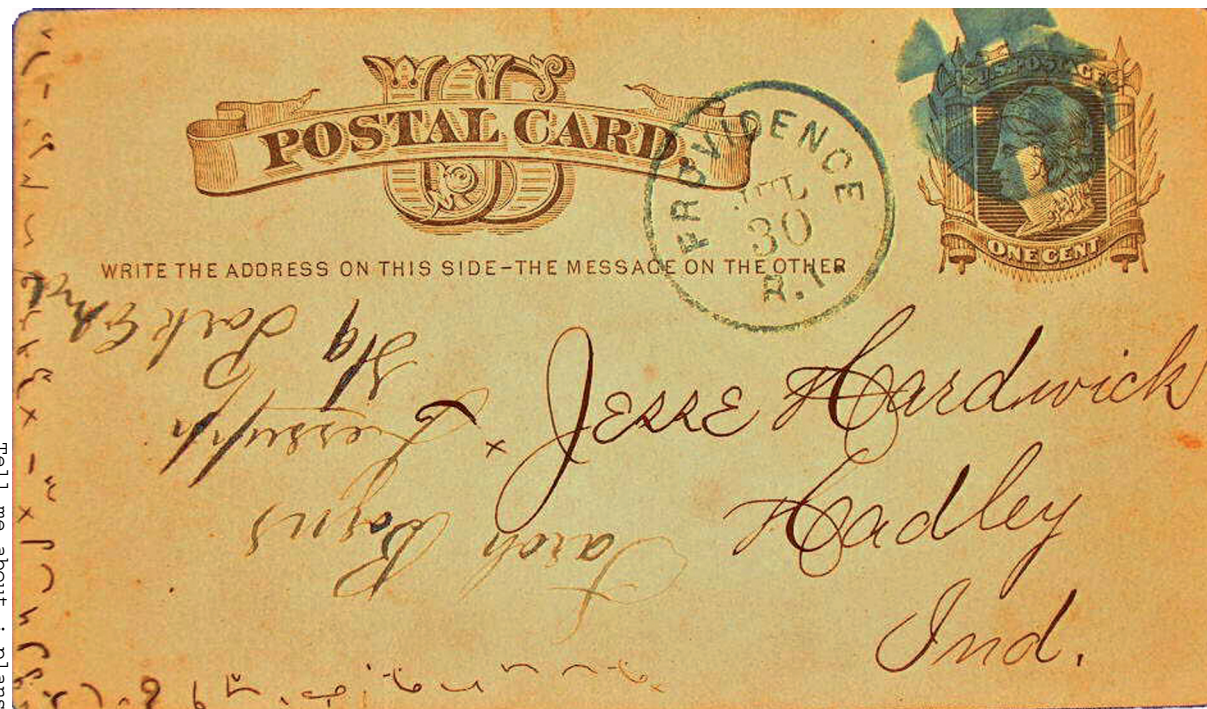


OCR vs. text2Pitman



OCR \rightsquigarrow

How old are you? It is time to close ; office now I'm ; office.



← text2Pitman —

Tell me about plans. How old are you? It is time to close office now I'm office.

Pitman Shorthand: Basics

stenem* = outline + diacritics

(t)(r)(t)(r) [e][i][ou][i]

word	pronunciation	metaform (4 segments)
territory	{ t * e . r i . t o u r . r i y }	(t)[e]&(r)&[i](t)[ou]&(r)[i]

strokes = consonant signs, e.g. (t), (r):

writing vowel signs:

in between:

*a glyph of one or more words phonetically written in Pitman shorthand = consonantal + vowel part

short forms:

. the, / and, | on, | but, \ to, ^ I, o you

punctuation marks:

⋈ period, ? question sign

strokes:

| t, also it; | d (firmly written t), also do;

/ r, also are; ∪ n, above line in: ∪

phrases:

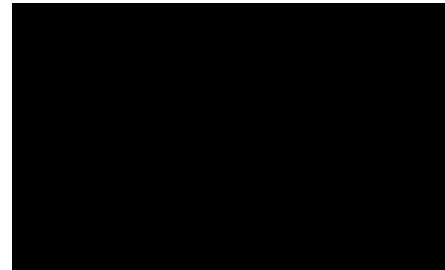
^ I do, | do you, / you are

“/ old, Father William,” . young man said,

“/ your hair has become very white;

/ yet o incessantly stand | your head —

| think, _ your age, it is ✓|?”



“∪ my youth,” Father William replied \ his son,

“I feared | might injure . brain;

|, now that I’m perfectly sure I have none,

Why, ^ | again / again⋈”

Lewis Carroll: Father William’s song
(from Alice in Wonderland)

dvitype-clone → DjVu*

(Writing)
Exercises →

Section Twenty ¹³³
EXERCISE 55 Write in shorthand, afterwards checking from the key (phrasing is indicated by hyphens):

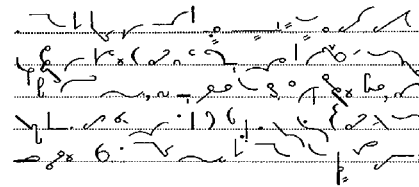
The cafe at-the top-of-the hill called Kate's Cook-In is well worth a visit as-those-who-have-been will tell-you. They serve you with very-good meals at low-prices and in-your-present state-of lack-of money, you-have good reasons for spending as little as-possible. At-all-events, you-will-be better-off taking one hot meal a day there than eating the poor fare that-they serve up in-your guest house. This-is a copy-of-the menu at-the cafe for Tuesday-of-last week.

vs.

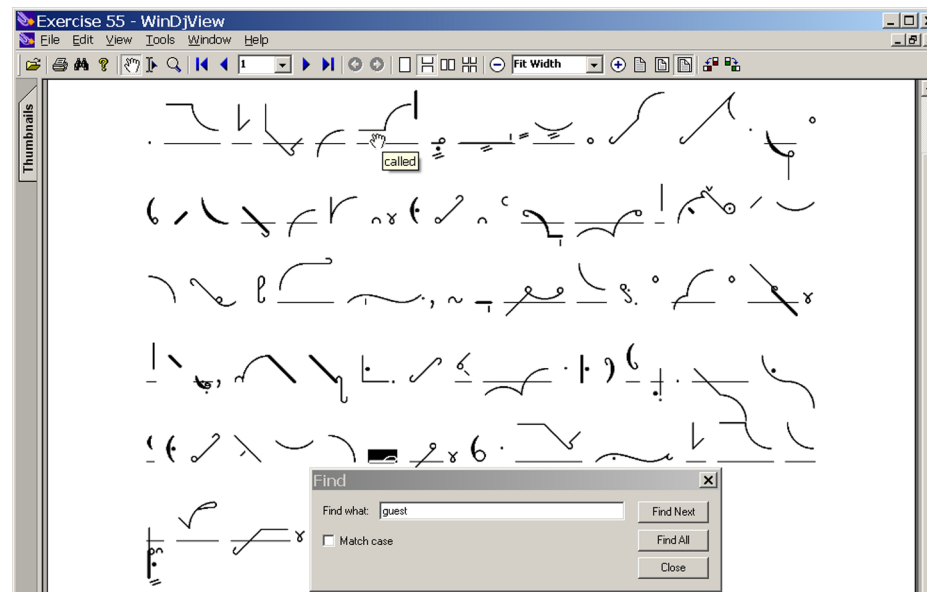
¹⁸⁶ *Shorthand: Pitman 2000*

time to spare, to inspect them. These gloves, which have been bought as part of the stock of a warehouse damaged by fire, are of very high quality and are being sold at a low profit margin.

Exercise 55



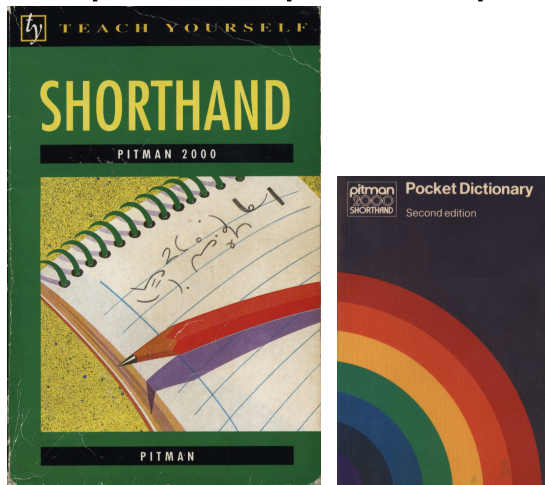
← Keys to
Exercises



* Annotated, searchable DjVu files (viewable also with a DjVu browser plugin)

Writing Pitman Shorthand with METAFONT and L^AT_EX

- Pitman shorthand* → Pitman 2000:
simplified alphabet, phonetic writing, short forms and phrases



- text2Pitman: <http://www3.rz.tu-clausthal.de/~rzsjs/steno/Pitman.php>
and [DEK.php](#), [Gregg.php](#), [Suetterlin.php](#)

*1837

†S. J. Šarman, Computing Centre, Clausthal University of Technology, Germany

Cave canem



(huf (hfm
C A V E C A N E M

old Roman cursive

3c K
cav-e canis
-em

Tironian notes*

(,k)(a,w)(e,) (,k)(a,n)(e,m)

DEK

(,ka)(,v)(e,) (,ka)(,n)(e,m)

Herout-Mikulík

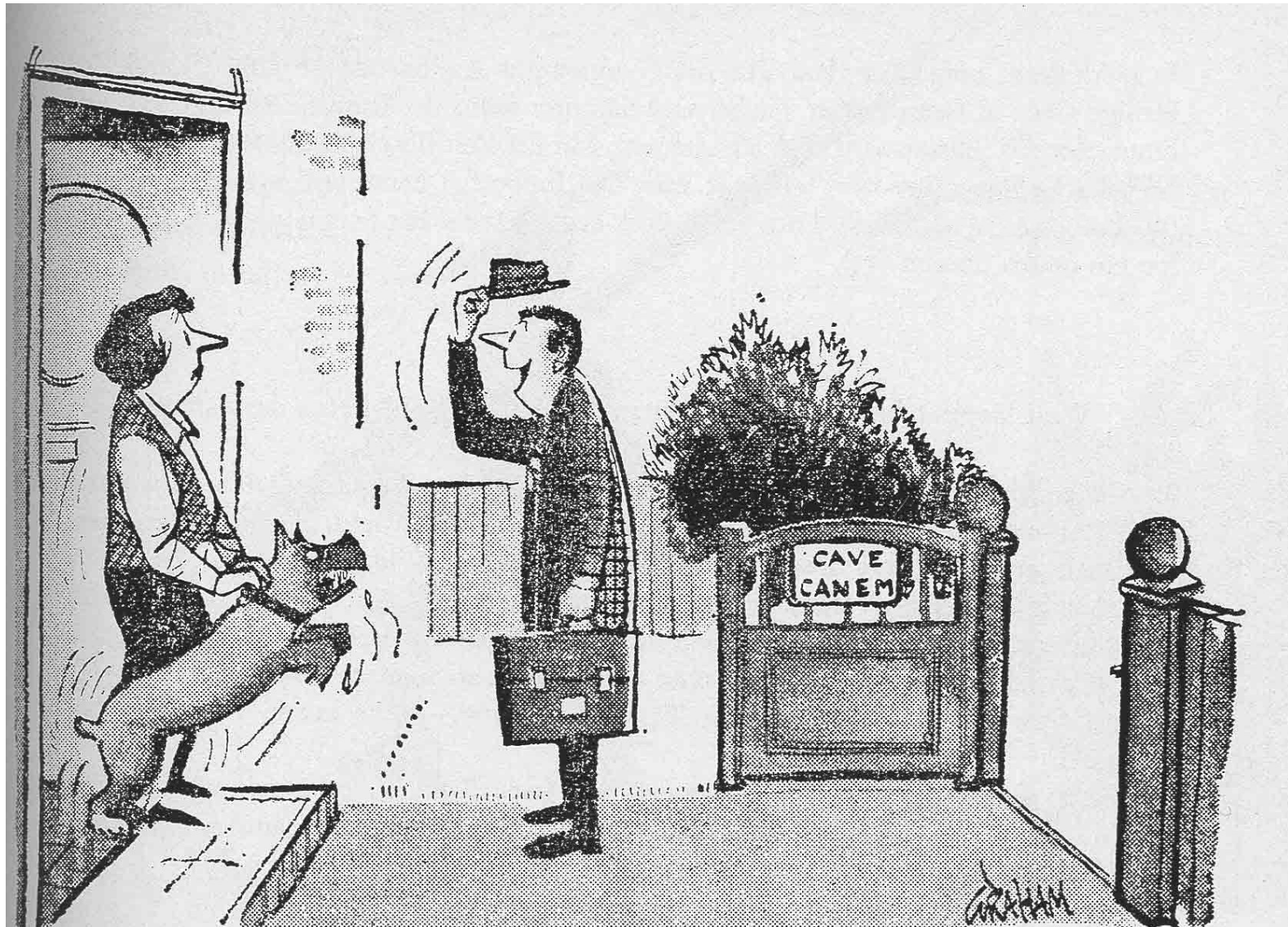
(k)-[a](v)-[e] (k)-[a](n)-[e](m)

Gregg

(k)[a]&(v)[e] (k)[a]&(n)[e]&(m)

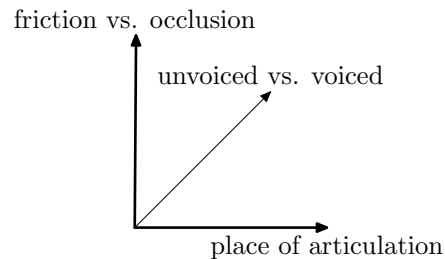
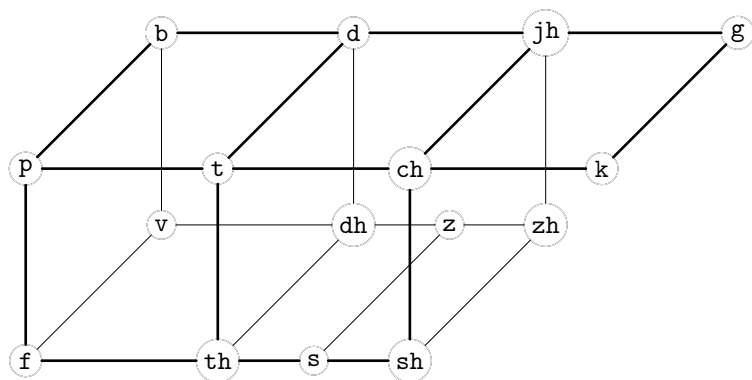
Pitman

*courtesy of Dr. Hellmann



“Mrs. Canem . . . ?”

Consonant Signs: Strokes



plosives

$\diagdown \diagdown$ \parallel \parallel — —
 (p) (b) (t) (d) (ch) (jh) (k) (g)
 (f) (v) (th) (dh) (s) (z) (sh) (zh)

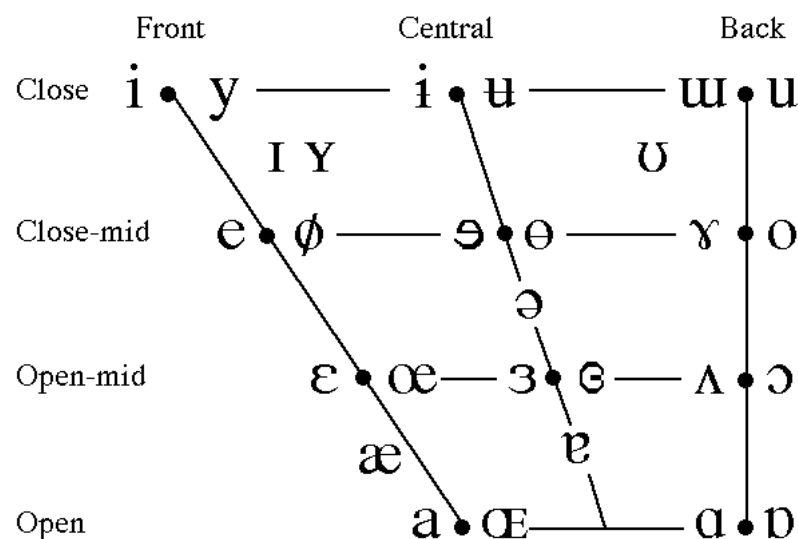
fricatives

$\diagup \diagup$ (()) $\diagdown \diagdown$

nasals			liquids		
\frown	\smile	\smile	\frown	\diagup	\diagdown
(m)	(n)	(ng)	(l)	(r)	(_r)

Trnka, B.: A Phonological Analysis of Present-day Standard English. Prague 1935

Vowel, Diphtong and Triphone Signs



Where symbols appear in pairs, the one to the right represents a rounded vowel.

Jones IPA vowel quadrilateral

triphone signs

“diary” $\text{ɪ} \text{a} \text{ɪ}$ “loyal” $\text{ɔ} \text{ɪ} \text{ə}$ “towel” $\text{t} \text{ə} \text{w}$ “fewer” $\text{f} \text{u} \text{ə}$ “idea” $\text{i} \text{d} \text{ə}$

	place	front	back
open	1st	$[\text{a}]$ $[\text{ah}]$ at pa	$[\text{o}]$ $[\text{oo}]$ odd saw
	2nd	$[\text{e}]$ $[\text{ei}]$ ed aid	$[\text{uh}]$ $[\text{ou}]$ up no
close	3rd	$[\text{i}]$ $[\text{ii}]$ ill eel	$[\text{u}]$ $[\text{uu}]$ took coup
diphtongs			
	1st	$[\text{ai}]$ “my” $[\text{oi}]$ “joy”	
	3rd	$[\text{ow}]$ “out” $[\text{yuu}]$ “few”	

$C ::= \text{b} \mid \text{p} \mid \text{d} \mid \text{t} \mid \text{v} \mid \text{f} \mid \text{dh} \mid \text{th} \mid \text{zh} \mid \text{sh} \mid \text{ng} \mid \text{n} \mid \text{m} \mid \text{l} \mid \text{r} \mid \text{w} \mid \text{hw} \mid \text{y} \mid \text{h}$









Segments*

size:







(1) [ei]:t (1) [ei] (1) [ei]:tr (m) [ii]:t (m) [ii] (m) [ii]:tr

prefixes:















$(p,r)[ei]$ $(f,r)[ii]$ $(p,l)[ei]$ $(f,l)[ii]$ $,s[u](p)$ $,st[e](p)$

suffixes:



suffixes:





, s(p,r) [ei] , s [uh] (p,l)
(p) [ii] , s
(p) [ii] , sis
(p) [ou] , st
(p) [ou] , sts

(m) [aa] , str
(p) [e] ; n
(p) [e] ; n, s
(p) [uh] ; f
(p) [uh] ; f, s
(p) [a] ; shn
(p) [a] ; shn, s

* $3 \times 2^4 \times 2^4$ forms possible

Stenems: Dis/joining Segments

morphological affixes



^com[o] (n)



(g) [ou] ~ing



[a] (n)+Upp

numbers, past tense

24



(_two_) (_four_)

[aa] (s)&(k)/(t)

(sh) [ou] / (d)

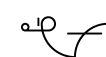
intersections



tax form



company boom



successfull company

left vs. right ,s



cassette



unsafe



traceable



desk



bestow

n/m, cusps



testimony



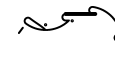
number



figure

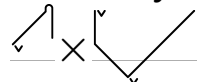


reply



stenographers

misc



writer x type ~



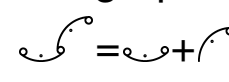
original



machines



statistics



senseless

text2Pitman

input: *Do you think, at your age, it is right?* → tokenizer →
 → do you_think_,_at_your_age_,_it is_right_?

Unisyn lexicon → stenemizer → mf run →

	token	pronunciation	metaform	stenem
1	,		(_comma_)	,
2	?		(_question_)	?
3	age	{ * ee jh }	[ei](jh)	/
4	at	{ * a t }	[a](t)	
5	do you		(d)&(_u_)	h
6	it is		(t),s	h
7	right	{ r * ai t }	(r)[ai]&(t)	✓
8	think		(th)	(
9	your		(_r))

mf-file: beginS(7);I(,r,,,);V(ai,-1);J;I(,t,,,);J;endS; %right

latex → dvips → gs → ppmtogif

h (, |) ✓ / h ✓ ?
 400


Phonetic Writing

Unisyn* multi-accent lexicon:



asked;;VBD/VBN; { * ah s k }> t > ;{ask}>ed>;89620

acted;;VBD/VBN; { * a k t }.> I7 d > ;{act}>ed>;3188





English homographs → Pitman heterographs:  (live),  (wind),





 latex;1,rubber;NN; { l * ee . t e k s };33

 latex;2,computing;NN; { l * ee . t e k };33

 read;1;VB/NN/NNP/VBP; { r * ii d }× read;2;VBN/VBD; { r * e d };94567

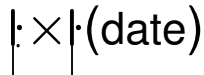

English homophones → Pitman homographs:

 right=rite=wright=write { r * ai t };70806  l=eye,  one=won,  not=knot

but: in × inn,  we ×  wee

ignoring **schwas**[†]?

@ backtransform

data	{ d * ee . t == @ }	d(ee,a)t(@,a)	(d)[ei]&(t)[a]	 (date)
poster	{ p * ou s t }.> @r r >	p(o,ou)st(e,@r)r	(p)[ou],str	

*<http://www.cstr.ed.ac.uk/projects/unisyn/>

[†]the most frequent “(non)vowels”

Stenemizer: pronunciation → metaform

cascaded **two-level** finite state transducers (FSTs)*:

	rewrite rule	
{ * ah s k } > t >	" ah" -> " aa", " " ["*" "{" "}"] -> 0	aa s k > t >
aa s k > t >	" > t >" -> "/" t"	aa s k / t
aa s k / t	conso @-> "(" ... ")",, vowel @-> "[" ... "]"	[aa] (s) (k) / (t)
[aa] (s) (k) / t	(Vowel) Conso (Vowel) @-> ... "&" _ (Vowel) Conso	[aa] (s) & (k) / (t)

ambiguities: { s t * ar r t } → ,st[aa](r):t $\not\sim$ Or ,s(t)[aa]&(r):t $\not\sim$
 { l * ee . t e k } → (l)[ei]&(t)[e]&(k) $\not\sim$ Or (l)[ei]:t&[e](k) $\not\sim$

context sensitive rewrite rules in phonology†:

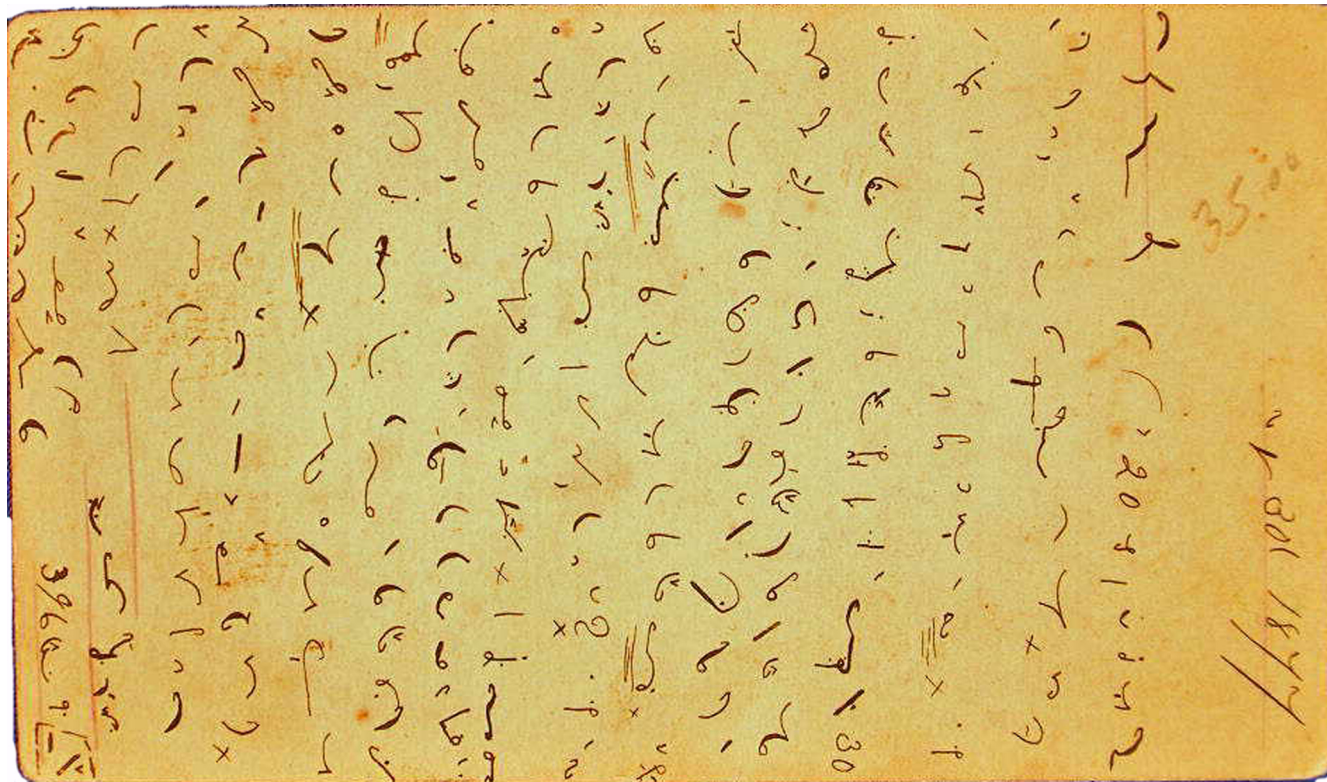
	rewrite rule	
{ * a k t }.> e d >	" e" -> " I7" [t d] " }.>" _ " d >"	{ * a k t }.> I7 d >
{ * ah s k } > e d >	" e d" -> t unvoiced " .>" _ " >"	{ * ah s k } > t >

*XEROX xfst (<http://www.fsmbook.com>)

†Chomsky and Halle (1968): English spelling is coming “remarkably close to being an optimal orthographic system for English”

traitor | { t r * ei . t == @r r } | (t,r)[ei] & (t,r) | × (t,r)[ei]:tr |.

Home Exercise*



*hint: turn the slide 90° to the left

Diary

London Review OF BOOKS

By Leah Price

Published: December 4, 2008

Stenography is dying out; so are stenographers. When I mention that I'm working on the history of shorthand, people tell me that their mother knew shorthand, or their grandmother, or their husband's first wife.

... Journalism degrees in Britain still include a speedwriting test; ...

In the US, court reporters have abandoned stenotype machines, whose keyboards use chord-like combinations to represent sounds, for a technique called voice writing. The 'writer' - really **a speaker - repeats testimony into a microphone** nestled in a hand-held mask that prevents her voice from being heard in court; **the recording is later transcribed, usually with speech-recognition software.** ... machine stenography takes three years to learn, voice writing six months.

... Gregg was to Pitman as Windows is to Linux, ...

The Handwriting Is on the Wall

Researchers See a Downside as Keyboards
Replace Pens in Schools

By Margaret Webb Pressler

Washington Post Staff Writer

Wednesday, October 11, 2006; Page A01

The computer keyboard helped kill shorthand, and now it's threatening to finish off longhand.

When handwritten essays were introduced on the SAT exams for the class of 2006, just 15 percent of the almost 1.5 million students wrote their answers in cursive. The rest? They printed. Block letters.

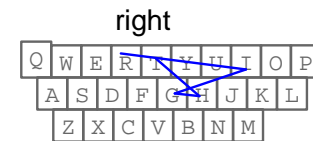
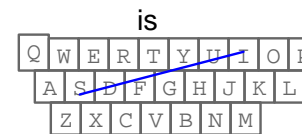
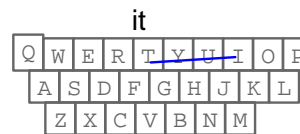
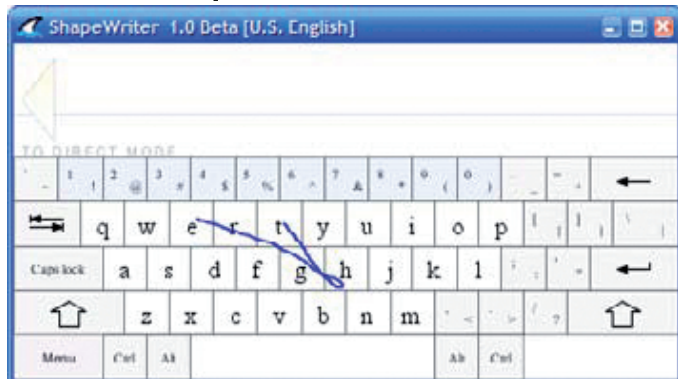
<http://www.washingtonpost.com/wp-dyn/content/article/2006/10/10/AR2006101001475.html>

SHorthand Added Rapid Keyboarding

Each pattern of a word is formed by the trajectory from the 1st to the last letter on a keyboard — scale and location independent

“the”

with ShapeWriterPro on iPhone:



compare with “are”=“a”+“r” in Willis shorthand (1602): $\wedge + _ \Rightarrow \wedge _$