

DIPLOMAMUNKA

Hasulyó László

Debrecen

2010

Debreceni Egyetem

Informatika Kar

Gyorsírás a billentyűzeten

Témavezető:

Aszalós László

PhD

Készítette:

Hasulyó László

PTI MSc

Debrecen

Tartalomjegyzék

TARTALOMJEGYZÉK	3
1 BEVEZETÉS	4
2 SZÓRÖVIDÍTŐ RENDSZEREK	4
2.1 SZÁMÍTÓGÉP ELŐTTI MEGOLDÁSOK (KÉZÍRÁS, ÍRÓGÉP)	4
2.2 ALTERNATÍV GYORSÍTÁSI LEHETŐSÉGEK	5
2.3 ANGOL NYELVŰ RENDSZEREK	7
2.3.1 <i>Easyscript</i>	7
2.3.2 <i>Abcz</i>	10
2.3.3 <i>Szövegszerkesztők automatikus javítás</i>	12
2.3.4 <i>Visual Studio kezelése</i>	13
2.3.5 <i>Gyorsíró billentyűzetek: velotype veyboard</i>	15
2.4 MAGYAR KUTATÁSOK	15
2.4.1 <i>Magyar nyelvészeti morfológiai kutatások. Alkalmazott nyelvészeti kutatások magyar nyelven</i>	15
2.4.2 <i>Magyar változatú gyorsíró billentyű</i>	16
2.5 A MAGYAR MEGVALÓSÍTÁS NEHÉZSÉGEI, MAGYAR NYELVBELI RÖVIDÍTÉSI NEHÉZSÉGEK	17
3 TERVEZÉS	19
3.1 MEGFONTOLÁSOK	19
3.1.1 <i>„online-offline”</i>	19
3.1.2 <i>Magyar szavak szerkezete</i>	20
3.1.3 <i>Myspell Hunspell Ispell</i>	24
3.1.4 <i>A dict fájl szerkezete</i>	25
3.1.5 <i>Aff fájlok szerkezete</i>	26
3.2 MIT LEHETNE RÖVIDÍTENI	30
3.2.1 <i>Előtag rövidítésére</i>	31
3.2.2 <i>Szótő rövidítése</i>	32
3.2.3 <i>Toldalékok rövidítése, rövidítési lehetőségei</i>	33
3.3 MÓDSZEREM HATÉKONYSÁGA	35
4 IMPLEMENTÁCIÓS TERV	36
4.1 A PROGRAM JELLEGE	36
4.2 SPECIFIKÁCIÓ	36
4.3 JÖVŐBELI FEJLESZTÉSI LEHETŐSÉGEK	44
5 ÖSSZEFOGLALÁS	45
6 IRODALOMJEGYZÉK	45

1 Bevezetés

A diplomamunkámban szeretnék bemutatni egy olyan rendszert, amivel a szavak beírási sebessége növelhető, ha nem kell begépelni a szó összes betűjét. Ez egy szórövidítő rendszer, ami a magyar nyelvre, szavakra van specializálódva. Angol nyelven már létező megoldások előnyeinek és hibáik alapján szeretnék egy hasonló magyar nyelvű rendszert megalkotni, illetve létrehozni egy ilyen rendszert.. Dolgozatomban a program tervet ismertetem, az implementációt későbbre halasztom. A magyar nyelv toldalékoló szerkezete miatt át kell értékelní a meglévő megoldásokat, és magyar nyelv specifikussá tenni. Ezért felvázolom a magyar szavak szerkezetét és azokból adódó nehézségeket. Felvázolok egy rövidítési megoldást, és megtervezek egy szövegszerkesztőt, amely kezelni tudja a rövidítéseket.

2 Szórövidítő rendszerek

Saját módszerem ismertetése előtt bemutatom a angol nyelvű módszereket, mert ezek előnyös vonásait érdemes átvenni, és lehetőség szerint ezek hátrányait ki kellene küszöbölni.

2.1 Számítógép előtti megoldások (kézírás, írógép)

A kézírásos gyorsírás azért alakult ki, mert az írás lassú volt a folytonos beszéd egyidejű rögzítéséhez; és ennek a kívánalomnak meg is felel. Viszont hátrányai is vannak:

- Elsajátításához évek szükségesek mintha újra meg kellene tanulni írni.
- Át kell írni utána természetes nyelvre, hogy mások is értelmezni tudják.
- Sőt ha más gyorsírni tudó ember akarja visszafordítani, akkor lehet, hogy nehézségekbe fog ütközni a másik ember egyéni szórövidítései miatt.

A később ismertetésre kerülő EasyScript angol rövidítésrendszert lehet használni papíron is, nem alkalmaz speciális írásjeleket a megszokott betűik helyett. (Létezik több hasonló rendszer is. (Stenoscrypt ABC Shorthand))

Az írógépek megjelenésével új igény és szakma jött létre a gépírás, amit ugyancsak tanulni kell hónapokon, éveken át, hogy nagyobb sebességre tegyen szert benne a gépíró. A tízujjas gépeléssel lehet kihasználni a gépírást igazán, de szerencsére itt nincsenek megtanulandó

jelek, mint a gyorsírásnál. Így bárki nekikezddhet gépírni két ujjal, ami a természetes írással enged párhuzamot vonni.

Sajnos történetileg az írógépek megjelenésénél a mechanika nem állt a helyzet magaslatán, így a pergő leütéseket nem sokáig bírták a betűkarok. Ezért a billentyű elhelyezkedést úgy alakították ki, hogy a lehető leghosszabban lehessen gépelni, megakadályozva a karok összeakadását átlagos használatkor. Ez az indok mára teljesen okafogyott lett a számítógépek térnyerése és a technika fejlődése miatt. A megszokás miatt már nem nyúltak hozzá a billentyűkiosztáshoz, noha Dvorak az angol nyelv számára elkészített egy jobb kiosztást, ami angol billentyűkiosztásra már letisztult és elfogadott formában van. A következő fejezetben ezzel foglalkozom.

2.2 Alternatív gyorsítási lehetőségek

Alternatíván azt értem, hogy nem a rövidítéssel oldja meg a gyorsítást, hanem egyszerűen gyorsabban gépelést tesz lehetővé.

Dvorak kiosztás

Mostanában jöttek olyan gondolatok, hogy meg kellene változtatni a kiosztást, ezekhez a technika is megfelelő lett, illetve régebben nem lehetett csak úgy átnevezni a billentyű gombjait a szoftver felkészületlensége miatt. Ezek a billentyűkiosztások nemzetenként, nyelvenként is változnak, ugyanis a használt szavak szerint is érdemes optimalizálni a betűk elhelyezkedését, ha már optimalizálunk.

Az angol kiosztású *dvorak* sebességnövekedésre példa: Egy kísérlet alkalmával a gépelőket tesztelték, addig kellett gyakorolniuk rajta amíg elérték a *qwerty* kiosztáson elérhető átlagos sebességet, ezt sikerült elérni 52 óra alatt. A tanulmány után a *dvorak* kiosztással 74% al gyorsabban gépeltek, és a pontossága a leütéseknek 68%al növekedett. (The Dvorak keyboard, 2003)

Tekintsük, hogy a szöveg betűi a billentyűzet mely sorában találhatóak:

Sor	QWERTY	Dvorak
Felső	52%	22%
középső	32%	70%
Alsó	16%	8%

Ami azt jelenti, hogy a középső részre – ami a quertz kiosztásnál az asdfghjklé billentyűket jelöli – tevődik az átlagos billentyűlenyomások 70 százaléka a Dvorak kiosztás esetén. Ez azért fontos, mert a középső billentyűket lehet lenyomni a leggyorsabban. Megfigyelhető benne az is, hogy szinte sose kell egy ujjal leütöni két egymás utáni különböző betűt. Ez az elgépéléseket is csökkenti.

Magyar „Dvorak”

Létezik magyar alternatíva a hagyományos kiosztásra: Szeghalmy Szilvia megoldására gondolok.

~ `	!	@	#	\$	%	^	&	*	()	{	}	← Backspace
Tab ↔	" ,	< ,	> .	P	Y	F	G	C	R	L	? /	+ =	 \ _
Caps Lock ↑	A	O	E	U	I	D	H	T	N	S	- _	Enter ↵	
Shift ↑	:	Q	J	K	X	B	M	W	V	Z	Shift ↑		
Ctrl	Win Key	Alt							Alt Gr	Win Key	Menu	Ctrl	

1. ábra. Az angol billentyűkiosztás

0	1	2	3	4	5	6	7	8	9	I	Q	C	← Backspace
Tab ↔	F	B	K	G	H	O	Ü	,	É	Ú	X	Ő	←
Caps Lock ↑	T	N	S	Z	L	A	E	U	.	Ű	Y	Í	
Shift ↑	M	D	J	V	P	R	Á	Ó	-	Ö	W	Shift ↑	
Ctrl	Alt							Alt Gr	Ctrl				

2. ábra. A magyar javaslat Szeghalmy Szilvia alapján (Szeghalmy Szilvia 2007)

A helyes az lenne, ha a gépirással ismerkedőknek már az első billentyűzet is ilyen lenne. Sajnos, aki már tud írni (tíz ujjal), annak újra kellene tanulnia, amit sokan nem vállalkoznának erre. Ezzel nem foglalkozom, mert ezt a billentyűkiosztás gyorsítaná ugyan,

de nem előfeltétele a gyorsírásnak. Így én ezzel a lehetőséggel nem foglalkozom, de kiegészítésként sokat segíthet benne.

2.3 Angol nyelvű rendszerek

Többféle rövidítési rendszer van, amiből csak párat ismertetek ebben a részben, de érdemes tudni, milyen rendszerek is vannak.

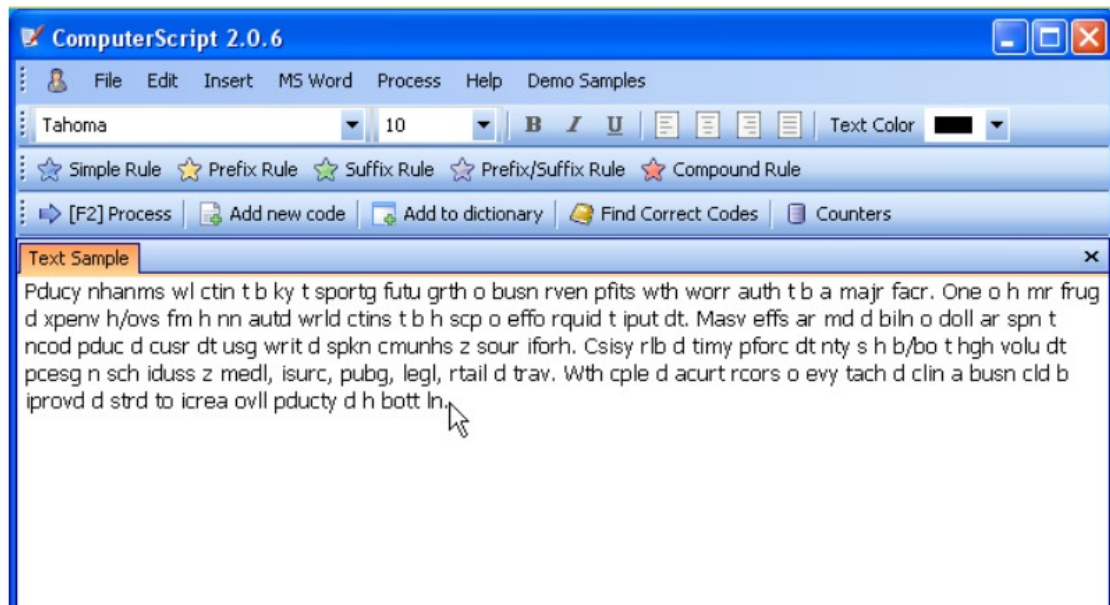
Felsorolva néhányat:

Speedwriting,	http://en.wikipedia.org/wiki/Speedwriting
ABC Stenoscrypt	(Stenoscrypt ABC Shorthand)
Easyscript	http://www.easyscript.com
Forkner shorthand,	http://en.wikipedia.org/wiki/Forkner_shorthand
AlphaHand	http://openlibrary.org/books/OL11506955M/Alpha_Hand_ABC_Shorthand
Baine's Typed Shorthand	http://www.spellingsociety.org/news/pyoneers/py2pt4.php
SuperWrite	http://en.wikipedia.org/wiki/SuperWrite
HySpeed Longhand	http://openlibrary.org/books/OL6134280M/Hy-Speed_Longhand_dictionary
Abbreviatrix	http://openlibrary.org/books/OL11438095M/Abbreviatrix_Shorthand_in_Four_Days
Handywrite	http://www.alyson.org/handy/handywrite.htm

2.3.1 Easyscript

Az EasyScript egy rendszer amit lehet alkalmazni az írásban (akár papíron is), és implementálták is: A ComputerScript program a rövidített alakú szöveget „kibontja” értelmes

szöveggé. Működése fordítónak tekinthető, ami a nem egyértelmű szavaknál felajánlja a lehetséges szavakat, amelyek abból a rövidítésből keletkezhetnek. (EasyScript Express 2001)



3.ábra. ComputerScript kezelőfelülete.

Stratégiája, hogy 5 alap kategóriába rendszerezi a szavakat, úgymint egyszerű, előtag, végtag, elő- és utótag, valamint összetett. 1,2 vagy 3 karakterrel jelöl egy egyszerű szót, és csak maximum két-három szó van hozzárendelve egy rövidítéshez, hogy csökkentse a tanulási mennyiséget.

Hogy egyszerűbb legyen a megértés, egy példával írom le.

Az *unsuccessful* szót szeretnénk rövidíteni. Ez a szó tartalmaz egy prefix/suffixet (elő és utótag) és egy simpla szót. A rendszer több lehetőséget ad a kezünkbe.

Copyright → co/ri , clubhouse → cl/ho

Az következő ábrán az egyszerűbb 3 karakter kezdetű rövidítéssel van lekódolt szöveg látható.

SOURCE TEXT:

It is estimated that up to 70% of small and medium size corporations in the United States are not in compliance with their State or Federal Requirements. They do not document their corporation meeting minutes, have them recorded, signed, and installed in their corporation book of record. One of the main reasons is that their meeting minutes incomplete and improperly recorded due to the lack of appropriate systems and tools to accurately record spoken information. Introducing EasyScript/ ComputerScript to write and type faster, easier, and more efficient than any method you have ever used. This revolutionary way to learn speed writing and typing will enable you to take meeting minutes quickly and accurately. It doesn't take long to learn, just a few hours, and you can become effective taking meeting minutes and more comfortable with this work assignment.

EASYSRIPT BEGINNER – 3 character root

It s estd th up t 70% o sm d me sz corhs n h US an' n cplic wi thr stt r fedr rquims. Ty dn' docm thr corh meeg mins, hv tm rcord, sigd, d istad n thr corh bk o rcor. On o h mn reas s th thr meeg mins icom d iproy rood du t h la o aprot syst d tls t rcor spok iforh. Itrog ES/CS t wt d tp fastr, easr, d mo effn tn ny mh u hv evr usd. Ths rvoly wy t lr sp wrig d tygp wl eb u t tk meeg mins quiy d acury. It dn' tk lg t lr, ju a fw hrs, a u cn bm effv tagk meeg mins d mo cforb wi ti asigm.



[View 2 Character Root \(Advanced\)](#)

5. ábra. 3 karakter szintű rövidítések.

SOURCE TEXT:

It is estimated that up to 70% of small and medium size corporations in the United States are not in compliance with their State or Federal Requirements. They do not document their corporation meeting minutes, have them recorded, signed, and installed in their corporation book of record. One of the main reasons is that their meeting minutes incomplete and improperly recorded due to the lack of appropriate systems and tools to accurately record spoken information. Introducing EasyScript/ ComputerScript to write and type faster, easier, and more efficient than any method you have ever used. This revolutionary way to learn speed writing and typing will enable you to take meeting minutes quickly and accurately. It doesn't take long to learn, just a few hours, and you can become effective taking meeting minutes and more comfortable with this work assignment.

INTERMEDIATE/ADVANCED – 2 character root

It s esd th up t 70% o sm d me sz crh* n h US an' n cpc wi thr stt r fdr rqm. Ty dn' dcm thr crh mtg mn, hv tm red, sgd, d isd n thr crh bk o rcr. On o h mn rs s th thr mtg mn icm d ipy rcd du t h la o apt sy d tls t rcr spk ifh. Itrog ES/CS t wt d tp fsr, esr, d mo efn tn ny mh u hv evr usd. Ths rvy wy t lr s/w d tpg wl eb u t tk mtg mn quy d acy. It dn' tk lg t lr, ju a fw hr, a u cn bm efv tkg mtg mn d mo cfb wi ti asm.

[View 3 Character Root \(Beginner\)](#)



6. ábra. 2 karakter szintű rövidítések.

A hatodik ábrán már a haladó felhasználók által írt szöveget lehet látni, illetve lehetne elolvasni, ha komolyabban foglalkoznánk ezzel a rendszerrel.

Ha megfigyeljük a haladó rövidítés másfél sorral rövidebb lett. De ezért természetesen kell fizetni is, mert a második rövidítés már teljesen olvashatatlan gyakorlatlan szem számára.

2.3.2 Abcz

Az Abcz rendszer szintén egy szórövidítési szisztéma, (Jon Knowles 2010)

. Az alapelv kísértetiesen emlékeztet az előbbire:

- vegyük a szó elején szereplő három betűt, valamint a szó utolsó betűjét. Innen jön a rendszer elnevezése is. Például: *absolutely* → *absy* *radicular* → *radr*
- Ez természetesen nem fedi le az összes angol szót, hiszen lesznek azonos kezdetű és végű szavak. Erre azt javasolja, hogy kezdjük el az abc betűivel jelölni sorra a különböző variációkat. A táblázat első oszlopa a rövidítés típusát jelzi. Például:

abcz	infn ← infection
abczb	infnb ← inflammation
abczc	infnc ← information
abczd	infnd ← infarction

- prefixeket is rövidíti, így *construction* → *cstrn* (a abcz), vagy *discussing* → *dcusg* (a abcz)

Rövidítés	Prefix
a	anti
c	com, con
d	dis
f	for
h	hyper, hypo
i	inter, intr, intra
n	neuro
s	sub, super, sup
t	trans

- Lehet egész mondatokat, mondatrészeket is rövidíteni a kezdőszavukkal: Pl. *thank you very much for the opportunity of allowing me to see* → *tyvmf*
- A végződéseknél, ragozásnál is egyszerű a megoldása. Pl. *Construct cstr* -nél

rövidítve	teljes szöveg
cstrs	constructs
cstrd	constructed
cstrg	constructing
cstrn vagy cstrj vagy cstrz	construction
larr	larger
lart	largest
trey vagy trel	tremendousl y

- Érdekességképen bevezeti a trigger (elsütő) karaktereket is. Ami azt jelenti, hogy a karakter után vagy előtte az általános szavakat lehet kategória szerint betűvel rövidíteni.
 - Család kategória f -el jelölve szó elején: fb ← brother fs ← sister
 - Anatómia kategória k -val jelölve a szó végén (abk) :fok ← forearm brk ← brachioradialis

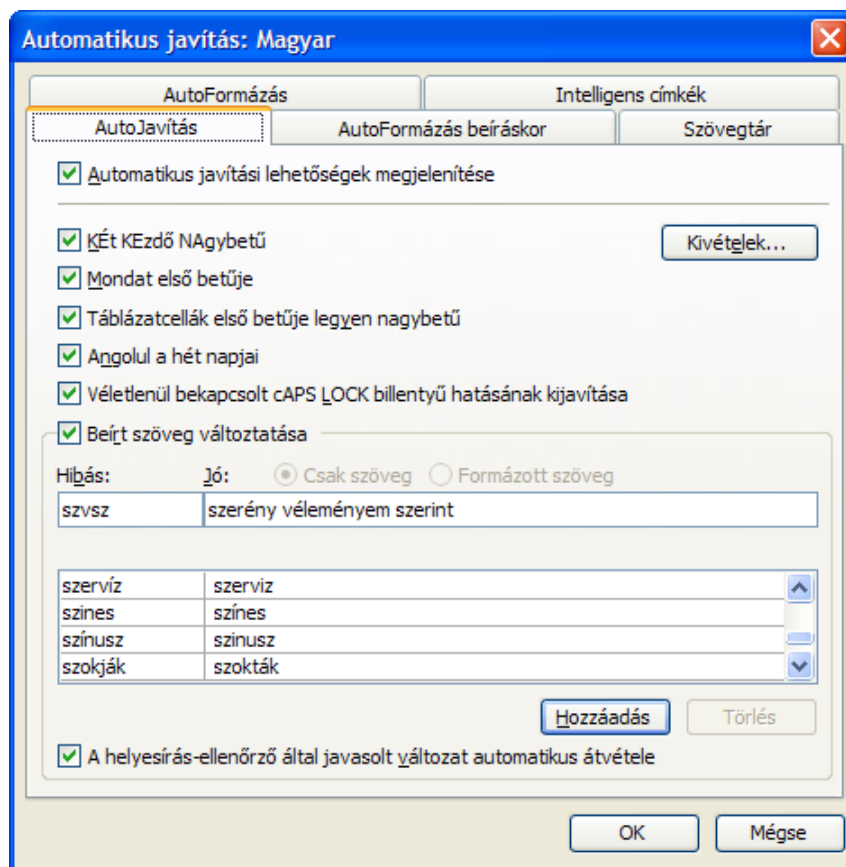
A 7. ábrán látható, hogyan lehet dátumokat rövidíteni. A magyarázat a következő: dátumot a n4y3 kombinációval és szóköz ütésével írja ki (n= november). A kötőjeles verziónál a n4y3 után 'h' rakva kötőjellel elválasztott dátumot generál (h = hyphenated), de ha szóköz helyett F1 -et ütünk utána, akkor már nem kell 'h' betűt utána írni. És ugyanaz 's' és F2-vel is (s=slashed). Ez a példa egy orvosi gyorsíró program rövidítés fájljának részlete.

```
Shorthand: DATES03_98SFWFINAL.SPF
[SPACE] n4y3 = November 4, 2003
[F1] n4y3h = 11-04-03
[F2] n4y3s = 11/04/03
```

7. ábra. Egy abcz típusú rövidítésfájl tartalma.

2.3.3 Szövegszerkesztők automatikus javítás

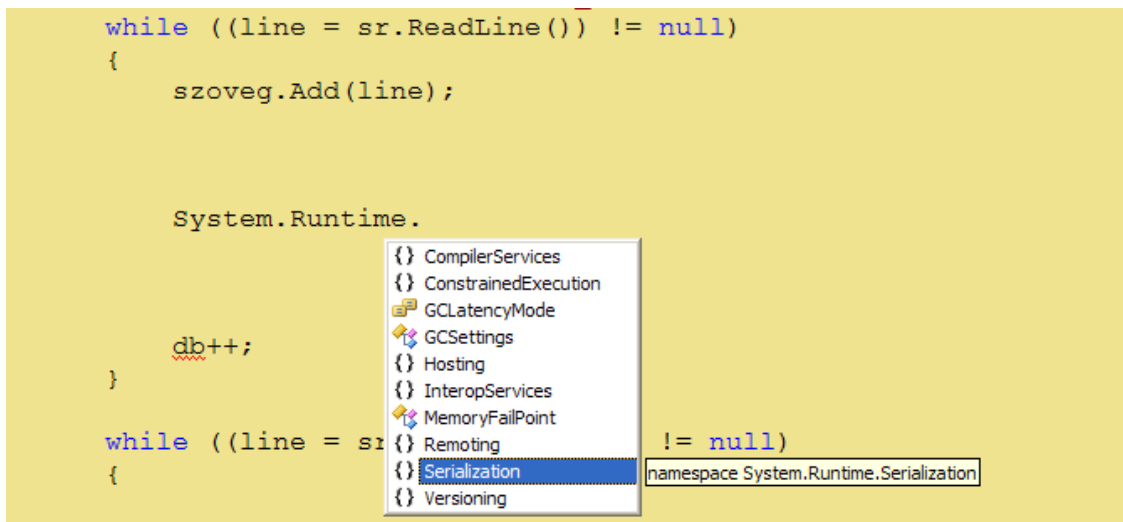
Mindenki találkozott már az automatikus javítással, ha komolyabb szövegszerkesztőt használt. Ide értem Microsoft Word, OpenOffice, sőt egyszerűbb szövegszerkesztők is megtaníthatóak erre a trükkre. Ez arra való, hogy a leggyakoribb *véletlen* gépelési hibákat kijavítsa, azaz a rövidítést a teljes kifejtett alakkal helyettesítse. A programmal kapott listát természetesen lehet bővíteni. Így az gyakran használt szavakra tudunk rövidítéseket létrehozni.



8. ábra Példa MS Word 2003 automatikus javításának szótára rövidítés beszúrása.

2.3.4 Visual Studio kezelése

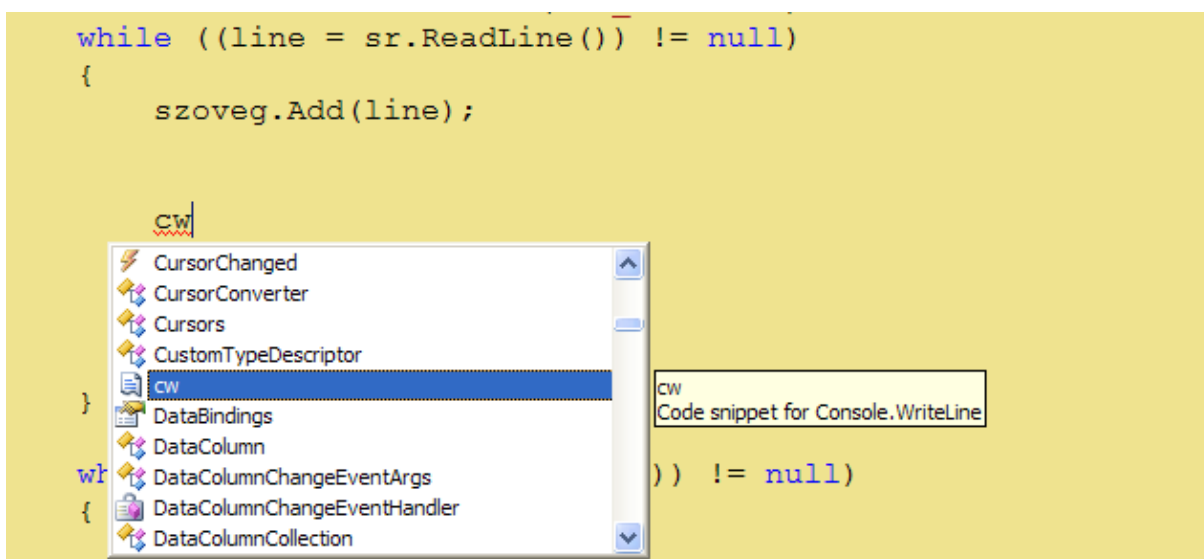
Meg kell említenem a programkódoláshoz segítséget nyújtó fejlesztői környezetek (integrated development environment IDE) működését is. Manapság már nehéz boldogulni nélkülük, a szabadjára engedett azonosító hossza miatt igen hosszúra nyúlt osztály és változónevek miatt. Sokat kellene gépelni, hogy kiírjuk a nevüket, nem ritka a 15 karakteres azonosító sem, mint például a *System.Runtime.Serialization.StreamingContext*.



9. ábra. Visual Stúdió segítése.

Az ilyen szavak begépeléséhez kitűnő segítséget nyújt az IDE azzal, hogy felajánlja a kezdőbetű, illetve minden begépelte betű utáni lehetséges folytatásokat. A korábban használt azonosítók a lista elejére kerülnek, ezt a listát általában ctrl+. kombinációval érhetjük el, de gyakran a rendszer automatikusan is felajánlja. A megfelelő elem kiválasztása után enterrel behelyettesítődik a szövegbe, ezzel tetemes gépelési munkát spórol meg. Nem beszélve a *Visual Studio Code Snippets*-nek hívott segítségére. Ezzel rövid szavak plusz tabulátor billentyű használatával gyakran ismételt sort vagy sorokat szúr be a szövegbe.

Például : a cw + tabulátor : a `Console.WriteLine()`; Hbegépelését eredményezi.



10. ábra. Code Snippets működés közben.

2.3.5 Gyorsíró billentyűzetek: velotype veyboard

A hagyományos QWERTZ billentyűvel ellentétben, ahol a gépíró rendszerint csak egy billentyűt üt le egyszerre, és ez egy karaktert szúr a szövegbe, addig a Veyboard-on (Velotype 2010) több billentyű egyszerre lenyomásával szótagok keletkeznek szavak helyett. Ilyen eszközökkel oldhatják meg az élő műsorok szinkron feliratozását.



11. ábra. A Veyboard. (forrás wikipedia)

2.4 Magyar kutatások

2.4.1 Magyar nyelvészeti morfológiai kutatások. Alkalmazott nyelvészeti kutatások magyar nyelven

Hunmorph (*Hunmorph 2009*) Ez egy project volt a Budapesti Műszaki Egyetemen (BME) -en, amelynek az volt a célja, hogy létrehozza a magyar web korpuszt. Egy minden korábbinál nagyságrenddel nagyobb méretű magyar nyelvű, tokenizált szöveggyűjteményt, az ez alapján készülő Szószablya Gyakorisági Szótárat, a szabadon elérhető (LGPL licencű) *hunmorph* morfológiai elemzőt, a *hunstem* szótővezőt és a *hunspell* helyesírás-ellenőrzőt és a programok által használt magyar helyesírási és morfológiai szótárat. Természetes a projektnek nem volt célja a gyorsírás, de az eredményeit kiválóan fel tudjuk használni.

Magyar Ispell/Myspell (Németh László 2002) szótármodul mely szerepet kapott az *Ispell* és az *OpenOffice Myspell* helyesírás-ellenőrzőben, segíti a magyar nyelvű szövegek helyesírás-ellenőrzését. Egy szó helyességének megállapításáról csakis úgy lehet meggyőződni, hogy tudjuk, hogyan lehet generálni szavakat. Nos, ennek a szótárat és todalékfájlját fogom

felhasználni a következőkben, mert nekem is szükségem van arra, hogy egy szó után mi következhet toldalékként.

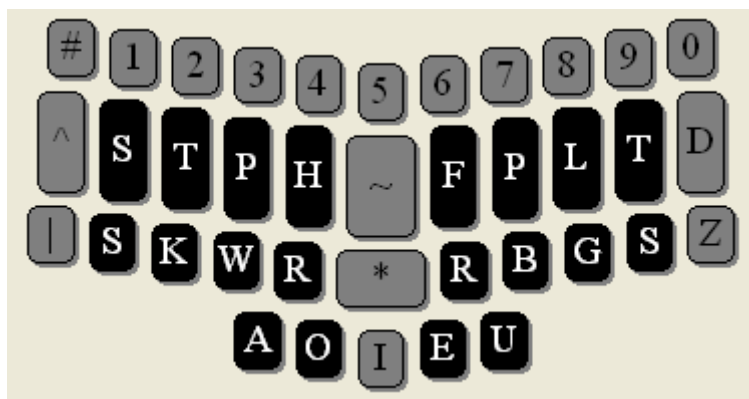
2.4.2 Magyar változatú gyorsíró billentyű

A gyorsíróbillentyű egy olyan módszer, amikor a szavakat fonetikusán írjuk le (kiejtés szerint) a speciális billentyűvel, több billentyű egyidejű lenyomásával. A gyorsíró képes a szótagokat, szavakat, több szavas kifejezéseket egyidejűleg egy lenyomással megjeleníteni. Ez „írás” a billentyűvel és nem gépelés. (Mecs Tamás 2007)

Lássuk a képét:



12. ábra Magyar szavakra is alkalmazható gyorsíró billentyű.



13. ábra. Gyorsíró billentyűzet kiosztása.

Az ABC azon betűi, amelyek itt nincsenek szerepeltetve, azokat e gombok kombinációival fogja megjeleníteni. Nem nagyon lehet a fejleményekről hallani, de találtam egy hírt miszerint már tesztelik: „Elkezdődött a modern szoftveres gyorsírás magyar nyelvi adaptációs munkája. Idén 2009. február 1-től a HARKE kollegák többedmagukkal gyakorolják a sajátos

kialakítású billentyűzet használatát és a program magyarítását.” (Mecsné Fullajtár Ildikó 2009)

Ezzel az elképzeléssel nagyon gyorsan lehet írni, csak van egy gond, hogy a biztos gyorsírás begyakorlására, a gyorsírásjelek kialakítására 24 hónapot becsülnek. (Magyar változatú gyorsíró billentyű 2010)

2.5 A magyar megvalósítás nehézségei, magyar nyelvbeli rövidítési nehézségek

Egyes nyelvekben, mint például az angolban, ahol a toldalékolás és előtagok alkalmazása szegényes, a szóelemzés gyakorlatilag egy szótár alkalmas kiépítésével megoldható. Nem így a magyarban, ahol a toldalékolt szó is kaphat toldalékot. (Ilyen esetekben relatív szótőről beszélhetünk.)

Magyarban a toldalékok szekvenciája eléggé kötött. A toldalékok (előtagok) is osztályokba sorolhatók aszerint, hogy milyen szóosztályokra alkalmazhatók, és milyen szóosztályba kerül az új szó. A különböző toldalékoknak lehet azonos az alakjuk (pl. *kutyá-nak*, *lát-nak*). Egyes toldalékok módosíthatják a relatív szótövet (tőváltozatok, hasonulások). (Naszódi Mátyás, 2001) A nehézség ott is belép, hogy a toldalékok hihetetlen változatosságban fordulhatnak elő, egy szótőnek rengeteg **helyes** ragozása is lehet, sőt van is. Például vegyük az *üzemel* szót: csak a (hu_HU.dic.openoffice.org) szótárban találtam a következő alapszavakat:

üzemelő üzemelés üzemeltető üzemeltetés üzemeltetető üzemeltetetés üzemeltetethető üzemeltetethet üzemeltetett üzemeltettet üzemeltethetőség üzemeltethető üzemeltethetlenség üzemeltethetetlen üzemeltethet üzemeltetett üzemeltetendő üzemeltet üzemelt üzemelhet üzemeli

Ezt szótárral és szavakat egyesével rövidíteni nem gazdaságos, ill ha a rövidítéshez semmitmondó karaktereket rendelünk hozzá, akkor kvázi megtanulhatatlan lesz a rövidítési rendszer.

A későbbi fejezetben felvázolt magyar szószerkezet eredményeit felhasználva (látsd Magyar szavak szerkezete) meg lehet becsülni, mennyi toldalékunk lesz.

Egy-két szám, ami érzékelteti a feladatot.

- A különböző igeragok számát 60-ra becsülik.
- Esetragok: 27-28 esetragot tartanak számon (az idejét múlt alakokkal ez még több is lehet).
- Egy átlagos szótó több ezer formában szerepelhet a szövegekben !!
- Igekötőkből körülbelül 56 van, ez a leggyakoribbak összeszámolásával jött létre.
- A magyarban minden névszó kaphat ragokat meg jeleket. A lehetséges ragok száma – modelltől függően – 17-30 lehet.
- A jelek: többes szám, birtok- és birtokos jel ettől függetlenül 14 lehetőséget ad, nem beszélve a családi többsről. Ez mintegy 600 lehetőség.

„A nagy szám nem innen keletkezik, hanem a képzőkből. Hiszen (majdnem) minden névszóból képezhetünk igét (-oz, -ít, -odik...), igéből folyamatos vagy befejezett melléknévi igenevet, vagy főnevet (-ó, -atlan, -ott, -ás), de képzők segítségével igéből keletkezhettek újabb igét (-gat, -tat, -hat), főnéből melléknevet, melléknévből főnevet (-os, -ság). Ha minden szóhoz négy különböző képző kapcsolódhat, és ezt a folyamatot háromszor megismételhetjük, akkor egy szónak $4^3+4^2+4^1+4^0=88$ képzett alakját kapjuk, ami ha névszó, 600 ragozott, jelzett alakja lehet. Ez önmagában több mint 10 000.” (Naszódi Mátyás, 2001)

Erre a helyzetre kívánok megoldást találni.

3 Tervezés

3.1 megfontolások

3.1.1 „online-offline”

A program működésének módjára több választás jöhet számításba.

- Az első a fordítóként működő gyorsítási lehetőség. Kifejtve úgy működik, hogy a begépelte szöveget kibontja a beírás (a teljes szöveg beírása) után. Ez az EasyScript-hez hasonlatos.
- A másik a gépelés közben prediktív felajánlások segítségével való gyorsítás. Amikor már sejthető, hogy a szó mi lesz, a felhasználó/gépíró interaktív formában választ gépelés közben. Ezt el se lehetett képzelni régebbi írógépen.

Az első megoldáshoz meg kell tanulni a rövidítéseket, és szórövidítési szótárra is szükség van. A másodiknál *bonyolult* program kell, ami felajánlja a lehetőségeket, és ami kellemetlen, rá kell nézni a monitorra, nem lehet *vakon* gépelni. Ami nyilvánvaló időkiesést jelent, de úgy is mondhatnánk, hogy „kezdők is elkezdhetik”, kevesebb a betanulási idő.

Az nyilvánvaló, hogy az első megoldás adhat rövidebb formát a szabadabb rövidítési lehetőségek miatt. Kérdés, akarja-e ezt valaki megtanulni, aki nem minden nap ír több oldalnyi szöveget?

Ha a kezdők kezébe adnánk egy programot, ami megtanítaná őket, hogyan kell a rövidítéseket használni a teljes szó előcsalogatására, akkor talán alkalmas alternatíva lenne a

sebességnöveléshez. Hiszen egy idő után nem kellene a tippeket/segítségeket néznie a monitoron, mert megjegyezné a legtöbbet használtakat.

Minden rövidítésen alapuló rendszernél van rövidítés felvétele rész, ami nélkül nem lehet igazán hatékony a rendszer. Különböző szótárak alkalmat adnak, hogy speciális témaköröknél betölthessünk tematikus szótárakat, olyanokat, amik nem szükségesek más témáknál, de a konkrét szövegben sokszor előfordulhatnak.

3.1.2 Magyar szavak szerkezete

Először is egy-két fogalomnak a magyarozatát írom le. A segítséget a fogalmak tisztázásához a Farkas Ernő, Naszódi Mátyás tanulmányából és Naszódi Mátyás Természetesnyelv-feldolgozás számítógéppel írásából vettem. (Farkas Ernő, Naszódi Mátyás 1990) (Naszódi Mátyás ,2001)

- **„Szónak** nevezzük a nyelv azon elemét, melyet írott szövegben egybeírunk (az ABC betűit, számjegyeket és esetleg kötőjelet felhasználva). Szóalaknak nevezzük ezek írásban megjelenő karakterlánc-formáját.” (Farkas Ernő, Naszódi Mátyás 1990)
- „A **morfológia** feladata leírni azt, hogy egy szó milyen morfémasorozattal állítható elő. A generatív leírás célja megmutatni azt, hogyan kell egy szóalakot előállítani. Az analitikus modell feladata: hogyan lehet egy szóalakból a benne szereplő morféákat megtalálni.” (Farkas Ernő, Naszódi Mátyás 1990)
- **A morféma fogalma:** A szavakat alkotóelemeire, morfémákra bonthatjuk, melyek a nyelv legkisebb önálló alakokkal és jelentéssel bíró egységei. A morfémaknak két alapvető típusuk van: tő- és toldalékmorféma.

A tőmorféma hordozza a szó jelentésmagját, s az ezekhez kapcsolódó toldalékok képezik a toldalékmorfémákat, melyek feladata a tőmorféma jelentésének megváltoztatása, módosítása vagy a mondatbeli viszony jelölése.

A tő- és toldalékmorfémák egymással szoros kapcsolatban vannak, a toldalékok legtöbbször a tőmorféma után találhatók. A tő után kapcsolódó toldalékokat szuffixumoknak, a tő előtt állókat pedig prefixumoknak nevezzük. Egyetlen toldaléktípus van, amely a tő előtt helyezkedik el: a felsőfok és a túlzófok jele.

A szóelemek típusai:

Tőmorfémák: Mind az igék, mind a névszók körében megkülönböztetünk egyalakú és többalakú töveket.

Szótó (tőmorféma), szabad morféma: Ez a szó jelentésének magva. Ha más szóelemektől függetlenül is értelemmel bír, akkor szabad tőmorfémának nevezzük. A kötött tőmorfémák önállóan nem használatosak: pl. *tav-ak, bokr-ok, alsz-ik, alv-ás*.

Félszabad morféma: Önállóan nem fordul elő, csak más szóhoz kapcsolva értelmes, jelentése járulékos, viszonyuló. Ide tartoznak a névutók, névelők, igekötők.

Toldalékmorfémák: (képző, jel, rag) Önállóan nem fordulnak elő, csak a szótóval összeolvadva. Három nagy csoportja van: képző, jel és rag. Ezek mindig ilyen sorrendben követik egymást. Kivétel: a középfok jele megelőzheti a képzőt (pl. *nagyobbít*), ritkán a rag is (pl. *háromszori*). A kötőhangot az utána következő toldalék részének tekintjük. Szerepe az esetleges mássalhangzó-torlódás feloldása. Sokszor ez határozza meg az adott szó szófaját, pl. *vörösek, vörösök, gyorsak, gyorsok*. A kötőhangok sokszor lekopott véghangok maradványai. (Jeszenői Péter 1998)

Képző A toldalékok közül elsőként követ(het)i a szótóvet. Több képző is állhat egymás mellett Nem köti meg a szó szófaját, nem zárja le a szóalakot (*levegőztet+nének*) Megváltoztathatja a szó jelentését, eredeti szófaját, illetve mondatbeli szerepét.

Jel Szerepe a szóalakban közbeékelt: a képző és a rag között áll! Több jel is társulhat a tőhöz. Megköti a szó szófaját, de azon nem változtat, csak a jelentésen módosíthat. Egy-egy nyelvtani jelentésmozzanattal gazdagítja a szótó jelentéstartalmát pl. idő-, mód-, számbeliség.

Rag A szóalak végén helyezkedik el, nem követheti semmilyen más toldalék. Meghatározza a szó mondatbeli szerepét (belefűzi a mondatba) Nem változtatja meg a szó szófaját Egy szóalak csak egy ragot tartalmazhat! (Los Csaba 2005)

A toldalékok kapcsolódási sorrendje:

SZÓTÓ + képző(k) + jel(ek) + rag ; van kivétel ritka esetekben:

Tő + jel + képző: *lent+ebb*

Tő + rag + képző (*hónap+onként*)

Ragok bővebben:

- **Esetragok:** Névszó végén álló toldalék, amely a névszóból határozót csinál (ideértve a tárgyat és alanyt is). Az esetragok meghatározzák egy egész névszói csoportnak a szerepét egy mondatban, vagy egy részkifejezésben.

27 esetet tartanak számon: t, ban, ba, ból, an, on, ra, ról, vá, nál, hoz, tól, ként, ért, nként kor, ig, szor, val, nak, ul, stul, képp(en), lag, onta, tt, szám(ra).

- **Igeragok:** Az ige végén álló toldalék, amely meghatározza az ige módját, tárgyasságát, az ige alanyának számát és személyét.

Igemód lehet: kijelentő, felszólító, múlt, feltételes, igenévi

A különböző igeragok számát 60-ra becsülik.

Képzők bővebben:

A képzők azok a toldalékok, melyek megváltoztatják a szó jelentését. Így meg lehet fogalmazni hogy miből mibe megy át egy szó ezt táblázatosan (nem teljes):

ige->ige:	-hat, -gat, -tat ...
ige->főnév:	-ás, -alom, -dalom, -at, -mány...
ige->melléknév:	-ó, -ott, -atlan, -andó ...
névszó->ige:	-al, -ez, -kodik ...
melléknév->ige:	-ít, -oz ...
főnév->főnév:	-ság, -ka, -né
névszó->melléknév:	-s, -talan, -szerű, -féle, -ú, -i, -beli ...
melléknév->melléknév:	-bb
számnév->melléknév:	-ad, -adik,

Az egy szóban található képzők számát a megérthetőség korlátozza. Így egy képző csak akkor ismétlődhet egy szóban, ha a korábbi képzett alak sajátos jelentéssel bír, így önálló képzetlen szónak tekinthető. Pl.: *biztonság-os-ság*.

Egy átlagos szótó több ezer formában szerepelhet a szövegekben.

Igekötők bővebben:

Az **igekötők** az igékhez kapcsolódnak és módosítják annak jelentését.

A leggyakoribb jelentésmódosítási típusok a következők:

- Az igekötő a cselekvés irányát jelzi. Pl.: néz-felnéz
- Az igekötő a cselekvés vagy történés időbeli lefolyásáról, állapotáról, befejezett szemléletéről vagy akár az ige tárgyával való speciális viszonyáról is tájékoztathat. Pl.: fekszik-elfekszik-kifekszik-átfekszik, lát-meglát, hull-lehull
- Az igekötő teljesen megváltoztatja az eredeti ige jelentését. Pl.: fog-hozzáfog
- Az igekötő állhat az ige előtt - ekkor egybeírjuk vele -, és állhat utána - ekkor különírjuk.

Az igekötő-használat a magyar nyelv jellegzetessége. Míg más nyelvekben - például az angolban - a cselekvés vagy történés időbeli megvalósulásának kifejezésére a nagyszámú igeidő használatos, addig ezeket a magyar az igekötőkkel fejezi ki.

Gyakoribb igekötők:

abba-	előre-	hozzá-	neki-	tönkre-
agyon-	fel-	ide-	oda-	túl-
alá-	föl-	keresztül-	össze-	újjá-
át-	félbe-	ketté-	rá-	újra-
be-	félre-	ki-	rajta-	utána-
bele-	felül-	körül-	széjjel-	végbe-
benn-	fölül-	közbe-	szembe-	végig-
benne-	fenn-	közre-	szerte-	vissza-
egybe-	fönn-	külön-	szét-	
el-	hátra-	le-	tele-	
ellen-	haza-	meg-	tova-	
elő-	helyre-	mellé-	tovább-	

Egy két példa:

El| lehet| etlen |it |és |ük

Igekötő szótó képző, képző, képző rag

Hegy |ség |ük |ból

Szótó képző jel rag

3.1.3 Myspell Hunspell Ispell

Myspell egy helyesírás-ellenőrző az openoffice Writer-hez, ami az openoffice.org Office csomagban van (a 2.0.2 verziójú OpenOffice.org lecserélte Hunspell-re)

A MySpell Ispell alapú a szóellenőrzést az *aff* és a *dic* kiterjesztésű fájljaival oldja meg. A *.dic* fájl egy szólista olyan szavakkal, amikhez betűk vannak hozzárendelve. Ezek a betűk a toldalékokat címzik, amelyek az *aff* fájlban vannak felsorolva. A *dic* fájlban egy szó csak egyszer van felsorolva. Ezzel helyet lehet spórolni, mert a szó összes lehetséges formáját képes felsorolni (mint fog fogás fogó fogott) .

A Hunspell egy továbbfejlesztett része a MySpell nek. Ez ugyancsak helyesírás-ellenőrző, alaktani elemző és generátor program, amit a morfológiailag gazdag, bonyolult összetett szavakat vagy karakterkódolást tartalmazó nyelvekhez fejlesztettek ki, amit alapvetően a magyar nyelvhez találtak ki.

A Magyar Ispell a Hunspell helyesírás-ellenőrző és morfológiai elemző programhoz készített, szabadon felhasználható magyar szótár. Egyedülállóan pontos helyesírás-ellenőrzést biztosít az ingyenes OpenOffice.org irodai csomag, a Mozilla Firefox böngésző és Thunderbird levelezőprogram, Google Chrome, Internet Explorer (kiegészítésként), Opera felhasználói számára.

Nekem csak a szógeneráláshoz kell a szótár és a toldalékfájl, így a továbbiakban nem részletezem az elemző szoftvereket. Ezért csak az *openoffice* szótárfájlját ismertetem és a magyar *Ispell*-ét.

3.1.4 A dict fájl szerkezete:

(Németh László 2009) Első sor a szavak számát tartalmazza. A további sorok szavakat tartalmaznak és '/' után azoknak a toldalékoknak a címkéjét, ami szóba jöhet nála. A címkék jelentését a magyar Ispel-ből vettem.

Például:

üzletvezető/CWMRmYD
üzletszerző/CWMRmYD
üzletiség/BVLITY
üzletiesség/BVLITY
üzleties/BDGVLT
üzleti/BDGVRL
üzletházak/IY
üzletház/AUY
üzletember/BVLIRYD
üzletelő/CDHWRMXY
üzletelés/BVLITYXD
üzleteltető/CDHWRMXY
üzleteltetés/BVLITYXD
üzleteltethető/CDHWRMX
üzleteltethet/EeX

A toldalékok címkéinek jelentése:

Ezek a betűk lehetnek címkék:

ABCDEFGHIJKLMNOPQRSTUVWXYZabcdefghijklmnopqrstuvwxyz

A: névszó mély
B: névszó magas ajakréses
C: névszó magas ajakkerekítéses
D: leg, legesleg melléknév prefix
E: ige magas ajakréses
F: melléknév mély
G: melléknév magas ajakréses
H: melléknév magas ajakkerekítéses
I: névszó morfonetikus alternáció (morfo1+morfo2)
J: névszó spec. morfon. alt. (morfo1), -hetett/-hatott igerag
K: névszó nem morfon. alt. mély
L: névszó nem morfon. alt. magas ajakréses
M: névszó nem morfon. alt. magas ajakkerekítéses
N: névszó spec. morfon. alt. (höz ön)
O: ige mély

P: ige magas ajakkerekítéses
 Q: ja
 R: je
 S: a
 T: e
 U: névszó ragok
 V: névszó ragok
 W: névszó ragok
 X: igekötők
 Y: szóösszetételben szerepelhet
 Z: idegen kiejtésű, vagy mozaikszavak kötőjeles raggal, kötőhangzó nélkül
 a: számnevek prefix, tízesek
 b: számnevek prefix, százaskok
 c: -szerű
 d: ige morf. alt. ragozás
 e: ige tárgyias magas ajakréses
 f
 g
 h: számnevek, ezeregyszáz stb. prefix
 i: -nyi mély
 j: -nyi magas
 k: főnév spec. melléknévképző, mély
 l: főnév spec. melléknévképző, magas ajakréses
 m: főnév spec. melléknévképző, magas ajakkerekítéses
 n: főnév spec. melléknévképző, morfo
 o: ige tárgyias mély
 p: ige tárgyias magas ajakkerekítéses
 q
 r
 s
 t
 u
 v: szerepelhet összetételekben első szóként (COMPOUNDFIRST)
 w: tiltott szavak kapcsolója (FORBIDDENWORD) első kapcsolónak kell lennie!
 x: szerepelhet összetételekben utolsó szóként (COMPOUNDLAST)
 y: összetett szó
 z

A szótár ilyen egyszerű, az *aff* fájlok szerkezete már egy kicsit bonyolultabb.

3.1.5 Aff fájlok szerkezete:

SFX I Y 351

SFX I e ésül e

SFX	l	e	ésért	e
-----	---	---	-------	---

SFX	l	e	éstől	e
-----	---	---	-------	---

SFX	l	e	éssé	e
-----	---	---	------	---

Példa:

A szó, amit lehet ragozni pl **ülőke** → ülőkésért ülökéstől ülökéssé

Vettem egy részletet a „hu_HU.aff „, fájlból. A betűk magyarázata pedig a következő

Az első sor magyarázata.

- SFX : SFX jelentése suffix azaz a szó végére kerül.
- L :Betű pedig a szabály nevét jelöli, ezek a szabályok egy betűvel vannak megcímezve.
- Y :Ez az oszlop mutatja hogy perfixekkel lehet e kombinálni.
- 351 : ez pedig hogy a megnevezett szabályból hány darab van. (magyarban elég sok van)

A második sor már mást jelöl az alábbi példában.

SFX	l	e	ésül	e
-----	---	---	------	---

Magyarázata:

- SFX : még mindig a szó végére kerül
- L : a szabály nevét mutatja hogy ha nem látnánk a szabály első sorát akkor is tudjuk melyik szabályt nézzük.
- E : ezeket a betűket vesszük le a szóból
- Ésül : ezeket a betűket rakjuk hozzá a szóhoz
- e :Ez az alkalmazás előfeltétele ez egy regex de mint most látjuk hogy elég ha a szó e betűre végződik. De lehetséges ilyen feltétel is: [^JLNRSZÉÍÓÖUŰÚ]T

Létezik egy másik fajta jelölés is (magyar Ispell használja), ami ugyanazzal a kifejezőerővel bír, csak máshogy jelöli a fájlban.

Például a felső szabályok ebben a jelölésben:

```
Flag *l
e > -e ,ésül
e > e,ésért
e > -e,éstől
e > -e,éssé
```

ami a szabály nevét flag *l -lel jelöli, A sor elején a regex áll , ha az teljesül akkor –e jelzi az e betű levételét a ’, ’ után van, amit hozzá kell rakni a szóhoz.

Ha nincs elvéve semmi, azaz a lenti szabály valósul meg:

```
L L > AK # hallak
[ ^L ] L > LAK # ajánllak
```

Még két megjegyzés a MySpell affix fájlokhoz

- Az első (nem kikommentezett) sor jelzi a fájlban használt karakterkódolást Pl. SET ISO8859-2
- Mind a szótárban és a affix fájlban nagybetűvel kellene szerepelnie a szavaknak.

Egy-két példa:

Szótárbejegyzés:

üzemel/EX

az EX azt jelenti, hogy a lenti szabályokat lehet rá alkalmazni.

- E: ige magas ajakréses

- X: igekötők

A rengeteg szabályt vonz ez a két címke, de most nem sorolom fel teljesen, csak az első párat.

flag *E:			
[^K]	>	VE	
[^K]	>	VÉN	
# E/1.			
[^K]	>	EK	# mer > merek (IK!!!!)
flag *X:			
.	>	ELŐ	
.	>	ELŐRE	# megy > előremegy
.	>	FEL	# megy > felmegy
.	>	FÖL	# megy > fölmegy
.	>	FÉLBE	# vág > félbevág
.	>	FÉLRE	# tesz > félretesz

Szótárbejegyzés:

üzelmek/IJY

Erre is rengeteg a szabály ezért csak részleteket emelek ki belőle.

SFX I Y 2414			
...			
SFX I	ok	aiba	ok
SFX I	ok	ai	ok
SFX I	ok	a	ok
SFX I	k	t	k
SFX I	ek	ünkül	ek
SFX I	ek	ünkön	ek
SFX I	ek	ünkéül	ek
SFX I	ek	ünkéért	ek
SFX I	ek	ünkévé	ek
SFX I	ek	ünkével	ek
SFX I	ek	ünkétől	ek
SFX I	ek	ünkét	ek
SFX I	ek	ünkéről	ek
...			
SFX J N 88			
...			
SFX J	ok	iban	ok
SFX J	ok	iba	ok
SFX J	ok	i	ok
SFX J	k	n	[^a]k
SFX J	ek	ül	ek
SFX J	ek	iül	ek
SFX J	ek	iért	ek
SFX J	ek	ivé	ek
SFX J	ek	ivel	ek
SFX J	ek	itől	ek

SFX J	ek	it	ek
SFX J	ek	iről	ek
SFX J	ek	ire	ek
SFX J	ek	inél	ek
SFX J	ek	inek	ek

...

PFX X Y 59

...

PFX X	0	újra	.
PFX X	0	újjá	.
PFX X	0	össze	.
PFX X	0	át	.
PFX X	0	által	.
PFX X	0	végig	.
PFX X	0	végbe	.
PFX X	0	vissza	.
PFX X	0	utána	.
PFX X	0	túl	.
PFX X	0	tovább	.
PFX X	0	tova	.
PFX X	0	tele	.
PFX X	0	szét	.
PFX X	0	széjjel	.

...

Megjegyzés: Természetesen a # jel után megjegyzések vannak, az értelmező ezeket kihagyja.

3.2 Mit lehetne rövidíteni

Az angolban a szót rövidítik és mivel az angolban, ahol a toldalékolás és előtagok alkalmazása szegényes, a szót elég rövidíteni, azaz nem változik a szó alakja. Természetesen vannak szómódosító toldalékok, de ez elenyészően kevés, főleg a magyar nyelvhez képest. A magyar nyelv tipikusan agglutináló nyelv, ez azt jelenti, hogy a szavakhoz több rétegben rakódhatnak végződés.

A szó előtagból, szótövből és toldalékokból épül fel. A toldalék képző, jel, vagy rag lehet, az előtag igekötő.

Ezt a három részt meg kell különböztetni, mert meg kell tudnunk, hogy előtag-e, vagy már magának a szónak a rövidítését írja.

A megkülönböztetésre három lehetőséget vázolok:

1. Az előtag kezdőbetűjével nem kezdődhetnek szavak ami meglehetősen korlátozná a szótövek rövidítését, nem beszélve az elgépelések miatt nagyon különböző, esetleg abszolút oda nem illő szót ajánlana fel a rendszer.

2. Az előtagot egy karakterrel elkülönítjük a szótótól. Ez lehet egy nem nagyon használt, de a billentyűzetten könnyen elérhető karakter.
3. Az előtagok speciális karaktereket tartalmaznak, ami utalhat (vagy nem utal) az előtag leírt formájára, de nem használhat csak betűket: pl. ö4 lenne az össze.

Én a másodikat választottam egyszerűség szempontjából. Részeket jelző karaktert már csak ki kell választani, erre két jelentkező van:

- A Pont (.) . Ha mégis ki kell írni a pontot, akkor a dupla pont megoldással ki lehet írni (..) .
- A caps lock billentyű tulajdonképpen érdemi használaton kívül van. Helyettesítését a Shift +caps lock végezné el .

Ezeket fogom én a következőkben **elhatároló karakternek** hívni.

Az elsőt választva így nézne ki egy szó: (könnyebb kiírni a pontot mint a caps lockot)

[Előtag rövidítése].[szótó rövidítése].[toldalékok rövidítése]

Ha nincs előtag az a szó rövidítéséből ki fog derülni, ha meg toldalék nincs, akkor nincs a szóban pont.

3.2.1 Előtag rövidítésére

Előtag rövidítésére ezt a táblázatot veszem alapul. De lehetővé teszem az átírását a gépelő által is, ugyanis az általa használt leggyakoribb előtagot érdemes egybetűs rövidítésre tenni.

Szó	Rövidítés	Szó	Rövidítés	Szó	Rövidítés
abba	a	fönn	fő	rajta	rt
agyon	ag	hátra	ht	széjjel	sz
alá	al	haza	hz	szembe	szm
át	át	helyre	hy	szerte	sze
be	b	hozzá	h	szét	szt
bele	bl	ide	i	tele	tl
benn	bn	keresztül	kl	tova	tv
benne	bn	ketté	kt	tovább	to
egybe	eg	ki	k	tönkre	tő

el	l	körül	kü	túl	t
ellen	el	közbe	kő	újjá	úá
elő	eő	közre	kz	újra	új
előre	er	külön	kn	utána	út
fel	f	le	l	végbe	vb
föl	fő	meg	m	végig	vg
félbe	fb	mellé	ml	vissza	v
félre	fr	neki	n		
felül	fe	oda	o		
fölül	fő	össze	ö		
fenn	fn	rá	r		

Magyar nyelvben érdemes a szótöveket, különösen a hosszú szótöveket kell rövidíteni. Például hosszabbak: Bicikli, konszolidál. Azt meg kell jegyezni, hogy az igazi szótövek általában rövidek, például rövidek: fut, jár, néz.

A rengeteg jel rag és képzővel terhelt szavakat nem érdemes egy az egyben rövidíteni, mert túl sok van belőlük. Rövidítsük ezeket külön így ezeknek a leírása is rövidebb lesz. A fent említett szótárak és toldalék (affix) toldalék fájlok segítségével nyújtanak abban, hogy a toldalék milyen kötőhangot igényel.

Pl. leves-ben, alma-ában.

3.2.2 Szótő rövidítése

A szótövet a 2+1, 3+1 módszerrel fogom rövidíteni. Ez azt jelenti, hogy a szó első két-három betűjét veszem, olykor egyértelműsítés miatt az első három mássalhangzót is érdemes venni például Rendszer → rndr. Alternatív megoldásként 3+2 is lehet.

A szótő rövidítése nem lehet értelmes szó ill. helyes szó, mert akkor kétértelmű lenne a leképezés, hiszen nem lehet tudni, hogy valaki nem akart rövidíteni és csak úgy beírta a szót, vagy tényleg egy rövidített szót írt bele.

Erre lehet kivétel, ha a szó után szóközt rak, akkor nem akart rövidíteni, de ha az elhatároló karakter billentyűt is megnyomja, akkor biztos rövidítés. Az értelmetlen szavaknál, amikben nincs elhatároló karakter használva az alábbi lehetőségek forognak fenn.

- Elírta a nem rövidített szót.
- Ez egy rövidített szó.
- Elírta a rövidítés szót.

A második szimpatikus nekünk, de sajnos a további kettővel is meg kell birkóznunk, ezt az interaktív segítségnek/kérdéseknek kell majd lekezelnie.

Egybetűs rövidítésekre a nagyon gyakori szavakat érdemes rakni. Ilyenekre gondolok, hogy: és →s , hogy →h , (a, az) →a

3.2.3 Toldalékok rövidítése, rövidítési lehetőségei:

Képzőknél:

Mit mivé lehet rövidíteni:

-ás -és	s
-hat -het	h
-andó -endő	dó, dő
-ság, -ség	g
-áz, -oz -öz	z
-at -et	at-et
-gat -get	gt
-áz -oz -öz	z
-talan -telen -etlen -atlan	-tlan -tlen
-tlen -tlan	
(nagy) -obb (kiss) -ebb	bb
leg (szó) +bb	lg

Jeleknél:

Birtokos személyjelek: -m, -d, -a, -e, -ja, -je, -unk, -ünk, -tok, -tek, -tök, -uk, -ük, -juk, -jük

Birtokos jel :é .

Ezek nem az igékhez mennek, ezért nem fogja megzavarni, ha egyforma rövidítések lesznek (Az „unk” igerag és az „unk” birtokos személyjelre gondoltam most.) de ez általában elkerülhető .

Felsőfok jelénél rövidítés: Leg [szó] -bb --lg elején határoló karakter és a végén már kirakhatja automatikusan a bb a szó végére. Illetve megoldható, hogy a rag elé jöjjön. Leg-nagy-obb-ra. Itt a lg kiírva és a -ra raggal befejezve érthetően nem a -ra után kell rakni a két b-t hanem elé.

Ragoknál:

Igeragok: -k, -m, -sz, -l, -d, -Ø, -ik -ja, -i, -unk, -ünk, -juk, -jük, -tok, -tek, -tök, -játok, -itek, -nak, -nek, -ják, -ik

Igék ragozásánál az utolsó kettő betűt használjuk a rövidítésre, ha egyértelmű, akkor az utolsót.

Az igeragokat értelemszerűen az igékhez lehet rakni, ennek kiderítéséhez a program a szótárból fogja tudni, hogy az egy ige, vagy nem. Ez azért nem olyan egyszerű, mert a szótó után a képzők szinte tetszőlegesen tudják változtatni a szó szófaját.

A **határozóragokhoz** egy táblát raktam, ez lesz a beépített rövidítése a programomnak.

Rag	Röviden	Rag	Röviden	Rag	Röviden
ban	b	ra	r	től	tl
ben	b	re	r	val	vl
ba	ba	ról	rl	vel	vl
be	ba	ről	rl	kor	kr
ból	l	nál	nl	an	n
ből	l	nél	nl	en	n
n	n	hoz	hz	ul	ul
on	n	hez	hz	ül	ul
en	n	höz	hz	lag	lg
ön	n	től	tl	leg	lg

A ragok teljesen kiírhatóak, mert ezzel rengeteg kellemetlenséget tudunk megspórolni.

Összetett szavak kezelése: az első szó általában tőalakban szerepel így szótó után elhatároló karaktert üt a felhasználó, akkor lehet tudni, hogy jöhet egy másik szótó, például Számítógépszerviz, fűgefalevél.

A ragozottnál vagy kötőjellel írjuk, vagy a toldalékban ütött elhatároló karakter jelzi az összetett szó következését, például számítógépes-szerviz.

Ragozottnál: levelesláda, fűteskapcsoló.

Apróbb észrevételek a rövidítésekben:

Például a Pistához Pistá-hz. Itt Pistá nincs a szótárban, holott Pista akart lenni, ennyi miatt már galibák lehetnek, ezért meg kell nézni az átalakult szótót, visszaalakítani a megfelelő szótóra, és így nézni. Mert nem lehet megkövetelni a Pistahz szerkezetű kiírást, habár ékezetek nélkül gyorsabban lehet gépelni.

Ezzel a megfontolásoknak végére értünk.

3.3 Módszerem hatékonysága

Szó	Hossz	Szó	Hossz
Napolhatóság	12	Naolhóg	7
üzletelhetetlenség	18	üzmlhtleng	10
üzemeltetethető	16	üzmltetethő	11
világvárosban	13	vilg.váosb	10
legmegfoghatóbb	15	lg.megfogh bb (a két b magától rakja oda)	10
üzemeltethetlenség	20	üzmltethleng	13
útbaigazítom	12	útbaigítm	9
újságkivágás	12	újg.k.vágs	10
ösztönözhető	12	ösznzhő	7

Összegezve a (rövidforma/teljes szó) hosszokat 87/130. Ez kb 66 százalékos arány, ami azt jelenti, hogy átlagban egy harmaddal csökkenti a szó hosszát. Ez alaprövidítésekkel van megoldva és az olvashatóság szempontból sem rossz.

Egy hosszabb szöveg rövidítésére példa. **Az eredeti szöveg:**

„Az operációs rendszer a híd az ember és a gép között” - Emlékszem, ezt mondtam, mikor egyszer papírra kellett vetnem, mit is értek az operációs rendszer fogalma alatt. Ha nagyon precíz akarnék lenni, természetesen hozzá kellene tennem, hogy nem feltétlenül csak az ember-gép kapcsolatot, hanem a gép-gép, program-program kommunikációját is lehetővé teszi egy OS. Viszont ebből a legtöbben mindig csak azt az oldalát tapasztalják, amikor mondjuk egy parancssort látnak villogó kurzorral, vagy egy "laza duplakattal" elindítják kedvenc stratégiai játékukat. Ha már ennyire körüljártuk az operációs rendszert, mint fogalmat, nézzük meg egy informatikus szemszögéből is, hogy mit csinál! A szakirodalom szerint két nézet terjedt el: az egyikben egy erőforráskezelővel, a másikban pedig egy virtuális géppel állunk szemben. No, eddig volt egyetlen ködös fogalmunk, és most kaptunk helyette még kettőt. A következő fejezetekben ezeket bagozzuk ki.” (Tized 2003)

A rövidített:

„Az opiós rndr a híd a embr s a gép közt” – Emlk.em, et mndtm, mikor egyszer pprk kelltt vetnm, mit is értek a opiós rndr fogm.a alatt.. Ha nagyn prez akar.nk lenni, trmsttsn h.. kllne tenm, h nm fittltlenul csak a embr-gép kapcsl.att, hanem a gép-gép, progm-progm kmmio.ját is lhtővé teszi egy OS.. Vitzt eből a lg.tő.n midg csk at a

odl.át tapszt,ják, amor mondjuk e. parc.sort látnak vilg.ó villogó krzr.al, v egy "laza dla.kattal" l.indt.ják kvnc strat.i jaek.ukat. Ha már enyire körl.jártuk a opiós rndr, mint foom.at, nézzük .m egy infom.s szmg.l is, h mit csinál! A szak.irdm szernt két nézet terjedt el: a egyik.b egy efk.vl, a másik.b pedig egy virtls géppel állunk szem.b. No, .dig volt egyetlen ködös fogma.nk, s most kaptunk heyt.e helyette még 2t. A köv.ző fejz.ekb ezeket bozgz.uk ki.

A leütött betűk szempontjából 663/809 ami 81 százalék körül van. Ha megfigyeljük a rövid szavakat illetve a rövid és ragozott szavakat nem rövidíti eléggé, ebből jön a magasabb százalékpont.

4 implementációs terv

4.1 A program jellege

Mindenekelőtt el kell dönteni, hogy milyen lesz a program jellege. Lehet beépülő program, egy „add on” az OpenOffice Writer –be vagy Emacs-ba, de lehet egy különálló program is.

A beépülő programban annyi könnyebbség van, hogy nem kell lekezelni a képernyőn végbemenő folyamatokat, minden eseményre reagálni stb. Hanem csak amire igényt tartunk arra az eseményre iratkozunk fel, de ismerni kell a befogadó program Application Programming Interface (API) -ét, kapcsolódási pontjait, vagyis hogyan kell a programmal kommunikálni, beleépülni.

Az önálló programnál csak magunkra számíthatunk, de nem kell ismerni azt a speciális másik programot, amibe be akarjuk építeni. Csak a programnyelv lehetőségei és korlátai állnak előttünk.

Én a másodikat választom, az önálló programot, a tervezés könnyebbé, függetlenebbé tételéhez. Független abból a szempontból, hogy nem kötöm egy konkrét programhoz sem. Holott ideálisan ez egy komoly szövegszerkesztő (OpenOffice vagy Microsoft Word) kiegészítés lenne, ami elég komolyan belenyúlna a gépelés menetébe.

Az implementálásra egy kellőképpen gyors objektumorientált nyelvet javaslok.

4.2 Specifikáció

A program leírása:

Ez egy olyan szövegszerkesztő, ami kezeli a rövidített szavakat, ajánlásokat és kész rövidítéseket tartalmaz a szavakra, vissza tudja alakítani a rövidített szavakat, és segít a rövidített szavak tanulásában.

Saját szótárak segítségével a szöveg utólagos kibontását is elvégzi önállóan akár emberi választások segítségével. Lehetőség van saját rövidítések szótárak alkalmazására.

Nem kezeli célzottan a helyesírást, (hibás szavakat is fel lehet venni) ezt már a kialakult helyesírás-ellenőrzőkkel kell megtenni.

Fogalomszótár:

Szövegszerkesztő: Amivel szöveget lehet szerkeszteni általános editálási műveleteket beleértve.

Gyors szövegszerkesztő: A tervezett program neve.

Elhatároló karakter: Az a billentyű, amit kijelölünk arra, hogy a szó előtagját elkülönítse a szótótól, és a szótövet a toldalékoktól.

Felhasználó: Az a személy, aki használja a programot.

Szótár: Ezek szótárfájlok, amelyek szavakat, vagy toldalékokat tartalmaznak. A szótárfájl és a toldalékfájl ad egy szótárt közösen.

Rövidítés szótár/ beépített rövidítésszótár: Olyan fájl, ami a szavak rövidítését tartalmazza. Más néven a beépített rövidítésszótár. Ennek a legenerálásához a tervezéskor megadott szabályok szerint van elkészítve. Ez egyszer kell csak legenerálni. Így futási időt takaríthatunk meg vele.

Saját/felhasználói szótár: Olyan fájl, amiben az eredeti szó és rövidítése szerepel.

Kibontott szöveg: Olyan szöveg, ami nem tartalmaz rövidítéseket, csak (helyes) magyar szavakat.

Config fájl: ebben tárolja a program a mentett konfigurációkat.

- Automatikusan megnyitandó fájl neve elérhetése.
- Automatikusan betöltendő felhasználói szótárak.

.rff kiterjesztés : Ez a rövidített szöveget tartalmazó, vagy nem teljesen kibontott szöveget tartalmazó fájl kiterjesztése.

.röv kiterjesztés: A saját rövidítésfájl kiterjesztése. Szerkezete:

Szó : szó rövidítése

Folyamatok:

Rövidített szöveg visszaalakítása: Ez történhet automatikusan és segítséggel.

- Automatikus visszaalakítás: A felhasználó elindítja a rövidített szövegre a visszaalakítást, és a szótárból, és rövidítés-szótárak segítségével visszaalakítja a szöveget. Ez addig automatikus, amíg egyértelműek az átalakítások.
- Segítségese átalakítás: Az automatikus visszaalakítás nem egyértelmű, esetleg nem talált rövidítéssel találkozik akkor a segítségese átalakítással fog futni a visszaalakítás ami a felhasználó döntésével alakítja vissza. Addig fut míg nem egyértelmű a transzformáció utána automatikusba vált vissza.

Automatikus szóátalakítás: A begépelés folyamatában a begépelte szó illetve szórész (előtag, szótő, toldalék) kiírása után azonnal visszaalakítja a szót a rövidített állapotából.

Szó tanítás: Egy hosszú vagy összetett szó rövidítésének átírása. A saját szótárfájlba írja az eredeti szót és rövidítését, ekkor ez lesz az elsődleges rövidítése a szónak.

Interaktív szótagválasztás: A begépelés pillanatában a szótag további értelmes folytatásaira ad javaslatot azzal, hogy a megfelelő betűket (rövidítéseket) és annak jelentését (kibontását) jeleníti meg a képernyőn.

Szöveg mentése: A szöveget attól függően hogy tömörített vagy kibontott lehet menteni. A kibontottat rtf, doc ,txt formátumban lehet menteni. A rövidítéseket tartalmazó rff kiterjesztésre mentődik.

Szöveg betöltése: Szövegszerkesztőhöz hasonlatosan. Ismert fájlformátumok rtf, rff, doc (esetleg).

Szótárfájl betöltése/mentése: Lehetőség van saját szótárfájl betöltésére és kimentésére.

Nyilvántartott adatok:

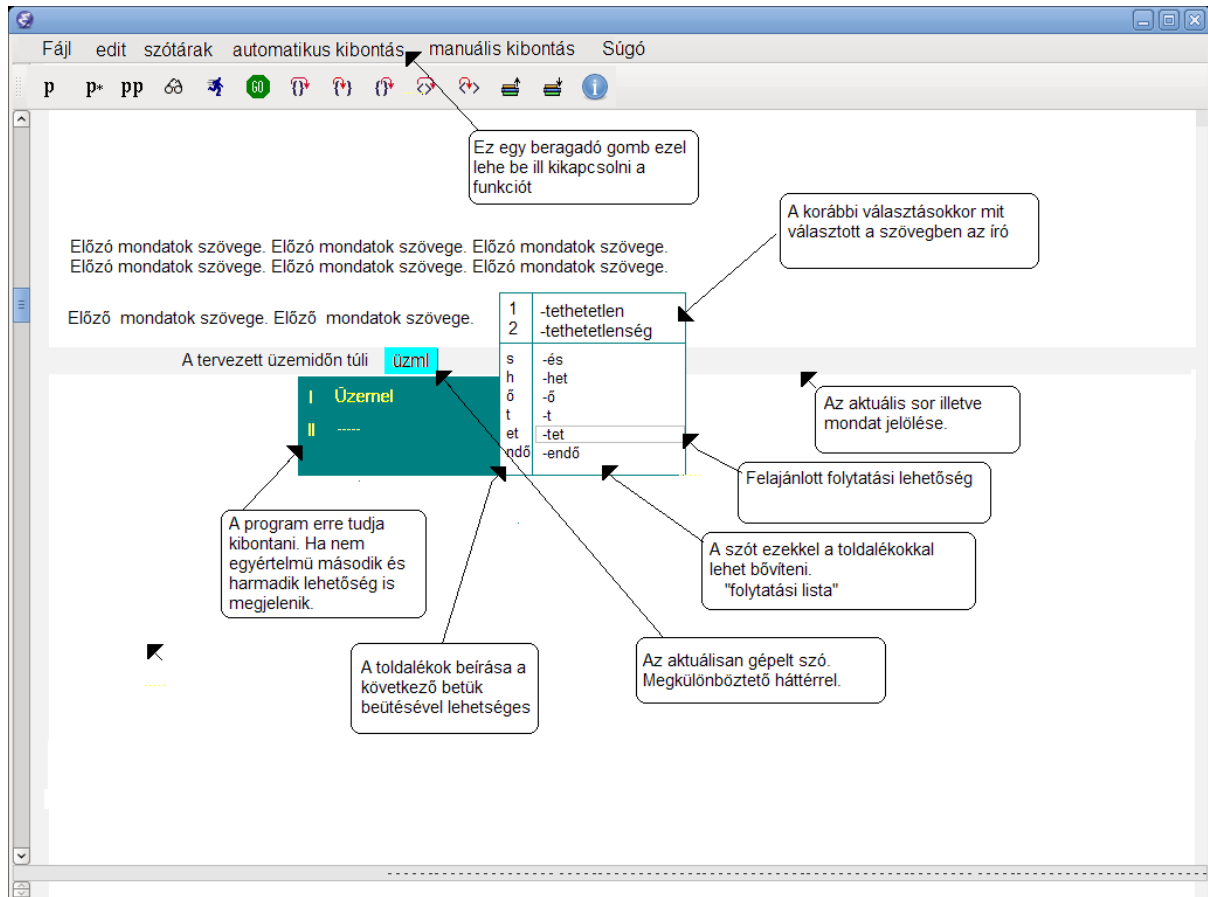
Ezek azok az adatok, amit a program futása során a programnak kezelnie kell.

A következőket érdemes berakni a memóriába és ezt is teszem, hiszen pár megabájtos adathalmaz a memóriában tartása manapság nem gond, kifejezetten ajánlott is.

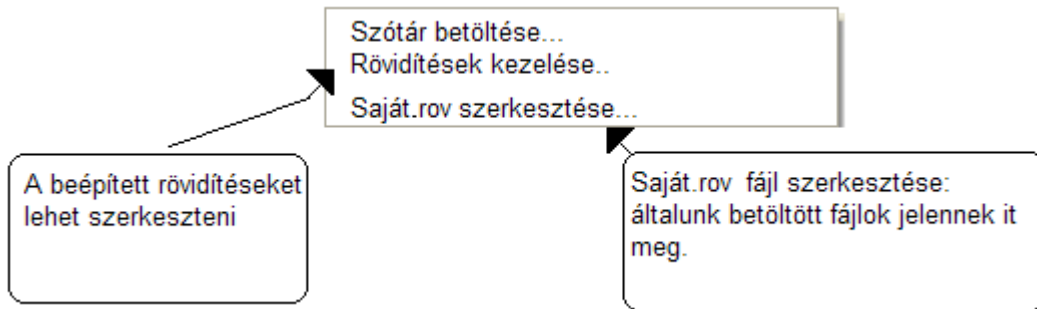
- A szótárfájlok tartalma a dic és aff fájlok.
- Rövidítésszótár tartalma. A beépített rövidítésszótár.
- Saját szótárfájl tartalma, azaz felhasználó által definiált rövidítésszótár.
- A beírás alatt lévő szó karaktereit.
- Az eddig beírt szavakat, magát a szöveget. (Ennek egy része lehet akár háttértárolón is.)
- Hol járok a szövegben. (Természetesen vissza is lehet menni a szövegben átírni a szavakat vagy újjakkal bővíteni a szöveget egy tetszőleges helyen.)
- A korábbi rövidített szavakra milyen toldalékot választott. (Ezt a választást lehet egy gyorsgombbal megismételni, a ragok újraírása helyett.)
- Konfiguráció. Vannak e automatikusan betöltendő felhasználói szótárfájlok, vagy automatikusan betöltendő dokumentum. Ezt a háttértárolón tárolja egy cfg fájlban.

Képernyőterv:

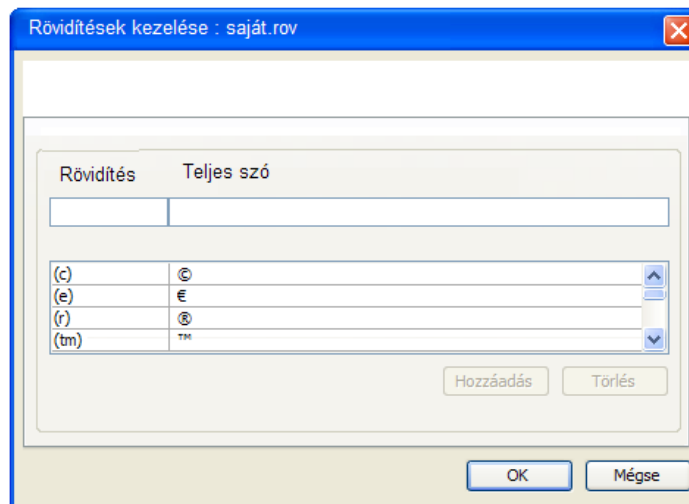
Képernyőtervemet az Emacs program felületének felhasználásával készítettem el.



14. ábra. A szerkesztő felülete.



15. ábra. A szótár menü kibontása.



16. ábra. Rövidítések kezelése.

Forgatókönyv (folyamatok működése):

Program elindulása: Megnyitás után betölti a szótárakat a memóriába, ha talál felhasználói szótárat, amit be kell tölteni a config fájl bejegyzése alapján.

Szó beírása: szó beírásakor nyilvántartja, hogy a szóban hol járunk. előtag |szótó |toldalékok résznél. Ennek függvényében ad folytatási lehetőségeket:

- Előtagonál nincs folytatási tipp ugyanis két szavas rövidítéseknél csak az eredményt írja ki. Ha többet ír mint 2 betű az azt jelenti hogy már szótövet gépel a felhasználó.
- Szótónél kiírja a folytatási lehetőségeket és az idáig beírt betűből következő szavak listáját. Ezt a szó alján elhelyezkedő szövegdobozban tünteti fel. Ez nem lehet több három-négy szónál, mert már akkor zavaró lehet. Ha ezek közül választani akarunk, akkor backspace billentyű és az alternatíva előtt lévő szám megnyomásával jelöljük ki. (képernyőterven most római számmal jelöltem) Erre az alsó listán első helyre fog ugrani a választott szavunk és a rövidítése beíródik az aktuálisan gépelt szó helyére. Két betű után kezd el tippeket ajánlani kivéve, ha van egybetűs rövidítés. Ha van, akkor kiírja. A folytatási listán szereplő felajánlott folytatási lehetőséget enter billentyű megnyomásával lehet beilleszteni. Ha szótó után üt elhatároló karaktert a felhasználó az összetett szó írását is jelentheti.
- Toldaléknál A korábbi választott toldalékok szerepelnek, egy maximum háromelemű listában, ezt számok beütésével lehet elérni. A toldalékok folytatási listán jelennek

meg. Előhívásához a táblázat baloldali oszlopában látható rövidítést kell begépelni. Lehetséges teljesen kiírni a ragokat is, ugyanis akkor a kibontott szót és a megadott toldalékot nézi meg, ha értelmes, akkor nem jelez hibát.

- Toldalék után is üthet a felhasználó egy elhatároló karaktert, ami azt jelenti, hogy összetett szót fog írni. Ezt új szótóként kell kezelni.

Szó végét szóköz billentyű leütése jelzi. Ha nem sikerült kibontani, akkor aláhúzza, illetve megjelöli a szót hogy általa hibás. Ha az automatikus kibontás be van kapcsolva, akkor kibontja a szót.

Szó és a toldalékainak meghatározása: A szótár, és rövidítés-szótárak segítségével megnézi, hogy fel tudja-e bontani a beírt szót. Ha megvan akkor a szótárból kikeresi a szót és a hozzárendelt címkéket. A címkéket kikeresi a toldalékfájlból és megállapítja a toldalék regex kifejezéséből, hogy melyiket lehet alkalmazni a szóra. Az alkalmazhatóakat egy listába rakja, amit majd meg kell jeleníteni.

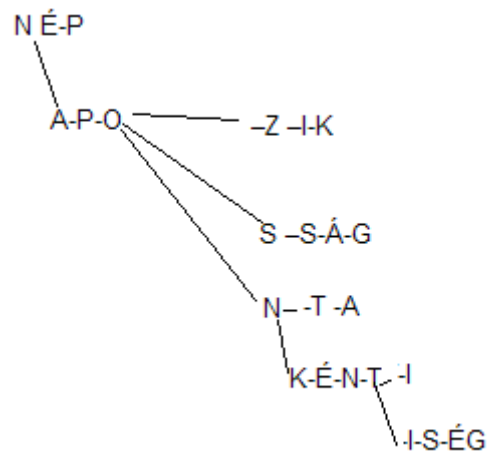
Felhasználói szótár betöltése: A felhasználó betölti a szótárfájlt a szótárak menü segítségével. Eztán ez a rövidítési szótár aktív lesz.

Felhasználói szótár szerkesztése: A szótár menüben kiválasztja a felhasználó a betöltött szótárát a [szótárneve].rov szerkesztése... menüponttal és elkezdheti szerkeszteni.

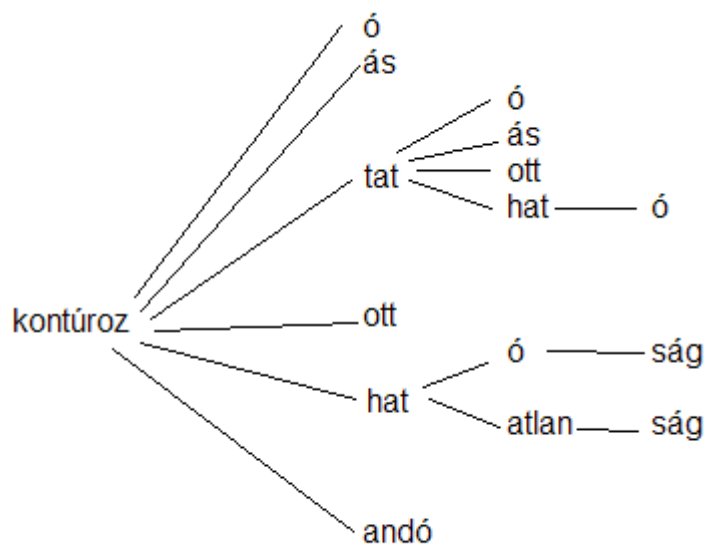
Gyorsítási lehetőségek: A szótárak elég nagyok tudnak lenni, ami természetesen könnyen befér a memóriába, de a keresés, ha sok adat/szó van az erősen erőforrás zabáló lehet. Sőt a felhasználói élmény illetve a használhatóság is csökkenthet, ha nem azonnal dobja fel a segítséget. Gyorsírásnál pedig a sebesség nagyon fontos tényező. A gyorsításra a szótárfájlok különleges adatszerkezete adhat lehetőséget. Itt ugyanis nagyon sok szó eleje ugyanaz. Erre egy speciális adatszerkezetet a szófát használva csökkenthető az eltárolt adatok mennyisége és keresési ideje.

Ez a fa lehet betűszintű - Amikor a fa elemei betűk. És lehet szótag, ill. toldalékszintű is.

Példa betűszintű fára:

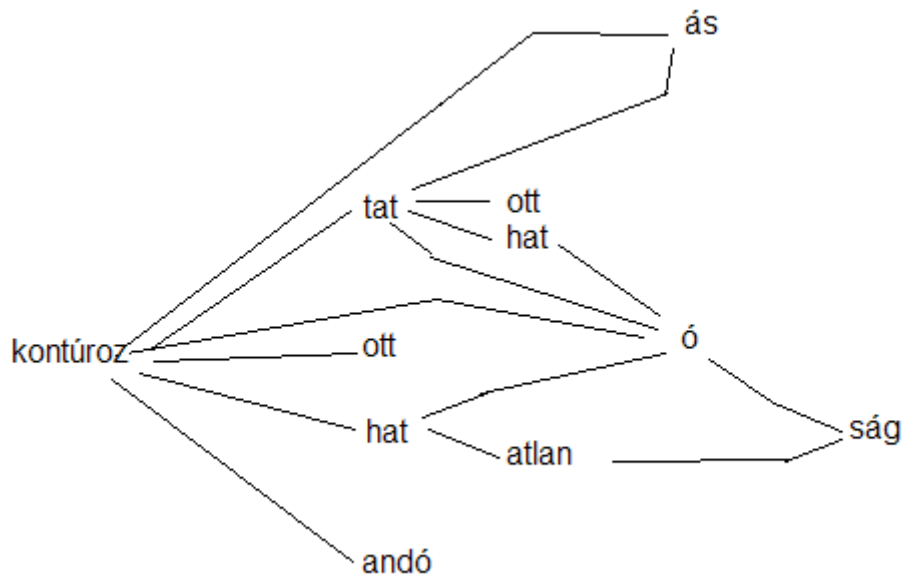


17. ábra. betűszintű szófa.



18. ábra. Toldalékszintű szófa.

Lehetne a ezt a fákat kicsit átalakítani ami azzal jár hogy fáról már adatszerkezeti szinten nem beszélhetünk de a tulajdonsága kárpótol minket.



19. ábra. „körösített” szófa.

Ezzel csökkent a szófa elemeinek száma, és a tömörítés végett itt lehet a betűfa átalakításával nyerni időt, ugyanis egy szó rövidítése általában az eleje és végéből vett egy két betű vagy a végéből vett egy-két mássalhangzó rövidíti. A keresés menete úgy néz ki, hogy az elején és a szó végén egyszerre indulunk és a kapott rész lesz a rövidített szavak kibontási halmaza. Így hamarabb találhatjuk meg a keresett szót mintha csak az elejéről indulnánk el. Hiszen mindig hasonlítani kellene a szóvégekkel, ami időbe telik. Ennél a szerkezetnél már tudjuk, hogy olyan szavak halmazában keresünk, aminek megfelelő a vége, már csak a köztes halmazt kell ellenőrizni.

4.3 Jövőbeli fejlesztési lehetőségek

Ez természetesen ez egy első verzió, ami csak az alap funkciókat tartalmazza így nagyon sok fejlesztési lehetőség adott. Gondolok arra, hogy lehessen a szavak rövidítését szóbeírás közben vagy közvetlen utána szerkeszteni, ne kelljen a menübe menni és ott átállítani. Extra billentyűk bevonásának lehetősége, amik egy gombnyomásra szavakat akár mondatokat is beírnak. Ezeket esetleg funkcióbillentyűkre rakni pl. F1, F2 stb. Új szavak felvétele az alap szótárba, amit megfelelő címkével lehessen ellátni. Itt speciális szakszavakra gondoltam. A legnagyobb fejlesztési lehetőség, amit érdemes elgondolni, mégpedig, hogy egy népszerű szövegszerkesztő beépülő moduljaként elkészíteni. Így a megszokott környezetről nem mondanánk le, csak az extra funkcióját élveznénk a programnak.

5 összefoglalás

Dolgozatomban sikerült egy szisztematikus rövidítési rendszert alkotni, ami könnyen bővíthető és papíralapú rendszerként is használható. Nem kell nagyon nagyszámú rövidítést megtanulnia a használójának, szabályok mentén könnyedén meghatározhatók a rövidítések vagy éppen a kifejtett alak. A számunkra nem tetsző rövidítéseket a program működése közben megváltoztathatjuk, az ismeretlen rövidítéseket pedig megtanulhatjuk.

A dolgozatomban ismertettem ezen túl mindazokat a hazai és nemzetközi törekvéseket, amelyek az adatrögzítés folyamatát próbálják meggyorsítani különféle eredményekkel. Továbbá röviden beszámoltam az általam felhasznált nyelvészeti kutatásokról, melyekre építettem a saját rendszeremet.

6 Irodalomjegyzék

Advantage Software , *Magyar változatú gyorsíró billentyű angol változata*, 2010,
<http://www.eclipsecat.com/> (letöltve 2010. április 9.)

Farkas Ernő, Naszódi Mátyás 1990: *Magyar nyelvű mondatok elemzése természetes nyelvű interface céljából* www.morphologic.hu/downloads/publications/nm/tenyel.ps
(letöltve 2010. március 14.)

Halácsy Péter (oldalkészítő), *Hunmorph - open source morphological analyzer*, 2009,
<http://mokk.bme.hu/resources/hunmorph> (letöltve 2010. március 18)

Jeszenői Péter 1998 *A szóelemek*, <http://users.atw.hu/e-tetel/NYELVTAN/03MORFEM.DOC> (letöltve 2010. március 14.)

Jon Knowles , 2010, *ABCZ Typing Abbreviation System*
<http://home.earthlink.net/~agjon/abczrule.htm> (letöltve 2010. március 18)

Levin, Leonard D, 2001, *EasyScript Express: Learn To Take Fast Notes in a Matter of Hours*,
Legend Publishing, ISBN-10: 1893726002

Los Csaba 2005: *A morfémák (szóelemek) szerepe és helyes használata a szóalak felépítésében*,

- http://realonline.uw.hu/erett/download/nyelvtan/11.a_morfemak_szerepe_es_helyes_hasznalata_a_szoalak_felepiteseben.doc (letöltve 2010. március 14)
- Mecs Tamás 2007: *Magyar gyorsíró billentyűzet, ppt prezentáció*,
infovilag.hu/data/files/129375568.ppt (letöltve 2010. április 9.)
- Mecsné Fullajtár Ildikó 2009, www.harke.hu halláskárosultak hírlevele 2009.ápr.25
http://www.harke.hu/hirlevel_2009apr25/ (letöltve 2010. március 20)
- Murlene Wallace Asadi, 1989, *Stenoscrypt ABC Shorthand*, Paradigm Publishing, ISBN-10:
1561184543
- Naszódi Máttyás , 2001, *Természetesnyelv-feldolgozás számítógéppel*,
sziami.cs.bme.hu/~gervarro/teny/tny.doc (letöltve 2010. március 14.)
- Németh László, 2002, *Magyar Ispell/Myspell* <http://www.szofi.hu/gnu/magyarispell/>
(letöltve 2010. március 18)
- Németh László, 2009, *Converting Affix Files: Understanding the Affix File Format*
<http://lingucomponent.openoffice.org/affix.readme> (letöltve 2010. március 14.)
- Németh László,2010, *Hunspell*, <http://hunspell.sourceforge.net/> (letöltve 2010. március 14.)
- Szeghalmy Szilvia, 2007, *Közel optimális billentyűkiosztás készítése a magyar nyelvhez*
:Debreceni Egyetem diplomamunka
- Speedwriting Overview*, 2006, http://www.easyscript.com/uses/speedwriting_overview.asp
<http://home.earthlink.net/~agjon/abczrule.htm>
- Tized (felhasználónevű cikkíró), 2003, *Az operációs rendszer és a kernel*
<http://prog.hu/cikkek/861/Az+operacios+rendszer+es+a+kernel.html>
- The Dvorak keyboard, 2003, <http://www.mit.edu/~jcb/Dvorak> (letöltve 2010. március 12.)
- Velotype Wikipedia bejegyzés 2010 *Velotype* , <http://en.wikipedia.org/wiki/Velotype>
(letöltve 2010. április 9.)