

Picked herring and strawberry ice cream

DESIGNING POLISH DIACRITICS

NEARLY ALL the languages of Western Europe are written using the Latin script: Germanic languages (English, German, Swedish, Dutch etc.), Romance languages (French, Italian, Spanish or Portuguese) and other languages such as Finnish (a Finno-Ugric language) or Welsh (a Celtic language). In Central and Eastern Europe, languages spoken primarily by members of Roman Catholic or Protestant churches use Latin script: this includes all Western Slavic languages (e.g. Polish, Czech, Slovak), some Southern Slavic languages (Slovenian, the Croatian version of Serbo-Croatian), Baltic languages (Latvian, Lithuanian) and Finno-Ugric languages (Hungarian, Estonian). Latin script is also used to write other languages spoken all around the world (e.g. Albanian, Aleut, Fijian, Greenlandic, Malay, Maltese, Navajo, Ndebele, Quechua, Samoan, Swahili, Vietnamese, Wolof, and many others).

In the early Middle Ages, people of Western and Central Europe spoke many languages, but most of the writing was still done in the Latin language, the *lingua franca* of those times. If a geographic or personal name from the region of today's France, Germany, or Poland had to be written down, a scribe tried to express sounds not present in the Latin language by approximating them somehow using the Latin letters. These 'rules of spelling' varied heavily from case to case. Gradually from approximately the 9th century on, and more rapidly during the high and late Middle Ages, the number of texts written in local languages increased. Finally, the Reformation movement and Gutenberg's invention of casting movable type boosted the market for publications written in Europe's local and national languages.

The invention of movable type emphasized the need for standardized and simplified character sets and spelling rules: the printers wished to minimize the set of characters they used, the publications were reaching a much larger audience, and the first kinds of publishing processes were introduced giving birth to professions such as typesetter and proofreader. Authors and printers tried to find efficient ways to write sounds not found in Latin. One approach was to use combinations of two or more letters (digraphs, trigraphs). Another method was to employ 'accented'

letters, equipped with diacritical marks – small signs (dots, curves, circles, dashes) attached to or placed above or below the base letterform, indicating a modified phonetic value (from the Greek word *diakritikós* meaning ‘distinctive’). Yet another technique was to use extra letters or letterlike characters. In a Europe split by national and religious borders, each language developed its own way of using Latin letters and diacritics. The sound of a palatalized *s*, as in the word ‘shop’, can be written as *sh* in English, as *sch*, *s*, *g* and *ch* in German (depending on dialects and context), as *ch* in French, as *sz* and *ś* in Polish, and as *š* in Czech. A set of more than 150 alphabetic characters is needed to write all Latin-scripted languages of modern Europe.

READ THE RECIPE FIRST

Type designers from Western Europe and North America, and more recently from other countries, have been refining and improving the style and appearance of Latin letterforms for centuries. With several thousand typefaces designed to date, the Latin letters have been used to set text in a multitude of languages, and have been on display and under the critique of designers and readers from many countries throughout many years. The Latin letter has become a truly global medium, serving cultures from around the world. On the other hand, the evolution of diacritics usually occurred in isolation. From language to language and from country to country, diacritical marks have gained different shapes and different positions.

Before the invention of laser printers, personal computers, and digital fonts, typefaces were developed for particular markets and printing equipment. Lead type was cut by punchcutters familiar with the specifics of a given language. In England, Germany, France, or Italy, printers bought type from local foundries. Polish printing houses imported type from Italy or Germany, but then resident Polish punchcutters localized the fonts, i.e. added the accented letters. There was no global typography. Even in the 20th century, typefaces were made for particular typesetting and phototypesetting systems. As a result of the political division of Europe after World War II, most typefaces used in the Slavic-speaking countries were developed or at least localized there.

Desktop publishing systems brought substantial changes to the font business. Large foundries ceased to exist or completely changed their profiles. Small companies and individuals started designing and releasing digital fonts. With Fontographer, everybody was able to try his or her hand at designing type.

The ‘digital revolution’ started in the USA and in Western Europe, so most digital fonts were designed there, but designers proficient in Latin

letterforms lacked the know-how to draw foreign diacritics properly. Specimens showing well-designed Slavic diacritics, written in ‘unintelligible’ languages like Czech or Polish were hardly available in the West. But market pressures forced foundries to release fonts quickly.

Consequently, the quality of diacritics is poor (with a few notable exceptions). French or German accents are usually acceptable, but Slavic diacritics are mostly mediocre. The design of those characters often does not match the overall typeface style. The accents are misplaced, their color and stress do not go well with the letters. Sometimes, accents are simply borrowed from a different typeface.

FIVE HUNDRED WAYS TO SERVE CHICKEN

The Unicode Standard favors unification of diacritical marks, which means that each font can only have one glyph of a given kind (unless advanced techniques like language-dependant OpenType layout features are used). However, typographic traditions in shaping or positioning accented letters vary among countries or languages. Instead of a separate *umlaut* sign for German (*mögen*, to like) and a *tréma* sign for Spanish (*güisqui*, whisky) or French (*Noël*, Christmas), the Unicode Standard encodes just one *diaeresis* mark (U+00A8), which is to be used for both *umlaut* and *tréma*.

The problem is most evident in display typography. In Germany, it is permitted to place the *umlaut* inside the capital U or the O, but Spanish or French readers may find it unacceptable. Similarly, the preferred angle, size, and placement of the acute mark (´, U+00B4) may vary from country to country, depending on national traditions. The Polish language uses an *acute*-like mark (´, Polish *kreska*) and a *dot* mark (˙, Polish *kropka*), while French uses *acute* (´, French *accent aigu*), *grave* (`), *circumflex* (^), and occasionally *diaeresis* (¨, French *tréma*). To visually differentiate *acute* from *grave* in French, the accents need to have a pronounced slant to the left or to the right, but do not necessarily have to be very tall. In Polish, however, the *kreska* only needs a little slant to the right, but should be tall enough not to be confused with *kropka*.

The Czech language also makes very extensive use of the *acute* mark (Czech *čárka*). When text is set in all capitals and there is little space above A or O, the accents need to be ‘compressed’. In Czech, the accent typically becomes very flat, nearly horizontal, and touches the letter or remains above it. In Polish, the accent gets reduced in size but remains rather vertical, with little slant to the right. It stays above, touches the letter, or even pierces the top of the letter (figure 1).

It is important to understand that diacritical marks are not an extra ‘gadget’ in written text. They are an essential element of the written lan-



Figure 1: In display typography, accents may need to be ‘compressed’ when text is set in all capitals. This can be done differently depending on language. Left: Czech lettering by Jaroslav Šváb (1967), right: Polish lettering by Witold Chmielewski (1957).

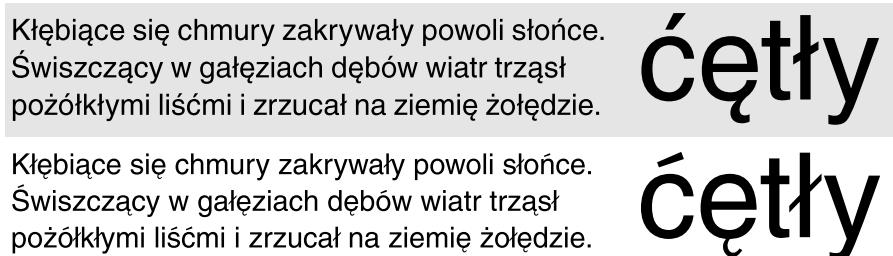


Figure 2: Diacritics need to harmonize with the design of the typeface (top). Poorly designed diacritics may ruin the overall appearance of the text (bottom).

guage; they need to harmonize with the overall design of the typeface just as much as the figures or punctuation marks do. The diacritical marks need to match the design of the basic letterforms in color, shape, stress and contrast. An accented letter must be designed in such a way that it is clearly distinguishable from another accented letter and from the basic letter. Poorly designed diacritics may ruin the overall appearance of a typeface and may seriously impair the legibility of the text. Figure 2 shows well designed and poorly designed Polish diacritics for Helvetica. Please note the bad spacing of the letters *łkły* in the word *poźółkłtymi* in the bottom example. The reader is likely to ignore thin and poorly drawn *ogoneks*, thus misreading the text.

COOKED IS BETTER THAN RARE

The oldest Polish sentence recorded in writing can be found in the Latin text *Księga Henrykowska* (Henry’s Book, 1270): *day ut ia pobrusa a ti poziwai* (‘giveth to me, I shall stir and you shall rest’). In today’s spelling, the sentence would be: *daj, ac ja pobruszę, a ty poczywaj*. We see that written language of the 13th century did not make use of diacritics. Over a century later, the essay *Orthographia Bohemica*, published 1406 or 1412 and attributed to the most important 15th-century Czech religious

reformer, Jan Hus, presented a modern, consistent phonetic method of writing the Czech language. The author suggested denoting long vowels with an acute mark and palatalized consonants with a dot, which later evolved into a *caron* (ˇ, Cz. háček) above the letter. Hus’s system greatly influenced other European languages by being the very first to constitute the principle of using diacritics to write distinctive sounds.

The Polish language of the 15th century was written very inconsistently. The scribes freely mixed phonetic rules of French, Czech, and German to express Polish sounds in Roman letters. Shortly before Gutenberg’s invention, in 1440, Jakub Parkosz published a set of spelling rules for Polish, but this had no lasting effect on writing practice in Poland. Some decades later (1512–14), Stanisław Zaborowski published two texts on Polish spelling. He adopted and extended Hus’s phonetic system, utilizing one dot above the letter to denote modified consonants (ć, ś, ń, ź) and contracted vowels (ą, ó), two dots above to denote palatalized consonants, and a *virgula* (a diagonal stroke, *kreska*) to denote nasal vowels and the *ew* sound (ł). 1511. Stanisław Murzynowski modified Zaborowski’s spelling, replacing the two dots with a *virgula* (which later evolved into an *acute*), and used digraphs like *cz, sz, rz* instead of the letters ć, ś, ń.

Throughout the first half of the 16th century, Cracow-based printers published numerous Polish texts in blackletter typefaces. Instead of cutting and casting new letters to follow Zaborowski’s and Murzynowski’s recommendations, the printers used Latin syncopes (contractions) that looked similar to the proposed diacritics. For example, rather than denoting nasalized *e* (ę) with *e virgula*, the printers used *e caudata*: an *e* with a long tail (Pol. *ogonek*), which once represented the syncope of the Latin diphthong *ae* (figure 3).



Figure 3: The Gros Canon typeface by Claude Garamond (c. 1530) included the *e caudata* character. The character once represented the syncope of the Latin diphthong *ae* but was later used to denote Polish nasalized *e*. Note that Garamond’s diacritical marks are very pronounced.

One of these printers was Jan Januszowski, who 1594 published a book, *Nowy charakter polski*, presenting three spelling proposals devised by the printer himself and by two other authors (both highly educated noblemen). For the purpose of this project, Januszowski designed and cut a new typeface that included all the diacritics needed in the book. In-



Figure 5: Examples of well-designed Polish diacritics.

1. Linotype Ergo (Gary Munch)
2. Bauer Bodoni (Giambattista Bodoni/Heinrich Jost)^a
3. Linotype Really (Gary Munch)
4. Fenway Italic (Matthew Carter)
- 5, 6. Bliss (Jeremy Tankard)
7. Interstate (Tobias Frere-Jones)^a
8. ITC Kabel (Rudolf Koch/Victor Caruso)^a
9. Lunatix (Zuzana Licko)^a
10. Triplex Italic (John Downer)^a
11. Silentium Pro (Jovica Veljović)
12. Linotype Really (Gary Munch)
13. ITC Souvenir (Ed Benguiat)^b
14. Valetta (John Hudson)
15. Bliss (Jeremy Tankard)
16. Waters Titling Pro (Julian Waters)
17. Flux (Monib Mahdavi)^a
18. Rusticana (Adrian Frutiger)^c
19. Excelsior (František Štorm)^a
20. Caffisch Script Pro (Robert Slimbach)
21. Shelley Allegro (Matthew Carter)^c
22. Monotype Script (unknown)^{ac}
23. Fenway (Matthew Carter)
24. Półtawski (Adam Półtawski)
25. Blocka (unknown)
26. Fenice (Aldo Novarese)^b

Diacritics designed by original authors unless indicated:

^aAdam Twardoch, ^bStefan Szczypka, ^cAndrzej Tomaszewski.

can be compressed by reducing its size and, in some cases, crossing the upper part of the letter with the stroke (figure 1, right example). The lower part of *kreska* should be aligned to the vertical axis of the letter, so that the accent appears slightly shifted to the right.

Ogonek is a diacritical mark which causes much trouble to type designers. It is not a floating accent, but rather becomes part of the lower right of the letter. In most cases, it has the form of a sickle or a hook, with one end attached to the body of the letter and the other end pointing to the right. It has a calligraphic nature, following the stroke proportions, the contrast, color, and form of the typeface. In particular, the contrast between thick and thin elements should be consistent (figure 5.2). *Ogonek* should be large enough to be recognized as such. In many cases, it may descend as deeply as the stems of *p* and *q*.

Obviously, *ogonek* does not necessarily have to be rounded. Blackletter types or display typefaces with unusual letterforms may have unusually formed, broken *ogoneks* (figure 5.9–5.10). Depending on how closed the letterforms are, the terminal may end downwards, leaving the inner space open (figure 5.7–5.8), or it may take a turn and end upwards, closing the inner space of the accent (figure 5.2, 5.10). *Ogonek* should not extend beyond the right edge of the base letter. In the vast majority of cases, *ogonek* should be connected with the body of the letter; in some display typefaces, however, it may be treated as a floating accent. In geometric all-capitals display sanserifs, *ogonek* can have the form of a simple diagonal stroke (since it was originally drawn as a *virgula*). To avoid unwanted overlapping effects in digital fonts, all letters with the *ogonek* accent should be outline characters rather than composite characters.

In uppercase *E*, *ogonek* should be attached at the baseline near the right part of the baseline stroke, and should be right-aligned with the upper beak (serif) of the letter.

There are three ways of attaching *ogonek* to uppercase *A*. You can attach it at the baseline to the outer serif of the right diagonal stroke of *A* (figure 5.24). In most cases, this produces poor results because the letter looks like it is flipping to the right (a heavy *ogonek* hangs on a thin serif). You can also attach it at the baseline to the middle of the right diagonal stroke of *A* (figure 5.23). This usually works fine. The third method is to replace the inner serif of the right diagonal stroke of *A* with *ogonek* (figure 5.12–5.14, 5.16, 5.26). If there is no serif, you should attach *ogonek* to the inner part of the right diagonal stroke (figure 5.15, 5.17, 5.25). In my opinion, this method works best, because removing the serif reduces visual noise in that part of the letter.

Drawing lowercase *q* is tricky because the danger of a visual clash is relatively high. Depending on the style of the typeface, *ogonek* may

be attached anywhere at the terminal of the letter *a* (not to the bowl!). Remember to avoid extending *ogonek* beyond the right edge of the base letter (figure 5.11 is just barely acceptable).

Drawing lowercase *ę* is most difficult because the bottom of *e* is round rather than flat. Please observe figure 2: you should always draw the letter *ę* as shown in the top example! *Ogonek* in *ę* should not be centered like a cedilla. It should rather be smoothly connected to the right part of the terminal (approx. at two-thirds to three-quarters of the character width), or to the end of the terminal (figure 5.1–5.11).

Ł is not composed out of the letter *l* and a *slash sign* (/), although this glyph is called *lslash* in PostScript glyph naming conventions. The lowercase letter *ł* is an *l* with a diagonal stroke in the middle. This diagonal stroke should never be too heavy or too large. In most cases, it should be as thick as the thin diagonal stroke of the letter *x*. The terminals of the stroke may be cut off vertically (figure 5.7) or diagonally (figure 5.8), depending on the shape of the terminals of the letter *x*. In most cases, the diagonal stroke should cross the letter *l* at its visual centre (which is usually higher than the geometrical centre). Typically, the right part of the diagonal stroke is slightly longer than the left part, especially if *l* has a bowled terminal or a spur at the ascender. The right terminal of the diagonal stroke should ideally end at the x height or slightly above it, but in some cases, it may end slightly below the x height. The rule of thumb for the angle between the baseline and the diagonal stroke is 30 degree (figure 5.1–5.11). Letter pairs *łw*, *ły*, *łt* should be kerned carefully.

The uppercase *Ł* is an *L* with a diagonal stroke. The middle of the diagonal stroke should be aligned at the height of the horizontal bar of capital *E* or slightly above it. The right part of the diagonal stroke is substantially longer than the left part, at least as long as the horizontal bar of *E*. The diagonal stroke should not be thicker than the thin diagonal stroke of the letter *X* (figure 5.12–5.18, 5.23–5.26).

In calligraphic styles, however, *ł* has a fairly long straight or curvy stroke, slightly ascending or nearly horizontal, and is placed just at the top or slightly above the top of the letter (figure 5.19–5.22). This script form of the letter *ł* is used in everyday Polish handwriting and should be used in script typefaces. The stroke in capital *Ł* is usually curvy or nearly horizontal and crosses the letter in the middle (see 5.21). In script faces, the double-letter sequence *łł* may cause a visual clash, so a ligature may be required (figure 5.22).

DINNER IS SERVED

Poorly designed diacritics ruin the appearance of a typeface and impair the legibility of text. Diacritic characters are an essential element of typographic communication, no less important than the basic letterforms. Drawing good accented letters is more time-consuming than building them from prefabricated composites, but in the long run, only a well designed typeface has a chance to survive on the market and to become a popular choice with users. My list of recipes is intended to help a type designer avoid most serious mistakes. After you have mastered the rules, you may start breaking them. Otherwise, you may end up serving pickled herring with strawberry ice cream to your readers.

Acknowledgements

Thanks to Zbyszek Czapnik, John Hudson, Ross Mills, Stefan Szczypka, Andrzej Tomaszewski, and Maxim Zhukov. Further information regarding Polish diacritics can be found at <http://www.twardoch.com/>

LANGUAGE
CULTURE
TYPE

INTERNATIONAL TYPE DESIGN
IN THE AGE OF UNICODE

Edited by John D. Berry

Graphis • ATypI