

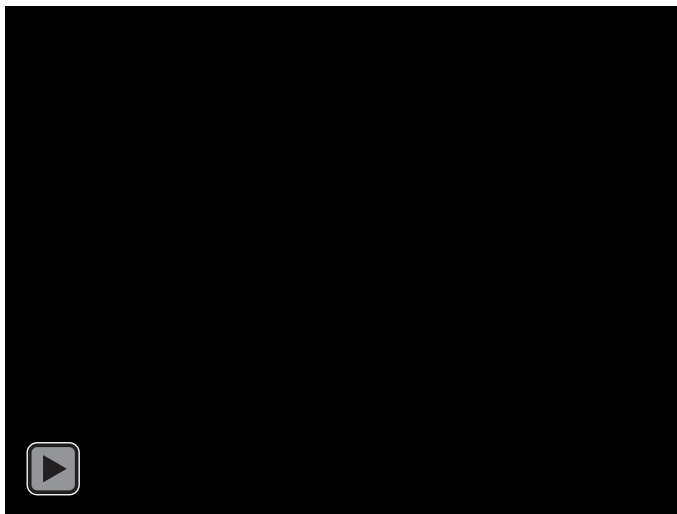
BESZÉDSZINTÉZIS VILÁGSZÍNVONALON

GORDOS GÉZÁVAL

Olaszy Gábor
BME TMiT

NJSZT Informatikatörténeti Fórum
2024. november 26.

Gordos Géza



Gordos Géza elindította a Beszédkutató Laboratórium szervezését a BME Vezetékes Híradástechnika Tanszék keretein belül 1970- ben.

SmartLab Intelligent Interactions

Laboratóriumunk magyar és idegennyelvű
gépi szövegfelolvasó fejlesztésével, az
ember-gép kapcsolat vizsgálatával és modern
gépi tanulóalgoritmusok kutatásával
foglalkozik.

A Gordos labor jogutódja 2024-ben
a BME TMI *SmartLab*
Intelligens Interakciók Laboratórium

Vezetője: Németh Géza



Kiss Gábor – Olasz Gábor 1980. A beszéd szintetizálás fejlődése és a számítógép szerepe. Mesterséges beszéd számítógéppel. I. Számítástechnika. 1980. október hó

Kiss Gábor – Olasz Gábor 1980. A gépi beszéd alkalmazási területei. Mesterséges beszéd számítógéppel. II. Számítástechnika. 1980. november hó

Kiss Gábor – Olasz Gábor 1980. Magyarul szól. Mesterséges beszéd számítógéppel. III. Számítástechnika. 1980. december hó

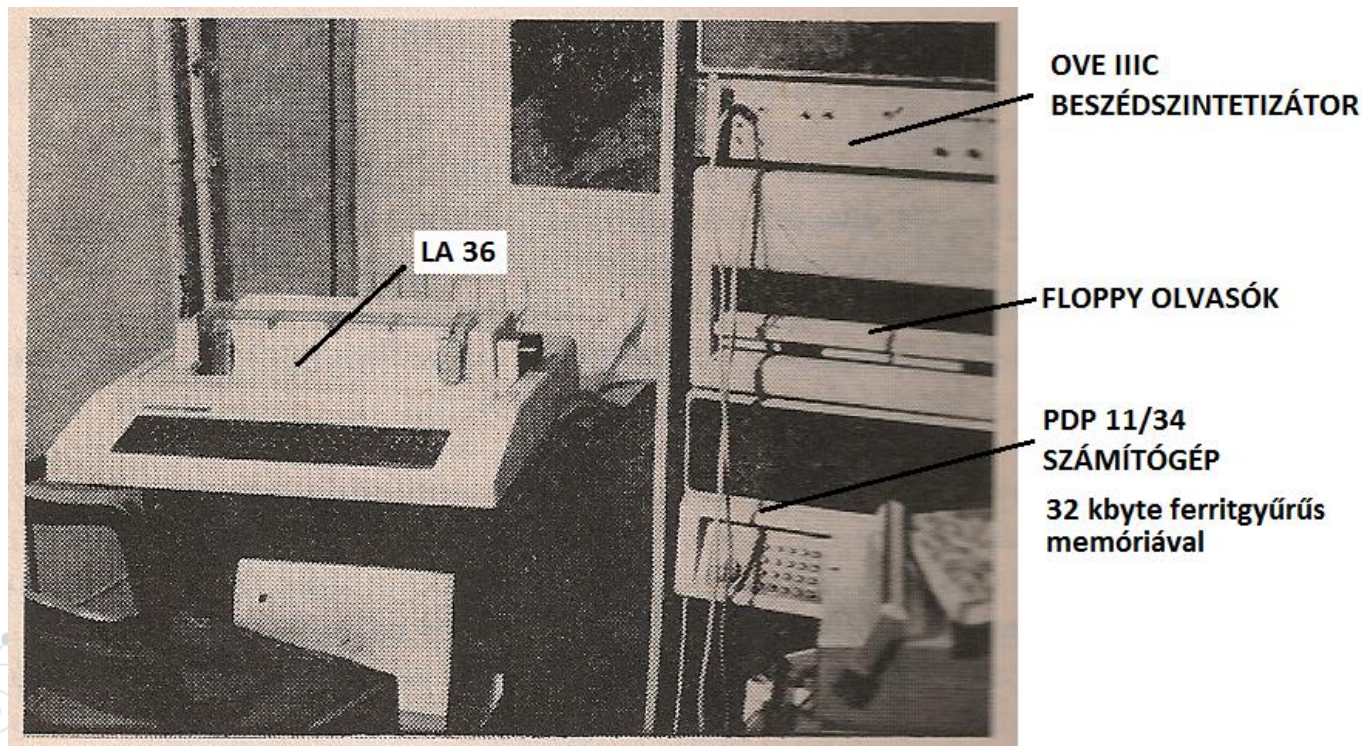
TALÁLKOZÁSUNK

1980

MTA Nyelvtudományi Intézet
Fonetikai Osztály és Laboratórium
Vezető: Bolla Kálmán

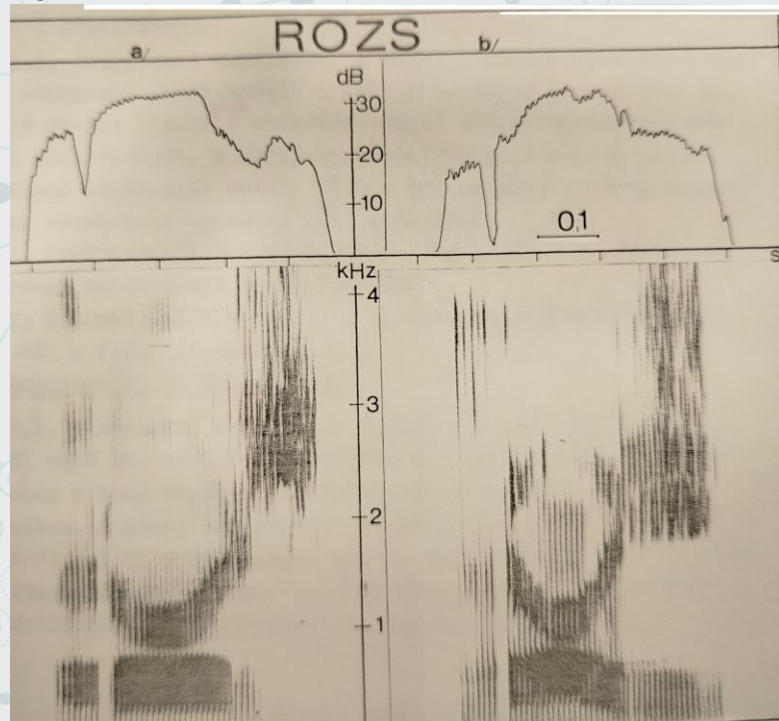
A FONETIKAI LABORATÓRIUM BESZÉDANALÍZIS SZINTÉZISSEL MÓDSZERE 1979

Fejlesztők: Olaszy Gábor, Kiss Gábor, Nikléczy Péter



Mottó: ha meg is szólaltatjuk beszédszintézissel a beszédanalízis eredményét, akkor tudjuk hallással igazán ellenőrizni, hogy jó-e az analízisünk minden részlete

1979



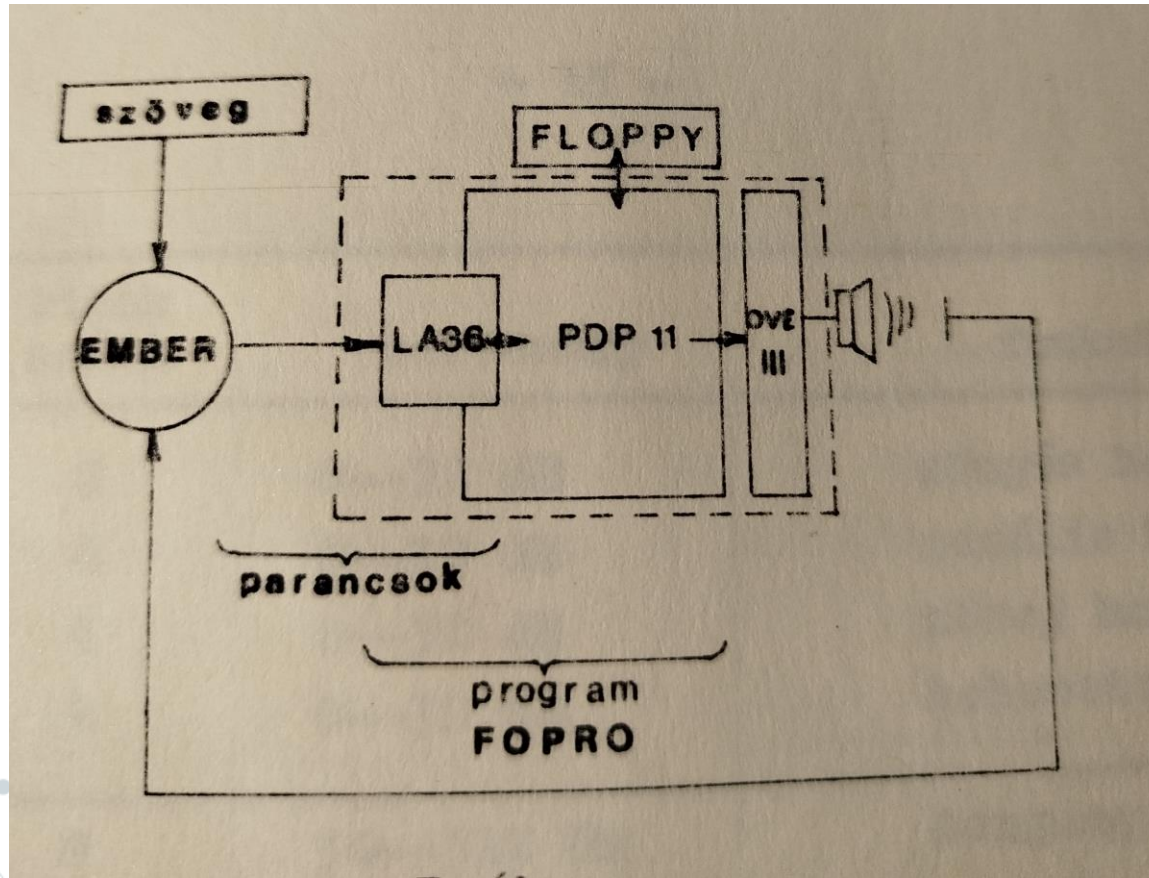
ROSZAH	SZINTETIZÁLÁS															81 VIII
DATUM:																
AZON, IDO	IDA	+ AD	FO	AN	+ N1	F1	B1	F2	B2	F3	B3	+ AC	AK	K1	K2	
RE1	40	10	0- -2	116		224- -267	96	1307	93	2851	500					
RE2	30	10	2- -4	116		291- -356	96	1307	48- -155	2851	500					
RE3	40	5	4- -0	116		400	96- -36	1307	155- -62	2851	500- -250					
RD	70	8	4- -10	116		400- -462		1345- -898		2016- -2329						
*D1	50	5	10	116		462		898		2397						
QZS	70	8	12	116		436- -300		951- -1600		2329- -2016						
QZSV1	50	10	6- -4	116		300- -252		1796		2329		10- -20	16	1796	4032	
ZSV2	50	10	4- -2	116		252- -212		1796		2329		20	16	1849	4032	
ZSV3	40	10	4- -0	116		212		1796		2329		20- -0	16	1849	4032	

a szintetizálási 9 db hangszet adatai

természetes

szintetizált

A FONETIKAI LABORATÓRIUM INTERAKTÍV BESZÉDSZINTETIZÁLÓ RENDSZERE (INBERE)



EZ A TALÁLKOZÁS SOK ÉVTIZEDES KÖZÖS MUNKÁT EREDMÉNYEZETT A FONETIKAI LABORATÓRIUM ÉS A GORDOS BESZÉDKUTATÓ LABORATÓRIUM KÖZÖTT

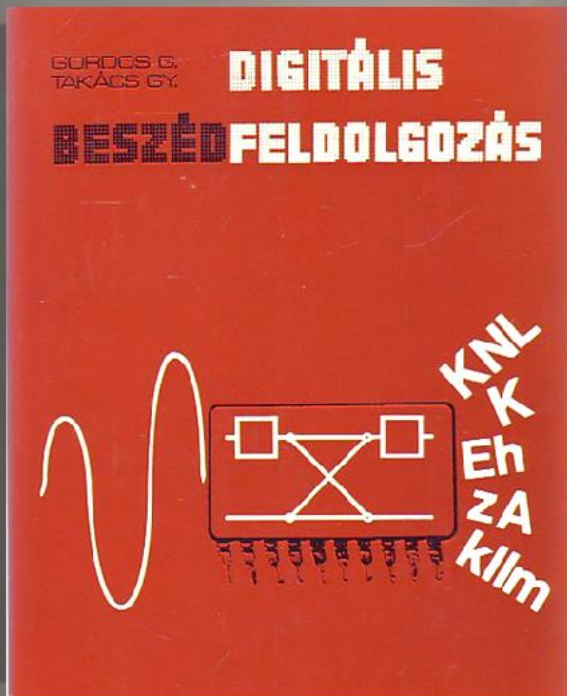


1. Gordos Géza javasolta, hogy adjak be doktori disszertációt a BME-re
2. Meghívott a BME-re dolgozni a beszédkutató csoportjába félállásban
3. A fonetikai laboratóriumban 1979-ben már működött az UniVoice magyar TTS rendszer első változata.
Adatbázisa 370 hangszelet adatait tartalmazta. Mérete: 1 kB.
A vezérlő programot Kiss Gábor írta Fortran nyelven

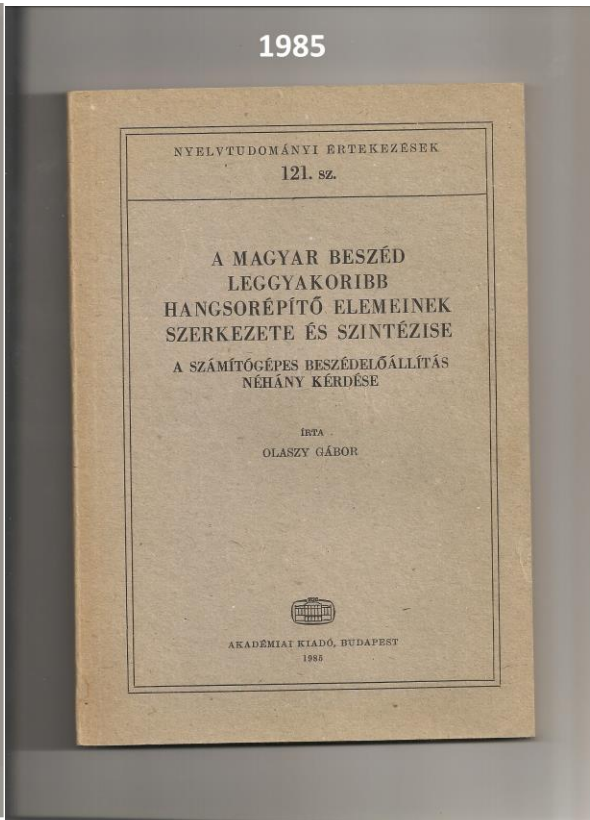


SZÁMÍTÓGÉPES BESZÉDKUTATÁS LENDÜLETET KAPOTT AZ 1980-90 -ES ÉVTIZEDBEN

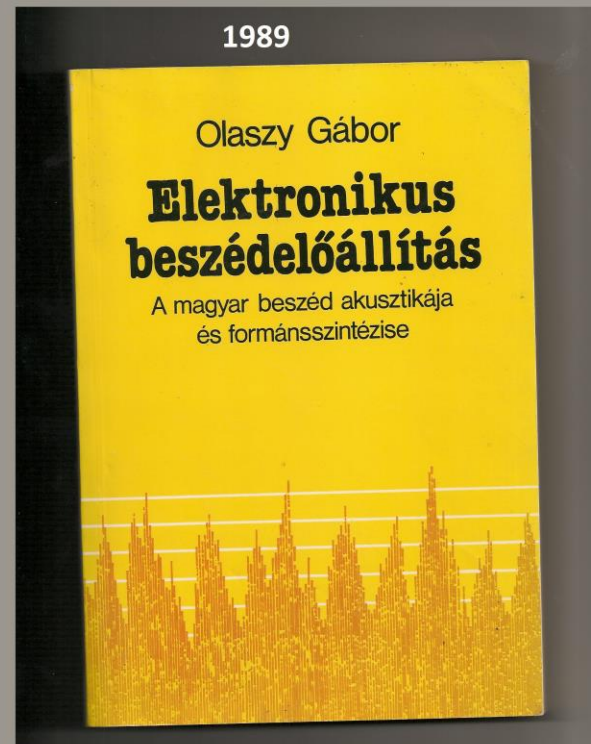
1983



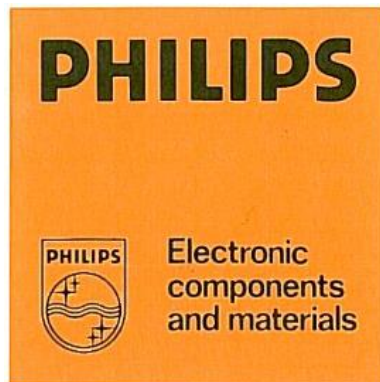
1985



1989

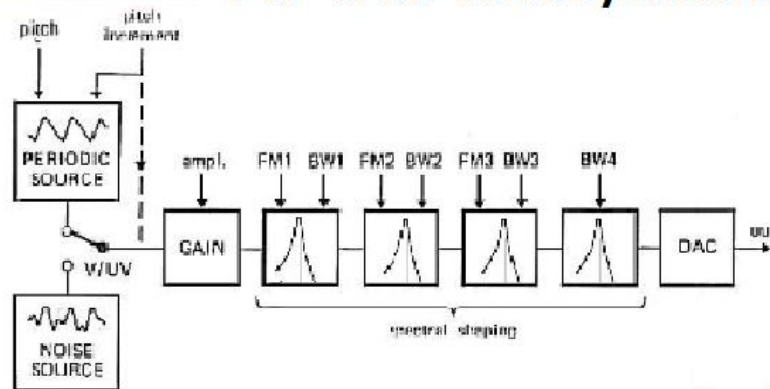


A FORMÁNSSZINTÉZIST FORRADALMASÍTOTTA A PHILIPS MEA8000 CHIPJÉNEK MEGJELENÉSE 1982-BEN



MEA 8000 voice synthesizer 1982

PCF 8200 voice synthesizer 1986



AZ ELSŐ MEA 8000 ALKALMAZÁS A GORDOS LABORBAN

1983-ban BME diplomaterv született a MEA 8000 felhasználásával

Németh Géza: VOXALARM - telefonos riasztó kötött szótáras beszédszintézissel
Konzulens: Takács György



BEINDULTAK A TOVÁBBI FEJLESZTÉSEK

Takács György fejlesztői rendszert tervezett a MEA 8000 alapú TTS hangszet adatbázisának elkészítésére SGS számítógépre

Multivox TTS development

Text <F3>: jo'napot .

Code <F4>: 1 22 4 4 4 4 20 3 3 12 4 4 14 1

	F1	B1	F2	B2	F3	B3	F4	B4	PI	FD	AM	FC	IT	Hertz	
1	335	39	2211	3000	2719	190	3253	700	-2.4	12	4	22	317	101.6	
2	335	20	1996	3000	2719	600	3253	700	0.0	12	7	22	317	101.6	c
3	335	39	2450	3000	2414	190	3253	265	0.0	12	8	22	317	101.6	c c
4	335	39	2328	3000	2719	600	2999	700	-2.4	12	8	22	317	99.2	c c
5	307	153	1801	800	2719	190	3528	700	0.0	12	10	22	319	99.2	c
6	307	153	1711	800	2719	190	3528	700	0.0	12	10	22	319	99.2	c c
7	335	153	1467	234	2414	190	3528	700	-2.4	12	9	22	319	96.8	c
8	365	153	1195	234	2414	190	3528	700	1.2	12	9	4	319	98.0	U
9	433	303	754	126	2719	190	3253	700	3.7	12	8	4	319	101.7	U
10	433	303	754	126	2719	190	3253	700	2.4	12	8	4	319	104.1	U
11	433	303	754	126	2719	600	3253	700	0.0	12	7	4	173	104.1	U
12	433	303	754	126	2719	600	3253	700	0.0	12	7	4	173	104.1	U
13	433	303	754	126	2719	600	3253	700	-2.4	12	7	4	173	101.7	U
14	433	303	754	126	2719	600	3253	700	0.0	12	7	4	173	101.7	U
15	433	303	754	126	2719	600	3253	700	0.0	12	7	4	173	101.7	U
16	433	303	754	126	2719	600	3253	700	-2.4	12	7	4	173	99.3	U
17	433	303	754	126	2719	600	3253	700	0.0	12	7	4	173	99.3	U

Jelek: Inton: Mondat 1

Item len=11 Buff len=59

<U>nice:MALE1 <D>ata:MALE <O>ut:wav Speed=? Rule editor <F5>

<A>mp=2 <144> dF1=0 dF2=0 dF3=0 dF4=0 Rule <↑ 1, 1+> = 490

1985-BEN ELKÉSZÜLT A SCRIPTOVOX TTS A BME BESZÉDLABORBAN A MEA 8000 BESZÉDSZINTETIZÁTORRA

Olaszy Gábor : hangszelet adatbázis, amely 255 különböző elemet tartalmazott a szöveg-beszéd átalakításhoz, valamint a prozódiai szabályok

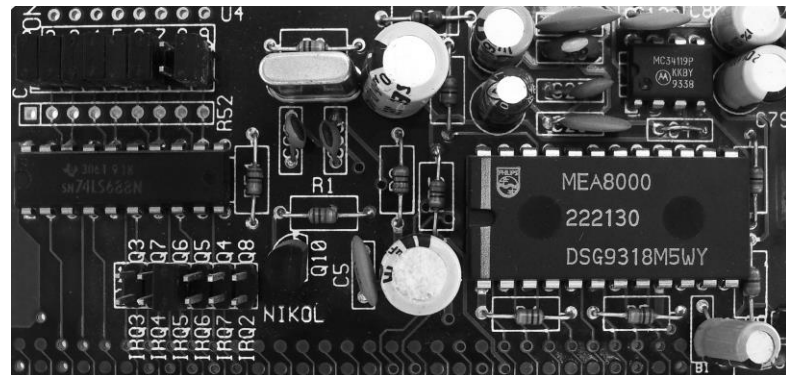
Podoletz György: ScriptoVox TTS program

Tihanyi Attila: áramkörtervezés és legyártás a demo alkalmazáshoz

Cikk:

Olaszy Gábor – Podoletz György.: A Scriptovox-MEA 8000 beszédelőállító rendszer felépítése és hangelemtára.

Kép és Hangtechnika 4. OPAKFI, Budapest, 1986.



MEGJÖTT AZ ÉTVÁGYUNK ÉS A COMMODORE 64 IS MEGJELENT A PIACON

Podoletz György átírta a ScriptoVox beszélő programot a Commodore 64-re
Tihanyi Attila elkészítette a Microvox-64 beszélő hardvert a Commodore géphez

ELKÉSZÜLT EGY KERESKEDELMI BESZÉLŐ SZÁMÍTÓGÉP ALKALMAZÁS

1985 Beszélő Commodore 64

Fejlesztő: BME

Gyártó és forgalmazó: Megamicro szövetkezet
(Podoletz György igazgató)

Hardver: Microvox-64 beszélő egység

Vezérlés: Commodore Basic program, 12 kByte

Hangkimenet: 3 pólusú dugó a TV-hez



European Conference on Speech Technology Edinburgh, Scotland, UK September 1987

ON THE SPEAKING MODULE OF AN AUTOMATIC READING MACHINE

G. Olaszy (1), G. Gordos (2)

ABSTRACT

The speaking module -- called SCRIPTOVOX -- of the automatic Hungarian reading machine was developed in the years 1983--86 by a four-member research team (3) of electrical engineers.

The speaking module using the general purpose, programmable MEA 8000 type integrated circuit for speech generation, converts any Hungarian text -- given in ASCII codes -- into good quality speech.

(Only the automatic Hungarian text-to-speech (TTS) conversion is discussed below, the grapheme-to-ASCII character converter -- developed at the Institute for Computer Research, Budapest -- is not discussed here).



A MULTIVOX, 8 NYELVŰ TTS RENDSZER FEJLESZTÉSE ÉS GYÁRTÁSA 1988-2002

Nyelvek: magyar, német, olasz, finn, holland, spanyol, eszperantó, arab

Nyelvi fejlesztés: Olasz Gábor és anyanyelvű segítők

Programozás: Mrena Rudolf

Hardver: Tihanyi Attila

Forgalmazó: Nikol Kkt.

Beszédszintetizátor: PCF 8200

Vezérlés: párhuzamos port

Vezérlő számítógép: PC –Windows

Bemutatás: CEBIT 1991-93 Hannover

Szervező: Gordos Géza

Díj: Software in Europe 1992

Az egyes nyelvek meghallgathatók: www.magyarbeszed.hu



CEBIT 1992, HANNOVER





A MULTIVOX SZABADALMI ÉRTÉKESÍTÉSE NÉMETORSZÁGBAN 1992

Németül beszélő képernyő felolvasó Apple Macintosh gépekhez

Know-how: BME

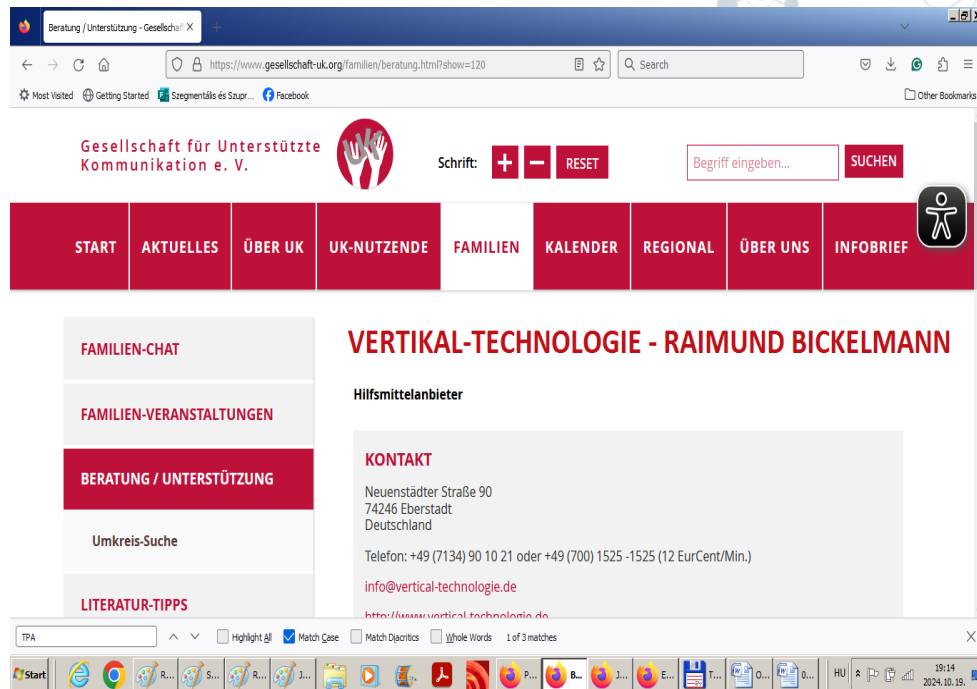
Vezérlő szoftver: Multivox-G program

Hardver tervező: Tihanyi Attila

Vezérlő hardver: TMS320 mikroprocesszor

TMS320 programozó: Osváth László

Forgalmazta: Vertikal-Technologie
Deutschland





A MULTIVOX SZABADALMI ÉRTÉKESÍTÉSE AUSZTRIÁBAN 1993

**Németül beszélő képernyő felolvasó
Amiga számítógépekhez
vakok segítésére**

Know-how: BME

Vezérlő szoftver: Multivox-G program

Hardver: TMS320 mikroprocesszor

Fejlesztő: Forschungszentrum Seibersdorf
Österreich

MULTIVOX - G
general purpose, phonetically based speech synthesis system
GERMAN VERSION 3.1

General system description
and
documentation

(C) TECHNICAL UNIVERSITY OF BUDAPEST, 1993

The MULTIVOX system is patented by the Technical University of Budapest. The system, in whole or in part, can be used outside of its intended application only by the written permission of the holder of the patent.
(c) Copyright: Technical University of Budapest
All rights reserved!

MULTIVOX-4, INGYENES, NEMZETI SZÖVEGFELOLVASÓ SZOFTVER 2002

CD MELLÉKLETEK/PCW/2002/12

Multivox 4

ingyenes!

A CD melléklet része a Multivox4 magyar nyelvű szövegfelolvasó szoftver, amely kis erőforrás-igényű (486, 33 MHz-től használható), és bárki által szabadon felhasználható. Le is tölthető a www.alpha.bme.ttt.pub/multivox4 webcímről. A program bármilyen alkalmazásban térítésmentesen felhasználható. Szabványos Microsoft SAPI 4.0 programozási felülettel látták el, és a legelterjedtebb operációs rendszereken alkalmazható (Win95, 98, NT, 2000, Me, XP és CE). A szintetizátor hangja minden hangkárttyával rendelkező PC-n megszólaltatható, egyéb kiegészítés nélkül. Az ingyenes magyar nyelvű szövegfelolvasó (32 különböző hangkarakterből lehet választani) széles körű érdeklődésre tarthat számot, segíti az alkalmazások kialakítását, a beszédet felhasználó technológiai kultúra fejlesztését. A programozható vezérlési felület alkalmas a programfejlesztőknek, hogy tetszőleges beszédalkalmazást készítsenek (játékokhoz, irodai és ipari alkalmazásokhoz, egyedi személyes szoftverekhez). Speciális területeken is jól használható, így például a mozgássérültek, a gyengénlátók és a beszédsérültek részére készített programok beszédkimenettel láthatók el. Fejlesztette: BME Távközlési és Telematikai Tanszék (1990–2000). Alapkutatás: MTA Nyelvtudományi Intézet, Kempelen Farkas Beszédkutató Laboratórium. A szoftver elkészítését az Informatikai Kormánybiztosság támogatta 2001–2002-ben. Jó szórakozást, jó programozást!

Magyar nyelvű beszéd szintetizátor mindenkinek



MISKA BÁCSI BESZÉL





ÁTTÉRÉS AZ EMBERI HANGALAPÚ HULLÁMFORMA-ÖSSZEFÜZÉSES TECHNOLÓGIÁRA

1995

Előny: természetesebb a hangzás

Hátrányok: bonyolultabb a TTS rendszer,

nagyobb memória igény kell

hangstúdió kell speciális szövegfelolvasással

hanghatár jelölés kell

periódus jelölés kell

minden beszélőre saját elembázis kell

speciális F0 módosítás kell a dallamhoz

speciális algoritmus kell a tempógyorsításhoz

A hangelembázis alapegysége: diád, triád, szó, mondat

Fejlesztői rendszer: interaktív hullámforma kezelés

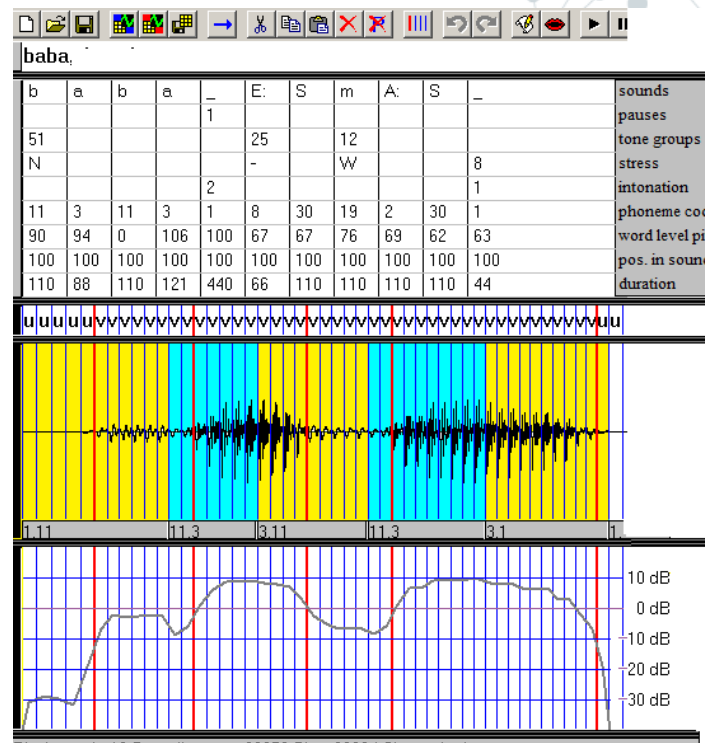
Fantázia név: **ProfiVox-diád, triád, korpusz**

Fejlesztői rendszer: Kiss Géza

TTS program: Olaszi Péter

Hangelem adatbázis, prozódia: Olaszky Gábor

Fejlesztésvezető: Gordos Géza, Németh Géza



ELKÉSZÜLT A PROFIVOX DIÁDOS TTS RENDSZER

1999



Rendező: BME TMIT

Elnök: Gordos Géza

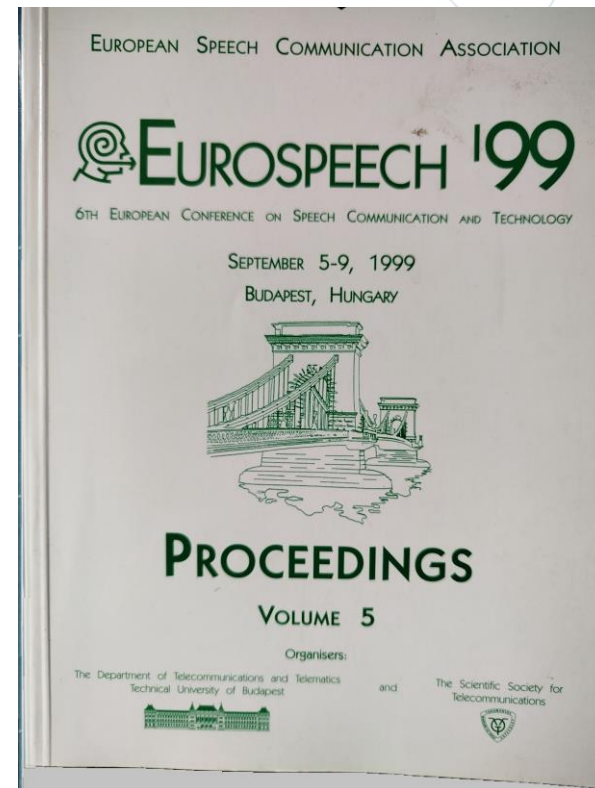
Titkár: Németh Géza

Előadások száma: 654

Résztevők száma: 1089

Magyar előadók száma: 15

[Interactive, TTS supported speech message composer for large, limited vocabulary, but open information systems](#)
[Gábor Olaszy](#), [Géza Németh](#), [Péter Olaszi](#), [Géza Gordos](#).
[interspeech 1999](#)





Készülék hírek ★

Tesztek ★

Alkalmazások ★

Tartozékok, kiegészítők ★

Tippek, tanácsok ★

Westel 900: bemutatkozik a Mailmondó

Kovács Attila - 1999. szeptember 7.

✕ Post

in Share



Szeptember 6-án a budapesti Eurospeech 99 konferencián a Westel 900 Rt. bemutatta Mailmondó nevű kísérleti e-mail felolvasó rendszerét, illetve az arra alapozott kísérleti szolgáltatást.

Sajtóhírek

SMSmondó: Egyedülálló magyar fejlesztésű szoftver westeles mobiltelefonokra

Budapest, 2003. október 8.

A Budapesti Műszaki és Gazdaságtudományi Egyetem Távközlési és Médiainformatikai Tanszéke, a M.I.T. Systems, a Nokia és a Westel együttműködésével a Symbian operációs rendszerű telefonokkal SMS-ek felolvasására képes új, magyar fejlesztésű szoftvert vehetnek igénybe a Westel ügyfelei. A szoftvert az arra alkalmas telefonokra díjmentesen telepítik a társaság szervizeiben. A készülék használója döntheti el, hogy olvasni vagy meghallgatni kívánja üzeneteit.

JAWS for Windows képernyőolvasó szoftver



Informatika a
Látássérültekért
Alapítvány

A JAWS for Windows képernyőolvasó program a billentyű leütések és a képernyőn látható üzenetek felolvasásával teszi lehetővé a számítógép használatát a látássérült emberek számára. A program kiemelten támogatja a Windows operációs rendszer, a Microsoft Office irodai programcsomag, több Internet böngészőprogram és sok más népszerű alkalmazás használatát.

AZ ELSŐ VOICE CHAT 2006

TELEFONOS
GYÓGYSZERINFORMÁCIÓS
ROBOT



RoboBraille.org/hu

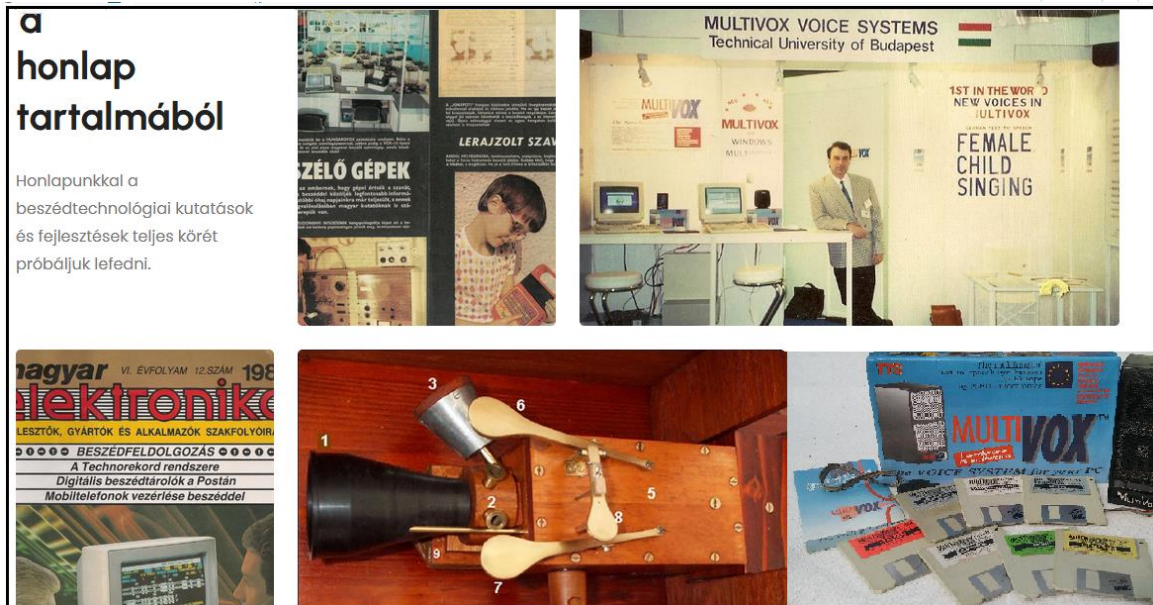


BME
TMIT



ProfiVox-diád, triád TTS egyéb alkalmazásai megtalálhatók és meghallgathatók az alábbi honlapon

A MAGYAR BESZÉDTECHNOLÓGIA MÚZEUMA



www.magyarbeszed.hu



Gordos Géza 1999-ben javasolta a BESZÉDINFORMÁCIÓS RENDSZEREK tantárgy tematikájának kidolgozását

Budapesti Műszaki Egyetem
Villamosmérnöki és Informatikai Kar

Villamosmérnöki Szak
Műszaki Informatika Szak
Választható tárgy

TANTÁRGY ADATLAP **és tárgykövetelmények** 1999. október.

Beszédinformációs rendszerek

..	Tantárgy kódja	Szemeszter	Követelmény	Kredit	Nyelv	Tárgyfélév
	VITMA 301	6	4+0+0v	5	magyar	1/1

A tantárgy tematikáját kidolgozta:

Név:	Beosztás:	Tanszék, Int.:
Dr. Gordos Géza	egy. tan.	TTT
Dr. Németh Géza	egy. adj.	TTT
Dr. Illényi András	tfmts.	TTT
Marosi Gyula	egy. ts.	TTT
Dr. Olaszy Gábor	egy. docens	TTT
Dr. Osváth László	egy. adj.	TTT
Tatai Péter	tmts.	TTT
Dr. Vicsi Klára	tfmts.	TTT

Állami kitüntetés 2004

Gordos Géza, Olasz Gábor és Németh Géza

a számítógépes beszédszintézisben elért kutatási eredményeikért

A MAGYAR KÖZTÁRSASÁGI ÉRDEMREND TISZTIKERESZTJE

kitüntetést kapta 2004. március 15-én

a Parlamentben



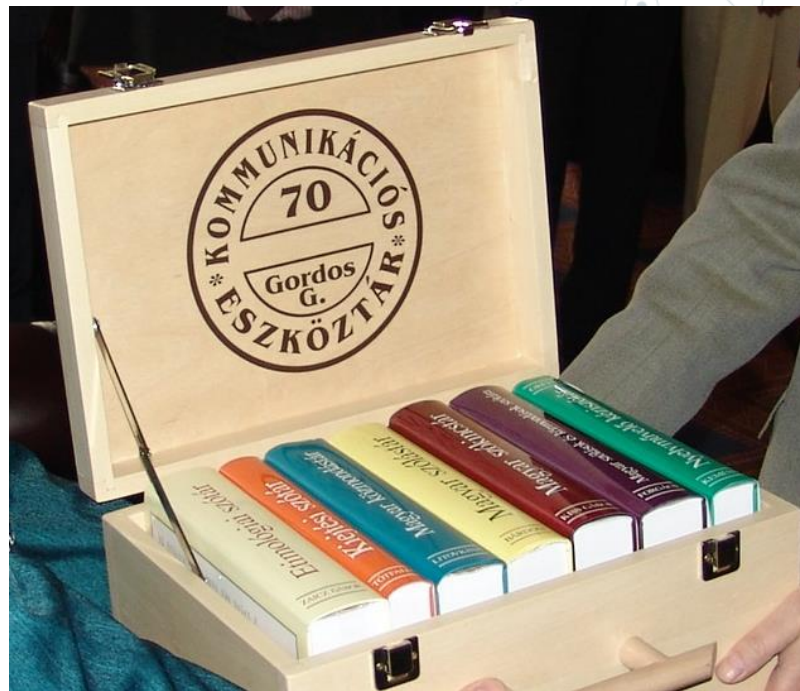
Gordos Géza 70 éves 2007



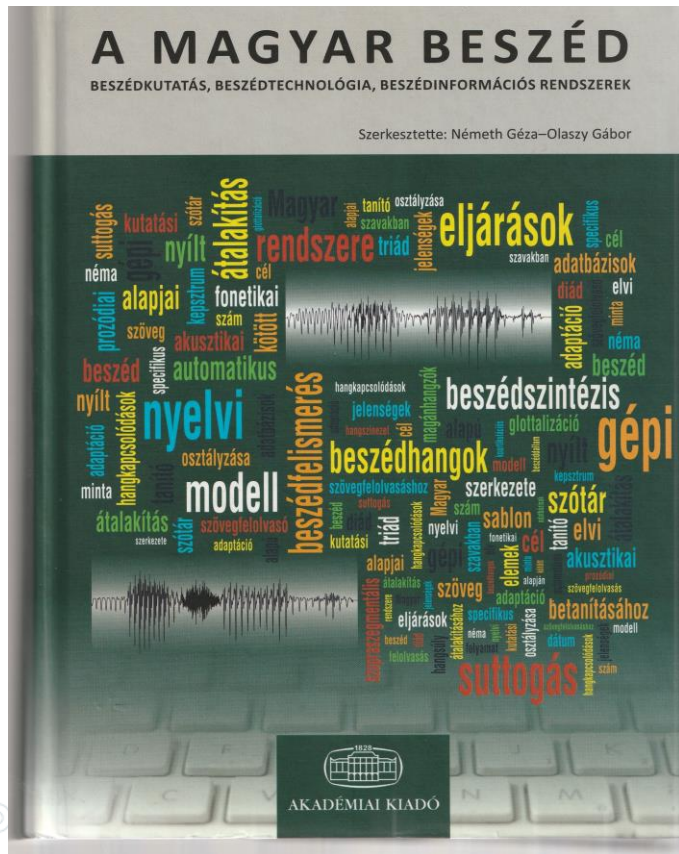
MTA Tudós Klub



Gordos Géza 70 éves



MEGJELENT A NAGY SZAKKÖNYVÜNK
MAGYAR BESZÉD 2010



Terjedelme: 708 oldal
Szerzők száma: 18



Áttekintő szerkesztő:



Gordos Géza (1937) a beszéd mérnöke, a műszaki indíttatású beszéd kutatás vezéralakja. A Budapesti Műszaki Egyetemen (BME) szerzett híradástechnikai szakos oklevelet 1960, dr. univ. 1966, kandidátus 1977, az MTA doktora 1995, dr. habil. 1995. 1960-tól a BME-n dolgozik. 1970-ben megalapította a Beszédfeldolgozási Laboratóriumot, amely 2010-ben is meghatározó szereppel bír a hazai beszéd kutatásban és az alkalmazások fejlesztésében. Számos szakmai kitüntetés tulajdonosa, nemzetközileg elismert kutató. Kutatási területei: távközlés, beszédkódolás, beszédszintézis, beszédfelismerés.

Szerkesztők és szerzők:



Németh Géza (1959) villamosmérnök, híradástechnikai szakmérnök. A BME Villamosmérnöki Kar Híradástechnikai Szakán végzett (1983), a BME-n doktorált (dr. univ. 1987, PhD 1997). A BME TMIT Beszédtechnológiai Laboratórium vezetője. Kutatási területei: beszédtechnológia, szolgáltatásautomatizálás, többnyelvű beszéd- és multimodális információs rendszerek, mobil felhasználói felületek és alkalmazások. Egyik fejlesztője a ProfiVox magyar szövegfelolvasó szoftvernek, valamint számos mobilos beszédinformációs rendszernek.



Olasz Gábor (1943) villamosmérnök, fonetikus. A BME Villamosmérnöki Kar Híradástechnikai Szakán végzett (1967), a BME-n doktorált (1985), a nyelvtudomány kandidátusa (1988), az MTA doktora (2003), habilitált (2004). Kutatási területei: a beszéd akusztikai szerkezete, fonetikai modellezés, szöveg-beszéd átalakító rendszerek tervezése, készítése, tesztelése. Egyik fejlesztője a ProfiVox magyar szövegfelolvasó szoftvernek, valamint számos beszéd kutatási honlapnak.

ÚJ IRÁNY A BESZÉDSZINTÉZISBEN: PROFIVOX – KORPUSZ 2010-2024

Utastájékoztató a MÁV állomásokon 2014-től



Előny: **természetes, emberi hangzás**

Hátrányok: nagy beszédatbázis

sok GB memória

speciális kereső függvények kellenek

hangstúdió kell 40 óranyi felolvasáshoz

hanghatár jelölés kell

periódus jelölés kell

minden beszélőre saját felolvasás kell

adott szűk témakörre ad természetes hangot

A hangelembázis alapegysége: **szó, mondat**

Fantázia név: ProfiVox-korpusz

Robot asszisztens szolgáltatás 2022-től

MÁV ELVIRA ROBOT BELFÖLDI UTAZÁSTERVEZÉSHEZ

06 20 499 4999

**Éjjel-nappal
hívható**



**beszélgető
robot**



Fejlesztő: Zainkó Csaba

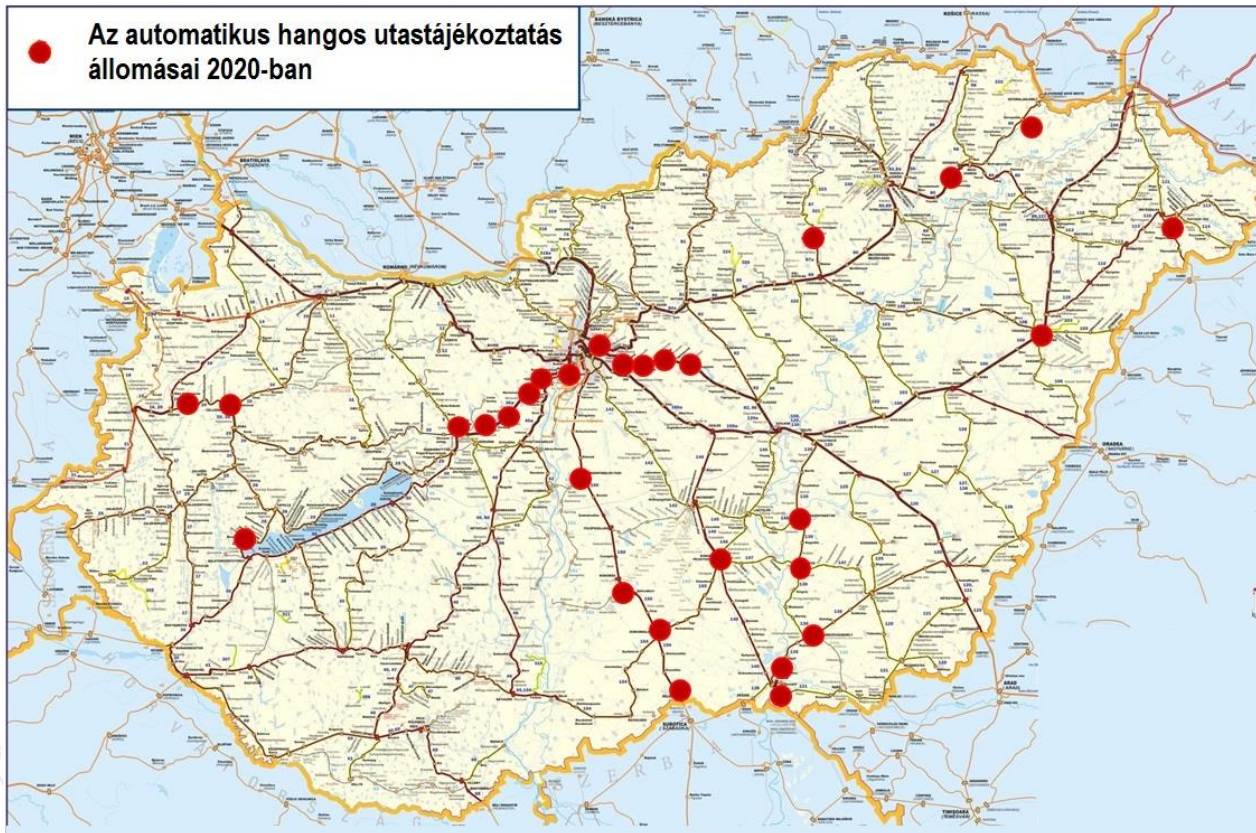
Informatikai karbantartó: Zainkó Csaba

TTS program: Zainkó Csaba

Beszédatbázis: Olaszy Gábor, Bartalis Mátyás

Fejlesztésvezető: Németh Géza

2020-BAN MÁR SOK MÁV ÁLLOMÁSON ÜZEMELT A BME UTASTÁJÉKOZTATÓ RENDSZERE



KORSZAKVÁLTÁS A BESZÉDSZINTÉZIS TECHNOLÓGIÁBAN. STATISZTIKAI PARAMETRIKUS ÉS NEURÁLIS HÁLÓS MEGOLDÁSOK 2010-2024

A Profivox HMM megoldás statisztikai elvű gépi tanuláson alapul és rejtett Markov-modelleket használ a szintetizálendő beszédjelet reprezentáló paraméterek létrehozására. Ezt a számítástechnika fejlődése tette lehetővé. Nem közvetlenül a hullámformával dolgozik, hanem a hullámformából spektrális és prozódiai jellemzők sokaságát nyeri ki. Komoly fonetikai, nyelvészeti tudásra nincs szükség.

A módszer kidolgozója: Gyires Tóth Bálint
Csapó Tamás

Előnyök: személyhez közeli hangzás, sokféle hang létrehozható
Hátrányok: nagy erőforrást és tudást igényel,
FEKETE DOBOZKÉNT MŰKÖDIK



NEURAL PROFIVOX

2020-2024

A jövő technológiája a mély neurális hálózatok alkalmazása a gépi beszéd előállítására. A természetes beszéddel összetéveszthető. A különböző modellek lehetőséget biztosítanak az alapfrekvencia és a sebesség állítására is, így teljes értékű beszédszintetizátorok motorjaként szolgálhatnak. A modellek az adott beszélő hangját élethűen visszaadják, de a modellparaméterek kombinálásával új beszélők hangjai is előállíthatók.

2022-re a BME SmartLab csapata többféle modellt is kidolgozott, amelyek nagyon jó minőségű szintetizált beszédet szolgáltatnak. Ez lett a Neural ProfiVox beszédszintetizátor.

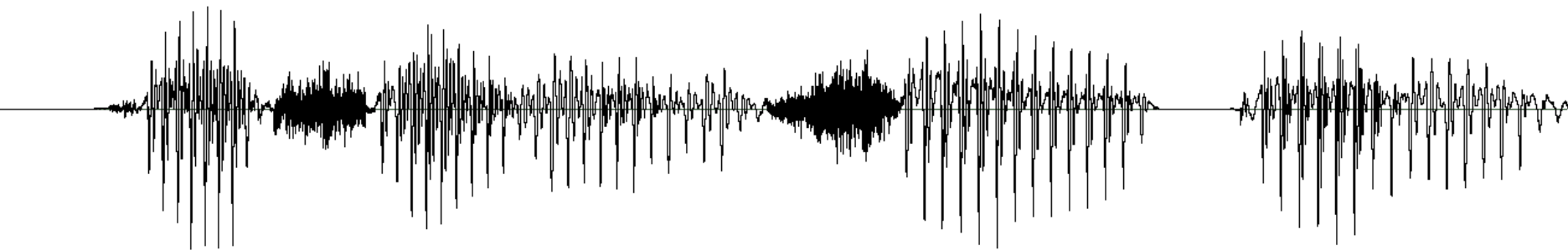
A fejlesztés vezetője **Dr. Zainkó Csaba**.

Előnyök: teljesen emberi hangon szól, személyhez köthető, sokféle hang létrehozható

Hátrányok: nagy erőforrást és komoly tudást igényel
FEKETE DOBOZKÉNT MŰKÖDIK







K Ö S Z Ö N Ö M S Z É P E N

