

An International Phonetic Alphabet

Dean R. Guenther

Washington State University, Computing Service Center, Pullman, WA. 99164-1220, USA

Bitnet: GUENTHER@WSUVM1 Internet: GUENTHER@WSUVM1.CSC.WSU.EDU

Janene K. Winter

American Mathematical Society, P.O. Box 6248, Providence RI. 02940, USA

Internet: JXW@MATH.AMS.COM

Abstract

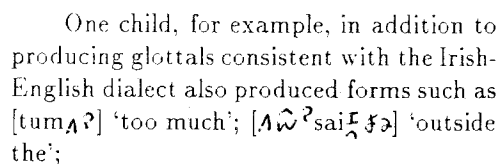
This paper discusses the development of an International Phonetic Alphabet (IPA) font for use with T_EX and the Computer Modern fonts primarily from the viewpoint of a manager. Included will be discussion on how to use this font and the direction we can go from here with it.

The Need for an IPA Font

Since 1983, T_EX has found extensive use at Washington State University (WSU). Although at many sites, mathematics typesetting is a primary reason for using T_EX, at WSU, the area of textual typesetting has always been the main thrust of T_EX usage. In fact, the pioneer project for T_EX at WSU was a critical edition of Robert Burton's seventeenth-century *The Anatomy of Melancholy* by Professors Thomas Faulkner and Nicolas Kiessling.

T_EX's popularity as a text processor at WSU grew each semester and, as a whole, T_EX earned rave reviews from its users. However, one complaint, or limitation, that was frequently echoed throughout the halls of the WSU Computing Service Center (especially from the liberal arts departments), was the lack of fonts available for use with T_EX which contained certain special characters that were often needed (such as the Old English "thorn" (þ) or any unusually accented character of a native American or some other dialect). These "complaints" provided the seed of thought for a future METAFONT font that would contain some of these special characters.

Around 1986, this "future font" was escorted from the recesses of the "some-day-it-would-be-nice" projects, to front-row status as several linguists brought to our attention the need for an IPA font to work with T_EX and Computer Modern fonts. Up to that time, for most linguists who wanted to use a diacritic or special character, it was frequently necessary to have the character hand drawn in, as illustrated in Figure 1. Early word processors allowed for the ability to superscript/subscript and overlay of some characters, and the introduction



One child, for example, in addition to producing glottals consistent with the Irish-English dialect also produced forms such as [tʊm_hʰ] 'too much'; [A^hɔ̃^h sɔi_h fɔ̃] 'outside the';

Figure 1: Hand drawn phonetic characters

of T_EX gave the phonetician much more flexibility as pointed out by Christina Thiele (Carleton University, Canada) in *TEX, Linguistics, and Journal Production*.

Still, there were frequently used characters that could not be created by combining existing characters or accents. In fact, some characters, such as the ejective (p̰) required the removal of bits, rather than the addition. Certainly we could have equipped each linguist on campus with a bottle of White Out, or perhaps we could have created a White Out font, but neither seemed as effective as a new IPA font.

There were already some IPA fonts available. However, they were not designed with METAFONT, and they did not always work well with Computer Modern. Figure 11 shows the ph10 font designed by Jean-Pierre Paillet in 1983 in consultation with William Cowan, editor of the *Canadian Journal of Linguistics*. It was a bitmapped font created by hand, designed to work with amr10. Figure 2, a sample from Richard Rhodes' article in the

(5) Vowel harmony

Orthography	Standard	Métchif
pouilleux	[puj ^ö]	[puj ^ü] ~ [p ^ü j ^ü]
mesure	[mez ^{ür}]	[miz ^{ür}] ~ [m ^ü z ^{ür}]
chevreuil	[šəvr ^ö]	[šuvr ^ü]
fusil	[füz ^{ij}]	[fizi]
musique	[müz ^{ik}]	[müz ^{ük}]

Figure 2: ph10 with amr10

Actes du dix-septième congrès des algonquistes, shows that this marriage worked well. When the crisper Computer Modern became available, ph10 proved to be too heavy. Figure 9 shows the use of ph10 with cmr10. This illustration comes from Terry Piper's article "On the Difference Between L1 and L2 Acquisition in Phonology" in the *Canadian Journal of Linguistics*. The only way to enhance the use of ph10 with Computer Modern would have been to tweak the pixels for each character by hand. One of these character bitmaps is illustrated in Figure 12.

The Development

To begin with, Janene Winter began experimenting with various characters, writing the code from scratch as she was attempting to learn METAFONT. These attempts created some rather unusual (and very non-IPA) characters during the learning process. This also created no small amount of frustration as Janene attempted to learn METAFONT on her own, for nobody else on campus knew or used METAFONT. Adding to this already less than ideal situation was the fact that there was no good graphics previewer available on VM/CMS so she could preview her work on the terminal. Each iteration of a character had to be printed on one of the University's laser printers, which was shared with about 2,000 other users. As can be imagined, this became a painstakingly slow process, not to mention all of the trees that went through the laser printer.

The first major breakthrough in the development of the WSU IPA was the discovery that the Computer Modern METAFONT code really was decipherable and could be used as a foundation on which to build from existing METAFONT characters, and whenever possible simply modify them to fit the required character. Shortly after this, a graphics terminal was obtained which allowed on-screen previewing of the characters via two preview programs. One from Malki Cymbalista (Israel) and the other from Georg Bayer (Germany).

With these two monumental accomplishments, it became obvious that a METAFONT IPA font could indeed be more than a "wouldn't-it-be-nice-if-we-could" project. Now that we were much more clear on the "how" were we going to do this, the next step was to specifically define "what" were we going to do—which characters, and which style or design of the characters would be used. Finding decent typographical examples of the IPA characters was more difficult than anticipated. Most books found in the campus library that showed the characters at all were at least 40 years old, and the largest character samples were printed in 10pt. Finally, Pullam and Ladusaw's *Phonetic Symbol Guide* was discovered. This proved to be a major find, as it had excellent IPA examples in a larger 19pt size. And more importantly, the equivalent of an x-height and baseline was also traced in for all characters illustrated.

The base design for the IPA was 10pt. Making additional sizes came with their own sets of problems. For example, going from 10pt to 9pt was usually fine, but going to 8pt would sometimes cause problems in some of the point and stroke definitions. Additional METAFONT errors were encountered when slanted faces were generated. METAFONT's "strange turning path" message became a rather familiar sight at that time. In most cases, even though the error messages were generated, the characters themselves looked fine. Eventually, the magic combination of correct parameter values and point definitions was found and all error messages disappeared.

Another pitfall that was not obvious at first was the misplacement within the character box of a few of the characters. These were simply Computer Modern characters rotated or flipped, or sometimes both. The first few attempts at transforming these characters often gave the correct visual result, but later examination would show that the baseline or height was wrong. The baseline problems would usually be easily detected by placing the character in a line of Computer Modern text. The height problems would not show up until the character was used with an accent.

These placement problems occurred because the Computer Modern METAFONT code was not changed. It was simply "transformed," and therefore, the original Computer Modern character's height, depth and x-height were being used. This problem was solved simply by changing all of METAFONT's point definitions that were relative to the character's height, depth or x-height, from the

A shibilant is a term occasionally found for a fricative corresponding to a “hushing” sound, e.g., IPA [ʃ] (more technically, a grooved laminal fricative).

Figure 3: Using the IPA Characters

original Computer Modern character, to the new transformed IPA character.

Transforms for the slanted fonts brought further difficulties because METAFONT slants a character before it executes any of its transformations. This caused the IPA characters that were created by rotating existing Computer Modern characters that were then slanted, to be slightly off center in their character boxes. It took a while to master all the various transforms.

Karen Mullen (University of Louisville, Kentucky, USA), Christina, and many others helped Janene in debugging the various characters.

Finally, Figure 10 illustrates the use of the WSU IPA with Computer Modern. This example comes from Claire Lefebvre’s article “Instrumental Take-Serial Constructions in Haitian and in Fon”.

How to use the IPA

To use the IPA you must first embed a collection of macros by entering

```
\input ipamacs
```

Each of the IPA characters has a control sequence defined in IPAMACS. The names chosen for these control sequences were based in large part on the standard names used in Pullam and Ladusaw. For example,

```
A shibilant is a term occasionally
found for a fricative
corresponding to
a ‘‘hushing’’ sound, e.g., IPA
[\esh] (more technically,
a grooved laminal fricative).
```

yields the illustration in Figure 3.

The IPA character macros are defined so they can be used with the Computer Modern characters and accents without the need for delimiting curly braces. By using the IPA definitions, you can use the Computer Modern accents with an IPA character in the same way you would accent a Computer Modern character. For example, Figure 4 was created by entering:

```
The superscript tilde is a
nasalization marker for vowels,
thus [\~\scripta] is a
nasalized [\scripta].
```

The superscript tilde is a nasalization marker for vowels, thus [ã] is a nasalized [a].

Figure 4: Accenting IPA characters

To use the IPA accents with Computer Modern characters or with another IPA character, you can define the IPA accent in the same way that the Computer Modern accents are defined. For example, an “over-ring” accent could be defined as,

```
\def\or#1{\edef\next{\the\font}%
\ipatenrm\accent"78\next#1}}
```

and may be used like,

```
The over-ring may be used over
letters with descenders as an
alternative to under-ring to
indicate devoicing, e.g. [\or g].
```

which will print as illustrated in Figure 5.

If the spacing or placement of the diacritic is not exactly what you desire with the basic definition, you can add kerns where needed. Or in some cases, you may want to space differently depending on what character is being accented. Perhaps your scheme for accenting characters with an “undercircle” requires less space between the undercircle and the character, and if the character to be accented is an “r” a completely different scheme would be used. For example, Figure 6 was created by entering:

```
\def\undercirc#1{\ifx#1r
\oalign{#1\cr\hidewidth
\kern.24em\undering
\hidewidth\cr\cr}
\else\oalign{#1\cr\cr
\hidewidth
\raise.1ex\hbox{\undering}
\hidewidth}}\fi}
```

A voiceless trilled r [\undercirc r] in certain Scottish dialects...

The over-ring may be used over letters with descenders as an alternative to under-ring to indicate devoicing, e.g. [g̊].

Figure 5: Using IPA accents

A voiceless trilled r [r̥] in certain Scottish dialects

Figure 6: Using IPA accents

Finally, Figure 7 illustrates using the IPA. The native American Makah script at the top of the figure appears on a plaque at an archeological site in the northwest corner of the United States. The translation for the Makah script is at the bottom of the figure.

Where to go from here?

The fonts are in roman, bold extended and slanted faces. And they come in 8, 9, 10, 11, 12 and 17 point sizes. The project is completed and stable. Figure 8 illustrates the full character set.

The question remains, where do we go from here? Certainly, the WSU IPA is not an exhaustive collection of phonetic IPA characters. Pullam's *Guide* lists 78 "major" characters, all of which are contained in the WSU IPA. The *Guide* also contains many "minor" characters, 50 of which have been included in the IPA.

With Janene no longer at WSU, and no one else willing to carry on the work, no further development on the IPA will be done here. There are more characters that can be added. Anyone willing to do so is welcome to use the existing base characters as examples. The arrival of T_EX 3.0 gives us the ability to assign more than 128 characters is available. It might be reasonable to finish the rest of Pullam and Ladusaw's character set. Also, with the addition of Professor Knuth's virtual font, new ways of mixing the IPA characters into existing fonts is possible. If you choose to modify the WSU IPA, I do ask that you call the new font something other than WSU IPA.

How to get the IPA

You can get the IPA on diskette or by anonymous FTP. Also, the IPA will be distributed to the various site coordinators who can determine whether to place it on their distribution.

To obtain the IPA on diskette from Jon Radel, send him a note and ask for the WSU IPA fonts. Jon's address is:

Jon Radel
P.O. Box 2276
Reston, VA
22090-0276

To get the IPA by anonymous FTP, you can connect to the machine BOBCAT.CSC.WSU.EDU. The address is 134.121.1.1. Log on as ANONYMOUS and change to the TEXT1 directory. You will see ten subdirectories which you can GET files from. The one you want is WSUIPA which contains the METAFONT source, tfm and 300pk files.

Bibliography

- Guenther, Dean, Alan Wittbecker, Janene Winter. "T_EX at WSU" *TUGboat*, 5(1), pages 24-25, 1984.
- Lefebvre, Claire. "Instrumental *Take-Serial* Constructions in Haitian and in Fon." *Canadian Journal of Linguistics* 34:319-337. [Example from p. 330.]
- Piper, Terry. "On the Difference Between L1 and L2 Acquisition in Phonology." *Canadian Journal of Linguistics* 32:245-259. [Example from p. 255.]
- Pullam, Geoffrey K. and William A. Ladusaw. *Phonetic Symbol Guide*. Chicago: The University of Chicago Press, 1986.
- Rhodes, Richard. "Métchif — A Second Look." Pp. 287-296 in *Actes du dix-septième congrès des algonquinistes*. William Cowan, ed. Ottawa: Carleton University. [Examples from p. 293.]
- Thiele, Christina. "T_EX, Linguistics, and Journal Production" *T_EXniques* 5, pages 5-26, 1987.

čabi·yʔakʔu k^wiči·ya· ʔuse·ʔiʔ yeʔitxpi·t
 ʔakyi·q q^widiččapa·tḅ yaqa·qey ʔadicaʔo·wis
 ʔicu·x^wadi· ʔiyaḅ q^wiʔi·q^wiʔḅ ʔuwadiʔ tuk^witap
 hubaqʔ k^wiči·ye·ʔiq ʔu·du·ʔaʔ taččiʔsiʔqa·
 tu·k^wi·ʔ q^wisi·ʔaʔ q^waʔaʔiʔqey ʔa·dʔok^wap
 wikaʔ hita·k^wačiʔ huʔačitwi·ʔiq wiki·d
 ʔaʔu· hidawaʔid yaqa·qey ʔu·ksda·qa·ʔ
 huʔačidiq yaʔaḅ batba·ʔas tu·k^wiʔcki·
 du·baqʔiʔ q^wiq^wiçaqaʔqey hu·ʔas q^wabitqey
 q^wiyowis q^wiʔsiʔ tu·k^witap batba·ʔasiq.

A major village of the Makah Indian Nation, the Ozette village was, over the past millennium, periodically covered by mud flows. The rapid burial created conditions that greatly retarded the deterioration of perishable artifacts. These conditions created a time capsule, unique in archaeology, in which the buried houses and their contents have survived in an almost perfect state of preservation.

Figure 7: American Macah script and translation

	'0	'1	'2	'3	'4	'5	'6	'7	
'00x	e	a	α	o	Λ	b	h	ʃ	"0x
'01x	β	β	ɛ	ε	Ɔ	đ	đ	đ	
'02x	d	d	ɟ	ð	D	ə	ɜ	ə	"1x
'03x	ε	ɜ	ɜ	ɜ	g	ɟ	G	ɣ	
'04x	ɣ	ɣ	h	h	h	h	u	i	"2x
'05x	ɪ	ɪ	I	I	J	ɪ	ɪ	ɪ	
'06x	l	ɮ	λ	λ	ŋ	u	u	ɲ	"3x
'07x	ŋ	ŋ	N	⊙	ə	ɔ	ω	ω	
'10x	∞	ɸ	ɸ	φ	ɾ	ɾ	ɾ	ɾ	"4x
'11x	ɹ	ɹ	R	ɸ	ɶ	ɶ	ɶ	σ	
'12x	t	t	t	θ	ɸ	ɸ	u	u	"5x
'13x	ɸ	u	M	χ	Λ	Y	z	z	
'14x	ʒ	ʒ	ʔ	ʔ	ɔ	ɶ	ɶ	ɶ	"6x
'15x	ɶ	ɶ	ɶ	ɶ	ɶ	ɶ	ɶ	ɶ	
'16x	˙	˙	˙	˙	˙	˙	˙	˙	"7x
'17x	˙	˙	˙	˙	˙	˙	˙	˙	
	"8	"9	"A	"B	"C	"D	"E	"F	

Figure 8: WSU International Phonetic Alphabet 12pt Roman

children had contact spoke a dialect of Irish-English which permits glottal stop in words such as *little* and *bottle*. Moreover, in connected discourse, word final /t/ is frequently glottalized, particularly if the next word begins with a vowel. It was likely, therefore, that the children were following, in varying degrees, the example of their teacher. All instances of glottal substitution which were acceptable in this teacher's dialect were, therefore, dropped from the analysis. Nevertheless, there remained in the data from these nine Ss, persistent if not widespread glottal replacement which may have resulted from various kinds of overgeneralization. One child, for example, in addition to producing glottals consistent with the Irish-English dialect also produced forms such as [tʉmʌʔ] 'too much'; [ʌʉʔsʌɪʔðə] 'outside the'; [wɪʔtʃəz] 'witches'.

Figure 9: Sample combining ph10 and cmr10

- (41) a. Derived LCS resulting from the association of sɔ 'Take' and d'ó 'Put':
LCS: [x cause [y undergo a change of location to z]/PUT
- b. Kòkú sɔ àsón d'ó hàsùn mè
Koku take crab put basket in
'Koku put the crab in the basket.'

Figure 10: Sample combining wsuipa9, wslipa9 and cmr10

	0	1	2	3	4	5	6	7
000	ɤ	d	t	ʎ	ʃ	b	ʒ	ʍ
010	Φ	Ψ	Ω	α	β	γ	δ	ε
020	ζ	η	θ	ι	κ	λ	μ	ν
030	ξ	π	ρ	σ	τ	υ	φ	χ
040	ψ	ω	ε	ϑ	ϖ	ι	ε	'
050	()	*	+	,	-	.	/
060	0	1	2	3	4	5	6	7
070	8	9	:	;	i	=	ʎ	?
100	ð	A	B	C	D	E	F	G
110	H	I	J	K	L	M	N	O
120	P	Q	R	S	T	U	V	W
130	x	Y	Z	[“]	^	·
140	i	e	ɲ	ɔ	d	ə	r	ŋ
150	h	i	j	ʰ	l	ŋ	ŋ	∞
160	ɒ	B	ɾ	ʂ	t	ʉ	ʌ	ʍ
170	ħ	ɥ	ɹ	ʃ	z	z	˘	˙

Figure 11: The ph10 font

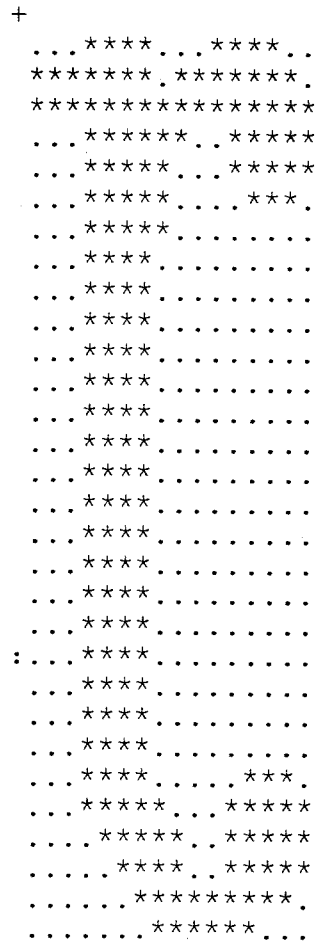


Figure 12: The pixels for the "r with right tail"