

Kedves Gábor! Több okból kifolyólag sokkal régebben ismerhetnénk egymást, mint ahogy: mindketten voltunk piarista diákok, aztán ott van még Balatonszemes, meg az is, hogy legkésőbb a huszadik század vége óta a lexicográfia egészen biztosan nem értelmezhető a számítógépes nyelvészet nélkül...

Mindettől függetlenül úgy alakult, hogy találkozni csak valamikor 2015 tavaszán találkoztunk – ám levelet „tőled” ennél sokkal korábban kaptam először. Emlékeim szerint ennek a történetét még nem meséltem el neked – ha mégis, ennek az írásnak annyi hozadéka mindenképp van, hogy most és itt a sztorit igyekeztem szépen illusztrálni is.

Van, ugye, az a hosszú múlttal bíró, nagy magyar lexicográfiai projekt, amelynek vonatkozásában az elmúlt évtizedekben így-úgy mindketten tevékenykedtünk – de valahogy különböző időszakokban, nem átfedésben.

Jómagam 2006-ban szembesültem először az említett szótári munkálathoz tartozó szövegtárral: egészen pontosan azzal, hogy abban miféle karakterkódolási megoldások is találhatóak. Valami ilyesmi tárult a szemem elé, amikor először találkoztam az általad még az 1980-as években kiötlött – és Kornai András által azon frissiben „Prószéky-kód”-nak elnevezett¹ – megoldás gyakorlati alkalmazásával.

Sparta bels43ö0-része s43ints43 illy veszély nélkü0l,

Mellyet Leónidás, gerjeszteni készü0l.

Ams43ares43t meg-vette tudod már pénzével,

Tö0b tanáts43okat is veszteget kintsével.

Nem akarja, hogy a' végezés meg-álljon,

'S tö0rvényeinkre titkon s43íve hóltig fájjon

'S mos43olyogva tes43zi els43ö4 mozdulását.

A' néző4k s43zemei formáján el-ves43znek,

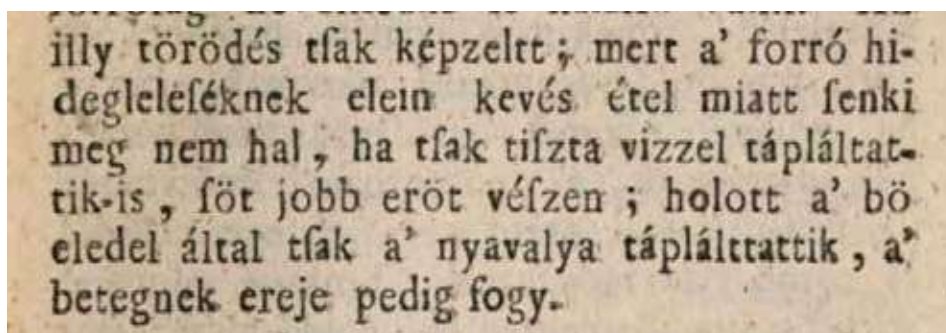
Egy pis43zs43zenés43t s43ehol miatta nem tes43znek

Ez volt az a látvány, amelynek a következtében számomra a Prószéky-kódok hosszú évekig egyet jelentettek a betűhűség egy sajátos értelmezésével. Amit itt láthatunk, az az, hogy ebbe a korpuszba – a Jóisten legnagyobb dicsőségére – sikerült néhány olyan – legfeljebb nyomdászattörténeti szempontból érdekes – graféma több százezernyi előfordulását fáradságos munkával rögzíteni, amelyeknek a megjelenítése egy paleográfiai projektben biztosan jelentőséggel bírhat – de egy alapvetően értelmező szótári munkálattal során minden bizonnyal nem. Számszerűsítsünk: a korpusz mintegy 390.000 Prószéky-kódjának 97 százalékáért 3 graféma „felel”: egy ilyen ő, egy ilyen ú és egy ilyen í.

¹ Kornai A. (1986). Szótári adatbázis az akadémiai nagyszámítógépen. *Műhelymunkák II.* (p. 67).

A hosszú s 270.000 előfordulása mellett rögzítésre került 79.000 ó meg 24.000 ú is. Ezen három betű mindegyikének fontos tulajdonsága, hogy semmiféle fonológiai relevanciával nem bírt a vele párhuzamosan ugyanekkor már létező s-hez, ő-höz vagy ú-höz viszonyítva: azaz egy-egy olyan hangot reprezentáltak, amelynek a jelölésére nemhogy a mai magyar ábécében megvan a megfelelő betű, de már 200 éve is megvolt.

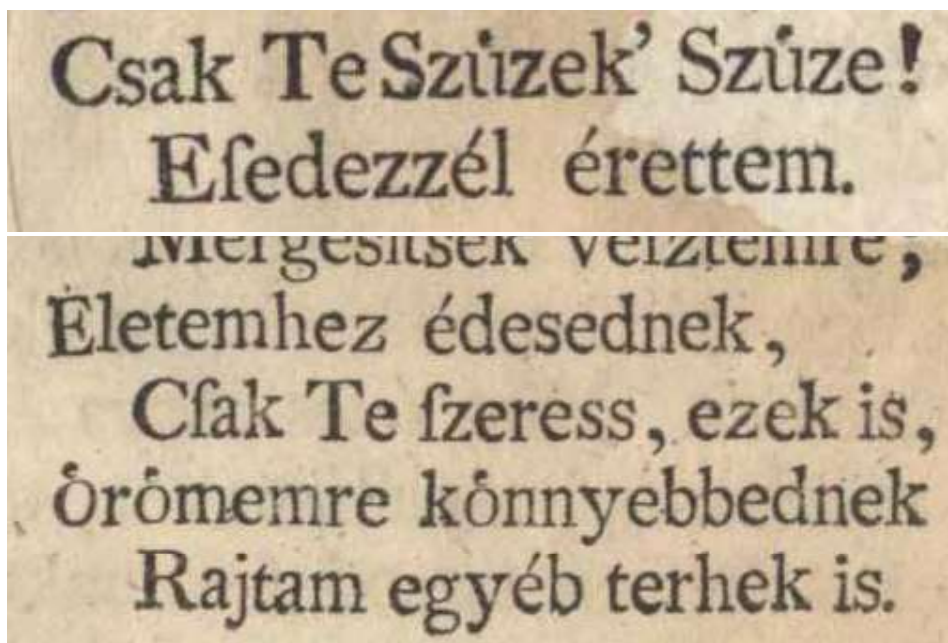
Az alábbi képanyag ékes bizonyítékul szolgál arra, hogy ezeknek a betűknek a 18. századi nyomtatványok lapjain való feltűnése vagy hiánya esetleg tipográfiai „divatirányzat”-okkal lehetett összefüggésben – vagy egyszerűen azzal, hogy az adott nyomdának milyen ólombetűből mekkora készlet állt éppen rendelkezésére a szedőszekrényben. Az eredeti nyomtatványok oldalképeinek tanulmányozása során egyértelművé válik: a hosszú s-t és társait nemhogy az egy időben keletkezett különböző kiadványokban nem szedték következetesen, de a használatuk még egy kötetben belül is ingadozhatott.



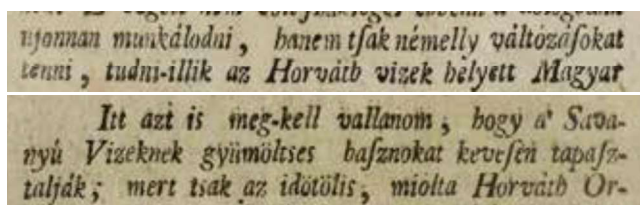
1. kép. Marikovszky Márton *A néphez való tudósítását* 1772-ben, Nagykárolyban nyomtatták: a *tsak* nem fordul elő máshogy, mint hosszú s-sel

ritésből tett fogadásokat az Egek tsak nevetik, és meg-nem büntetik: Addig várakoztatta, és reménytette tehát haszontalanul a' fel, gyúlladott Hermenét, hogy emez végtére szerehmes nyughatatlanságától kényszerítettett utánna tudakoztatni, a' ki a' kimlöi által tsak hamar ki-tanúlta, hogy a' Hertzeg egészségének leg-kevesebb hibája se volna, és tsak szánt-szándékos hallasztással támogatná-el

2. kép. Mészáros Ignác *Kártigámját* szintén 1772-ben nyomtatták, de Pozsonyban: szabad szemmel is jól láthatóan rendelkezésükre állt néminemű hosszú s, a *tsak* mégis mindig így lett szedve



3. kép. Cserey Farkas *Isten anyjának dicsérete* című művét is 1772-ben nyomtatták, ám azt Bécsben: a 135. oldalon *Csak*, a 139.-en *Csak*...



4. kép. A magyarországi orvosvizekről és a betegségekben azokkal való élésnek szabott módjairól című, 1783-ban, Nagykárolyban nyomtatott mű előszavában szintén van *tfak* és *tsak* is...

Már önmagában ez a tény erősen meggyengíti a rögzítésük létjogosultságát! Ám ha ehhez még hozzávesszük a korpuszhoz rendelt lekérdezőszövegnek azt a sajátosságát, hogy ha a keresőkifejezés nem betű- és számkarakterrel

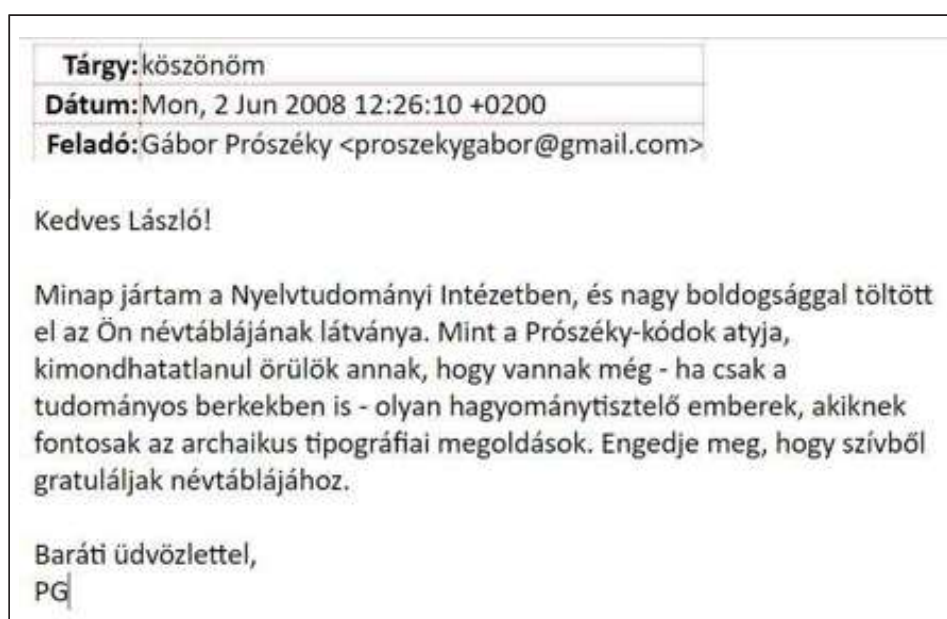
terről betű- és számkarakterre pontosan egyezett meg a korpusz-előfordulás sztringjével, akkor az adat jó eséllyel nem került fel a találati listára, akkor talán érthetővé válik, milyen léptékű problémákhoz vezetett a betűhűség. Aki ebből a korpuszból ki szeretné nyerni például a *dicsőséges* szó összes adatát, az jól teszi, ha nem feledkezik meg a *dits43ös43s43éges* karakterláncról sem...

Reményeim szerint a fentiekből is kitűnik, hogy egy pillanatig sem a kódrendszert kárhoztattam. Sokkal inkább az olyan grafémákhoz való megszállott ragaszkodást, amelyek rögzítése (1) a projekt fő csapásirányát, a nyomtatott forrásokból vett idézetekkel való illusztrálást tekintve érdektelen volt; (2) feleslegesen túlbonyolította a lekérdezőszöveget.

Idővel az s43 lett a felesleges zavart okozó Prószéky-kód szinonimája: feltehetőleg egyszerűen az előfordulási gyakorisága miatt. Annak következtében, hogy heti szinten adtam hangot azon az álláspontomnak, miszerint „*Ceterum censeo...*”, azaz ezek az s43-ak

a korpuszból feltétlenül eltávolítandóak (értve ezalatt, hogy egy kétperces művelettel s-sé egyszerűsítendőek), a kollégáim a felesleges grafémákkal való bajlódás iránti ellen-szenvenet úgy interpretálták, hogy nekem feltehetőleg már eredendően fenntartása-im voltak ezekkel a szám-betű kombinációkkal, sőt talán már a feltalálójukkal szemben is ellenérzések alakultak ki bennem...

2008 egy kora nyári napján aztán az alábbi levél landolt a nyelvtudományi intézeti postafiókomban.



Ismételten csak az emlékeimre hagyatkozom, de szerintem egyetlen percig sem feltételeztem, hogy ezt a levelet valóban te küldted volna, azaz hogy PG az „igazi” Prószyk Gábor lenne. Vicces kedvű kollégámat, Gyórfy András barátomat gyanítva a háttérben, kiballagtam az irodám ajtaja elé, sejtve, hogy a tréfa nem merült ki egy proszekygabor@gmail.com-os fiók regisztrálásában. Nem kellett csalódnom... Innen nézve azt azért sajnálom, hogy arról, amit az ajtón találtam, nem maradt fenn eredeti dokumentum: de hát akkortájt a mobiltelefonok többségének még nem volt kamerája. (Hogy Gyórfy András honnan vehette, hogy az e-mailjeidet „PG”-ként írod alá, arról semmilyen elképzelése nincs sem neki, sem nekem.)

Kedves Gábor, Isten éltesen sokáig!
A hetven az új negyven (vagy lehet, hogy harminc: már nem is tudom).
Barátsággal: Laci



5. kép. A Nyelvtudományi Intézet 109-es szobájának ajtaja, 2008. június 2. (rekonstrukció: Gyórfy András)