

## FEHÉRBOTOS PINGVINEK, AVAGY MAGYARUL BESZÉLŐ LINUX

Előadásom az első magyar, látássérülteknek készült Linuxon futó beszélőprogramról szól. Röviden elmondom a fejlesztés fázisait és előadásom végén konkrét alkalmazásokat mutatok be.

A fogyatékos embereknek, ha lépést akarnak tartani a társadalommal, ki kell alakítaniuk a megfelelő módszereket, melyek segítségével hozzájuthatnak a számukra is nélkülözhetetlen naprakész információkhoz. Ma ennek leghatásosabb módja a különféle számítógéphálózatok használatának minél jobb elsajátítása.

A látássérülteknek készített szoftverek jelentős része az internet használatának megkönnyítését célozza. Néhány fontos terület még a tanulás segítése, a szórakozás (zenehallgatás, olvasás) valamint a másokkal való kapcsolattartás.

### Számítógéphasználat látás nélkül

Köztudott, hogy azok az emberek, akik csak kevéssé, vagy egyáltalán nem tudják használni a szemüket, a látás hiánya miatt kiesett információkat a többi érzékszerv intenzívebb használatával próbálják pótolni. A számítógép kezeléséhez a hallás és a tapintás hívható segítségül. A text-to-speech (TTS), írott szöveget beszéddé alakító programok teszik lehetségessé a képernyőn megjelenő információk meghallgatását. Ez a leginkább elterjedt módszer a számítógép használatára, mert gyors, hatékony s nem igényel külön hardware beruházást. A segédprogramok másik nagy csoportját alkotják azok a meghajtóprogramok, melyek pontírasú kijelzőkön teszik lehetővé a számítógép által közölt információk elolvasását. A pontírasú kijelzők általában 40-80 karaktert tudnak egyszerre megjeleníteni, s navigációs billentyűk segítségével választható ki, hogy a képernyő egy adott részén lévő, a felhasználót éppen leginkább érdeklő információ legyen rajtuk pontírásban olvasható. Ezek az eszközök meglehetősen borsos áron szerezhetők be, s ez komoly akadálya az elterjedésüknek. Ezen kívül használatukhoz szükség van arra, hogy a felhasználó jól – azaz pontosan, s kellő sebességgel – tudja olvasni a braille írást. Ezeket az elvárásokat az érintettek jelentős része sajnos nem tudja teljesíteni.

## MS-DOS

Ha PC-ről beszélünk, akkor azt mondhatjuk, hogy az első széles körben elterjedt operációsrendszer ezen a hardveren a MS-DOS volt. Ez az operációsrendszer a maga karakteres felületével a vak emberek számára könnyen elsajátítható volt. A kilencvenes évek elején készült el a látássérült körökben rendkívül népszerűvé vált Brailab PC. Könnyű kezelhetősége, rövid reakcióideje, stabil felhasználói programjai révén rászolgált az elismertségre. MS-DOS-os környezetben remekül működött.

## WINDOWS

A Windows térhódításával a magyar látássérült felhasználók nehéz helyzetbe kerültek. Sokáig megfelelő segédeszköz híján mindenki úgy boldogult, ahogy tudott. Leginkább amerikai gyártók által készített szoftverek voltak használatban, de ezek magyarul nem, vagy különféle kényszermegoldásokkal, nehézkesen voltak használhatók. Jelentős előrelépésnek tekinthető a ma leginkább használt JAWS program honosítása, mely sok technikai problémát megoldott. Azonban a végrehajtani kívánt feladatok jelentős része csak részben, vagy egyáltalán nem megoldható a Windows grafikus környezetében a vakok számára.

## LINUX

A Linux operációsrendszer egyre nagyobb teret hódít magának. Egyre többen választják, mert számos olyan tulajdonsága van, mely rendkívül előnyös mind amatőr felhasználók, mind szakemberek számára. Ezek közé tartozik a nyílt forráskódú szoftverek nagyon magas aránya, a rendkívül fejlett hálózati infrastruktúra, a biztonság, a rugalmasság, a személyreszabhatóság, az alacsony hardware igény, stb. A Linux környezet karakteres felülete a látássérültek számára különösen előnyös. Az operációsrendszer karakteres felületén szinte minden olyan feladat megoldható, amit az egyébként grafikus felületen futó programokkal a látó emberek néhány egér klikkeléssel el tudnak végezni. Megfelelő képzettség esetén Linux alatt a vakok közel azonos hatékonysággal tudják használni a számítógépet, mint látó társaik. Ezt felismervén, külföldön régóta léteznek Linux segédprogramok látássérültek számára (Csehországban például már évtizedes hagyománya van a fejlesztésnek). Non-profit szervezetek, alapítványok jöttek létre, melyeket állami intézmények, vállalatok támogatnak. Az utóbbi években a szabványosítás ezen a területen is megkezdődött. Ennek köszönhetően a különböző komponensek kompatibilisak egymással. Hazánkban eddig Linux környezetben működő, vakok számára készült, magyar TTS nem létezett, én most elkészítettem ennek első változatát, melynek neve: Linbrailab.

## A Linbrailab (magyar TTS) elkészítésének folyamata

Linuxon futó TTS már sok nyelven működik. A legjobban persze az angol anyanyelvű felhasználók vannak „elkényeztetve”, de szinte az összes európai nyelven megszólalhat már egy alkalmasan felkészített program. Ez az eredmény annak a belga kezdeményezésnek a hozadéka, mely a kilencvenes évek közepén indult útjára. A TCTS Lab projektjének a neve Mbrola (Multi Band Resynthesis Overlap Add). A fejlesztők egy keretrendszert biztosítanak, melynek a segítségével mindenki elkészítheti a saját anyanyelvére jellemző hangokat egy adott formátumban, melyet egy erre a célra előállított lejátszó program azután fel tud dolgozni. A rendszer bemutatása meghaladja az előadás kereteit. A fejlesztés eredeti célja a beszéd kutatás, különösen a prozódia megfelelő generálásának támogatása volt, annyi nyelven, amennyin csak lehetséges. A gyakorlatban azonban a szoftver nagyon jól bevált képernyőolvasó programok beszédkiemeneteként is.

Mivel a MBROLA nem közvetlenül szövegből állítja elő a beszédet, így szükség van még egy közbülső láncszemre, mely a kimondandó szövegből előállítja azokat a beszéd generálásához szükséges paramétereket, mellyel már a MBROLA beszédkészítő alkalmazása felhívható.

Ez a hiányzó láncszem, a Festival nevű rendszer. Az edinburghi egyetemen fejlesztett, több nyelvű, saját TTS motorokat is tartalmazó komplex szoftverkörnyezet alkalmas arra, hogy a kívánt paraméterek előállítása után meghívja külső beszéd-szintézist előállító programként a Mbrola alkalmazást. Jelenleg a Festival a leggyakrabban használt beszédkiemenetet előállító eszköz Linux/Unix környezetben.

Az előzőek ismertetése azért volt fontos, mert hasonló elven működik a magyar Linuxos TTS program is. A KFKI Beszéd és Rehabilitáció-technológiai Laboratóriumában, Arató András vezetése alatt az elmúlt évtizedekben több magyar TTS-t fejlesztettek ki. Készséggel bocsátották rendelkezésemre egy windows-on futó TTS forrását. Ez alapján elkészítettem a Linuxos változatot, a Linbrailabot. A Linbrailab az Mbrola és a Festival együttműködésének mintájára készült, azzal a különbséggel, hogy a Linbrailab esetében a Festival az eredeti input szöveg beolvasása után rögtön felhívja a Linbrailabot, mint külső programot. Ez azt jelenti, hogy nincs szükség a szöveg semmilyen előfeldolgozására.

Ezen a ponton joggal tehető fel a kérdés, hogy akkor a Linbrailab esetén mi szükség van a Festivalra? A Festivalba való beágyazás azért volt fontos, mert ez biztosítja a kompatibilitást a világszerte létező egyéb, látássérültek számára készített segédprogramokkal.

A Linuxos brailab szoftver úgynevezett formáns szintézissel állítja elő a beszédet, ehhez az angol matematikusról, James Klatt-ról elnevezett Klatt modellt használja. A program lelke két bináris fájl, a binadata.bin és a binahank.bin. Ezekben találhatóak a hangok előállításához szükséges paraméterek, illetve a hangkapcsolatok generálásához nélkülözhetetlen adatok. A beszédgenerálásban nem a tiszta hangzók előállítása a nehéz, hanem a különböző hangok egymáshoz kapcsolódó szakaszainak megfelelő kialakítása. A beszéd előállításában szerepet játszik még néhány egyszerű prozódiai szabály, mely az alaphang frekvenciájának folyamatos változtatásával próbálja létrehozni a mondat dallamát. Részletesebben sajnos nincs mód a jelenlegi előadásban a beszédgeneráláshoz kapcsolódó nyelvészeti problémák boncolgatására.

**A Linbrailab interaktív tétele: hozzákapcsolás képernyőolvasó programokhoz**

Egy TTS elkészítése (jelen esetben a Linbrailabé) önmagában még nem elég a számítógép interaktív használatához. A képernyőn megjelenő összes szöveg szolgai (folyamatos) felolvasására nagyon ritkán van csak szükség, és rendkívül lelassítaná a számítógép használatát, ha a gép mindig mindent az első betűtől az utolsóig felolvasna. Szükség van egy programra, amely lehetővé teszi a képernyőn megjelenő információk közötti szelektálást. Ezt a programot nevezzük képernyőolvasó programnak. Többféle képernyőolvasó program létezik.

A Linux karakteres felületéhez az Emacs környezetet választottam. A grafikus környezetben a Gnome felületen működő Gnopernicus nevű képernyőolvasó programot használom.

### 1. Karakteres felület - Emacs környezet - Speech dispatcher

Néhány évvel ezelőtt indult a Speech-dispatcher nevű programcsomag fejlesztése. A cél, melyet a létrehozók maguk elé tűztek, ambíciózus, s nagyon fontos lépés a TTs-ek egységes kezelése felé. Hasonlóan egy grafikus könyvtárfájllhoz, melyben implementálják a különféle objektumok létrehozását, s csak egyetlen függvény felhívására van szükség, például egy kör megrajzolására, a Speech-dispatcher

olyan rutinokat tartalmaz, mely a különféle hardver és szoftver TTS-eket egységesen kezeli. Ez lehetővé teszi, hogy a felhasználói programokat meghangosító alkalmazások magasabb szinten kezelhessék a beszéd kimeneteket, nem kell az egyes modellek konkrét implementációjával részletesen foglalkozni.

A Speech-dispatcherhez készítettek Emacs interfészt, s mivel a Speech-dispatcher kimenete lehet a Festival, így a rendszer már magyarul is meg tud szólni.

## 2. Grafikus felület – Gnome – Gnopernicus képernyőolvasó program

Miért kell látás nélkül grafikus felületet használni?

Alapvetően azért, hogy alkalmazkodjunk a környezetünkhöz, a többiek által használt programok egy részéhez nem férhetnénk hozzá más módon. Van, aki kényelmi okokra is hivatkozik; aki nem szeret sokat gépelni, itt menüből választhatja ki a kívánt funkciót.

A Gnopernicus programról

A Gnopernicus program a Gnome felületen futó képernyőolvasó, nagyító program. Tervezésekor gondoltak a teljesen vakok, az aliglátók és a gyengénlátók szükségleteire is. A program képes kezelni TTS motorokat, illetve számos braille kijelzőt is.

A Gnopernicus használata

A program alapkonceptiója, hogy az összes gyakran használt funkciót legyen könnyen hozzáférhető. Ennek érdekében a felhasználó a numerikus billentyűzeten minden szolgáltatást megtalál. Mivel több funkcióval rendelkezik a szoftver, mint amennyi billentyű rendelkezésre áll a kijelölt területen, a tervezők úgynevezett layereket, szinteket vezettek be. A felhasználónak először szintet kell választania, majd az adott szinten belül már a kívánt funkcióhoz tartozó billentyűparancsot használhatja. Az eljárás hasonló, mint a telefonáláskor használt előhívószám. A Gnopernicus esetén ez úgy értendő, hogy a funkciókat logikai kategóriákba sorolják, s először kategóriát kell választani. A tervezők azt is elősegítik, hogy a felhasználó könnyebben megjegyezhesse a különféle billentyűkombinációkat, hiszen mindegyik parancs társítható egy kategóriához.

A számbillentyűzeten 17 billentyű található, ebből azonban a numlock-nak mindig bekapcsolt állapotban kell lennie. A maradék 16 billentyű felhasználható a kívánt célra; ez tehát azt jelenti, hogy összesen elméletileg 16 szintet definiálhatunk, s mindegyik szinten 15 funkciót helyezhetünk el. A szintek közötti váltást úgy oldották meg a tervezők, hogy a „0” billentyű lenyomása után 5 másodpercen belül le kell ütni a kívánt szinthez tartozó billentyűt. Ezért nem lehet a „0”-hoz funkciót rendelni. A program jelenlegi fejlettsége 11 különböző layer használatát teszi szükségessé. Egy-egy szinten logikailag összetartozó funkciókat helyeztek el. Így például a 8-as billentyűhöz tartozó layerhez rendelték a beszédkimenet vezérlésének paramétereit. Itt található az a funkciók, melyekkel állítható a beszéd tempója, hangereje, hangmagassága.

#### A Gnopernicus komponensei

Az alábbiakban röviden összefoglalom, hogy az eddig felhasználói szemszögből vázlatosan bemutatott grafikus képernyőolvasó program milyen belső szerkezettel rendelkezik, felsorolom a szoftver moduljait, tömören ismertetve, hogy milyen funkciót látnak el.

#### Az SRCore komponens

Az SRCore a program központi modulja. Ez a modul gyűjti össze az információkat a különféle input forrásokból, rendezzi azokat, majd a beállított szükségleteknek megfelelően kiküldi őket a három lehetséges kimenethez: braille, beszéd, nagyító. Szintén ez a modul tartja nyilván az egyedi felhasználói beállításokat, s értesíti a megfelelő modulokat az esetleges őket érintő változásokról.

#### Az SRUtil komponens

Ennek a modulnak a segítségével tudnak a program különböző komponensei egységes módon egymással kommunikálni. Az információcsere egymásnak küldött üzeneteken keresztül történik. Ebben a modulban található az összes előforduló eseménytípus definíciója, illetve a megfelelő struktúra, amely az események feldolgozásához szükséges összes adat tárolását biztosítja.

#### Az SRLow komponens

Az SRLow komponens figyeli a rendszereseményeket, valamint a felhasználó által kiadott parancsokhoz saját eseményeket generál. Minden generált esemény tartalmazza, hogy melyik objektum hatására váltódott

ki. Ezen információk együttesen segítenek a felhasználónak jobban megérteni egy adott alkalmazás felhasználói felületét.

#### A braille komponens

A braille komponens egy dinamikus, osztott objektum, mely inputját BML nyelven kapja. A BML a Baum által kifejlesztett XML dialektus. A modul képes kezelni számos braille kijelzőt, s az ő feladata az is, hogy a képernyőn megjelenő szövegeket a braille kijelzők számára elfogadható formátummá alakítsa át. A modul a Gnopi modul segítségével konfigurálható.

#### A nagyító

Lehetővé teszi a képernyő egy részének felnagyítását, s biztosítja a kommunikációt a felhasználó, s a felnagyított képernyőrészlet között. Tartalmaz egy fontnagyítási algoritmust is.

#### Beszéd

Ez a modul tartalmazza azokat a rutinokat, melyek segítségével a program képes kommunikálni a TTS motorokkal. A modul szintén a gnopi modul segítségével konfigurálható, ezzel lehetővé téve a felhasználónak, hogy számos TTS paramétert a saját ízlésének megfelelően állítson be.

#### A Libke komponens

A komponens felel a billentyűzet figyeléséért. A beérkezett üzeneteket stringekké konvertálja. Háromféle üzemmódban képes működni: betű, szó és úgynevezett auto mód. A modul szintén konfigurálható a gconf segítségével vagy a gnopi modullal.

#### Gnopi komponens

A Gnopi modul tartalmazza a Gnopernicus program felhasználói interfészét, mellyel a különféle komponensek egymástól függetlenül konfigurálhatóak, illetve ez a modul indul el, amikor a Gnopernicus programot futtatjuk.

#### BrlMon komponens

Ez a braille nyomkövető illetve tesztelő modul. A braille eszközről érkezett információkat tárolja karakteres formában. Kliens / szerver programként működik. Az SRCore komponenssel tartja a kapcsolatot, ahol a Core a szerver, a Brlmon pedig a kliens.

Előadásom nagy területet fog át, ezért nem állt módomban az egyes területeket részletesen bemutatni. Célom nem a témával kapcsolatos összes fogalom, s technikai eredmény teljeskörű ismertetése volt, hanem egy szubjektív keresztmetszet megjelenítése a Linux világában jelenleg meglévő, látássérültek számára rendelkezésre álló módszerekről. Gyakran egy jól elkészített, fogyatékosként is használható szoftver nagyobb segítség, mint egy ténylegesen kézbevehető, a maga fizikai valóságában is létező segédeszköz, például egy fehérbot, beszélő karóra vagy egy világító nagyító. Ha valaki azt a célt tűzi ki maga elé, hogy fogyatékos emberek számára készít, vagy alakít át már meglévő alkalmazásokat, egy külön szakterület kell, hogy megismerjen. Tudni kell az érintettek szükségleteit, s alapos technikai felkészültségre is szükség van. Beszámolóm alapvetően azért készült, hogy felkeltsem azok érdeklődését, akik úgy érzik, hogy tudnak tenni valamit azért, hogy a látássérültek is felfedezhessék a Linux világot. Segíteni sokféleképpen lehet, ha valaki kedvet érez hozzá, kérem, keressen meg, szívesen szolgálok bővebb információkkal.