

# A gépi beszédfelismerés kezdeti kutatásai

Vicsi Klára

BME TMIT Beszédakusztikai Laboratórium

# Első beszédfelismerő Magyarországon - 1971

Készítő: Prof Tarnóczy Tamás - MTA Akusztikai Kutatólaboratórium

Analóg erősítőkből és szűrősorból álló berendezés

Folyamatos lassú szöveg bemondásakor az **egyes magánhangzókat mutató** lámpasor elemeiből az a lámpa gyulladt ki, amelyik magánhangzó éppen elhangzott.

A berendezést a Budapesten megrendezésre került 7th ICA kongresszuson mutatták be.

# MTA TTK BÉKÉSY GYÖRGY AKUSZTIKAI KUTATÓLABORATÓRIUM

(Budapest, XI., Budaörsi út 45.)

Vezető: Dr. Illényi András, PhD (1981-1995)

Munkatársak:

*Fehér köpenyben Dr. Illényi András,  
Horváth Attila műszerész,  
Körmöczy László technikus,  
Dániel István tud.s. munk.,  
Berényi Péter tud.s. munk.,  
Török Kálmán műszerész,  
Vicsi Klára tud.munk.,  
Kosik László elektromérnök,  
Illényi András mellett  
egy fizikus kolléga*



Az 1980-as évek közepétől kezdve a beszéd kutatás végig kiemelt és előremutató témája volt a Kutatólaboratóriumnak, Vicsi Klára vezetésével. Központi téma a kötött szótáras gépi beszéd felismerés volt.

## 1985-ben elkészült az első magyar kötött szótáras beszéd felismerő rendszer.



*Vicsi Klára  
tudományos munkatárs*



*Tihanyi László  
matematikus, szoftverfejlesztő*



# A kész termék

Személyfüggő, kötött szótáras berendezés.

Szótárkészlet:

100 szabadon választható szó vagy tőmondat.

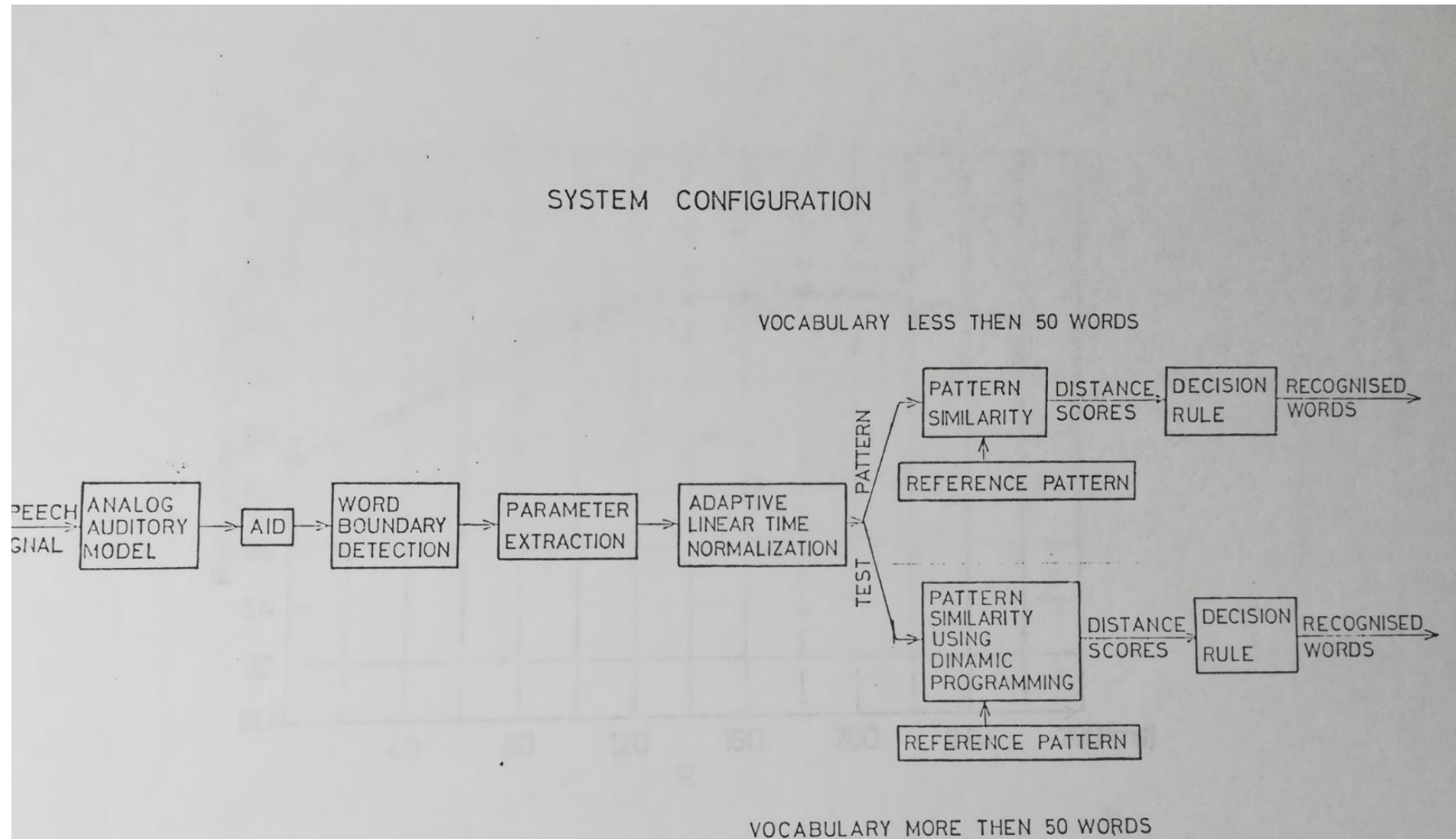
Többfelhasználós rendszer:

egyidőben a felhasználók száma 10.



A műszer a Budapesti Elektroakusztikai Gyár megrendelésére készült

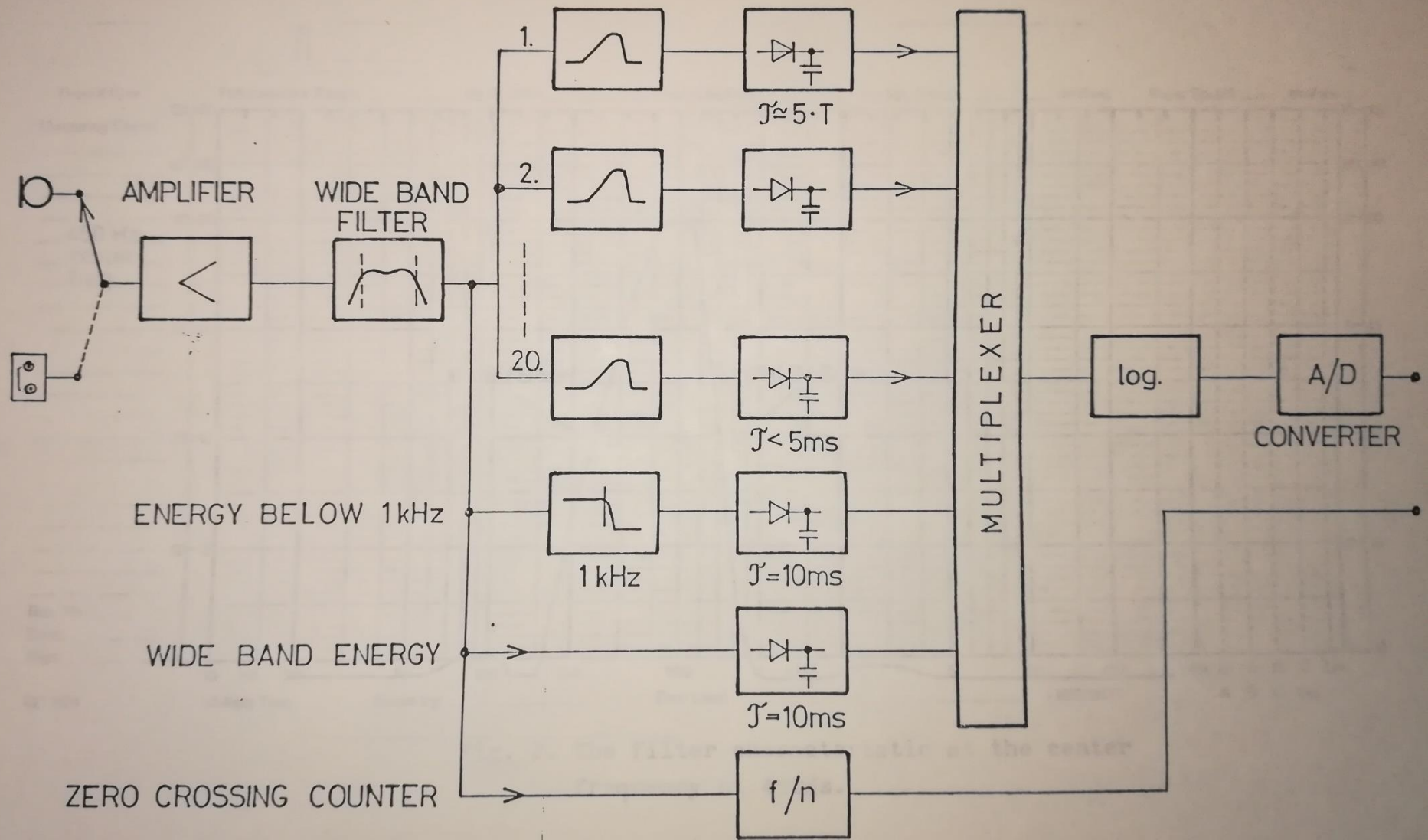
# A rendszer felépítése



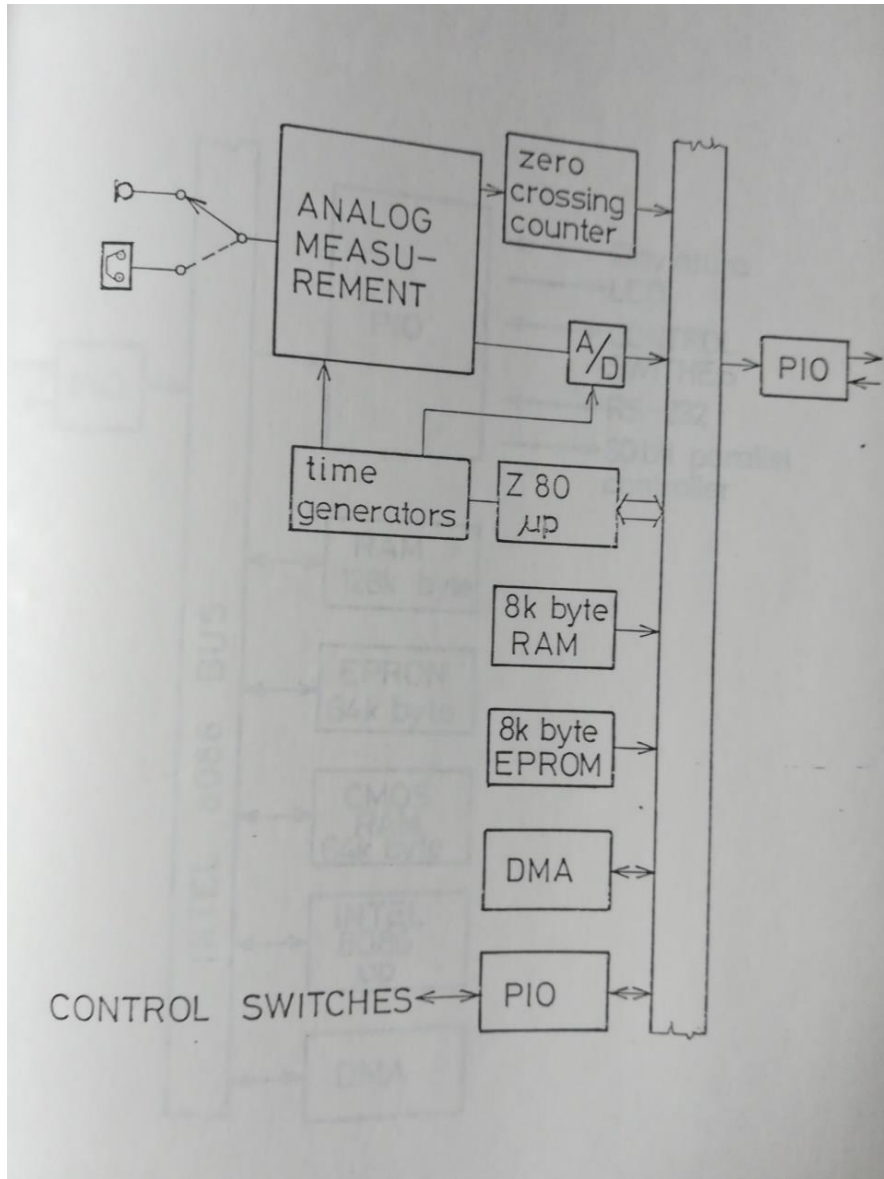
# ANALOG PART OF THE PREPRO- CESSOR

FREQUENCIES ANALYSIS  
IN 1 BARK

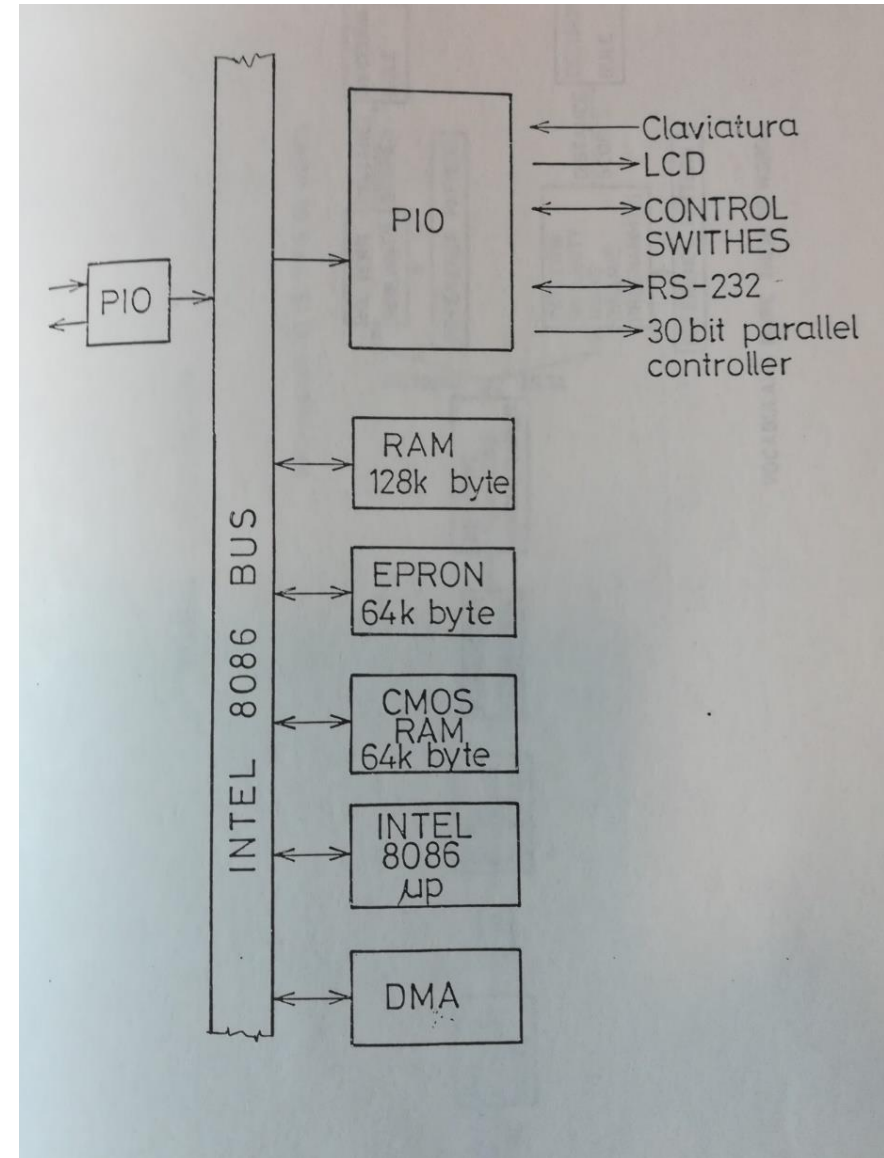
RECTIFF



## A preprocessor digitális része

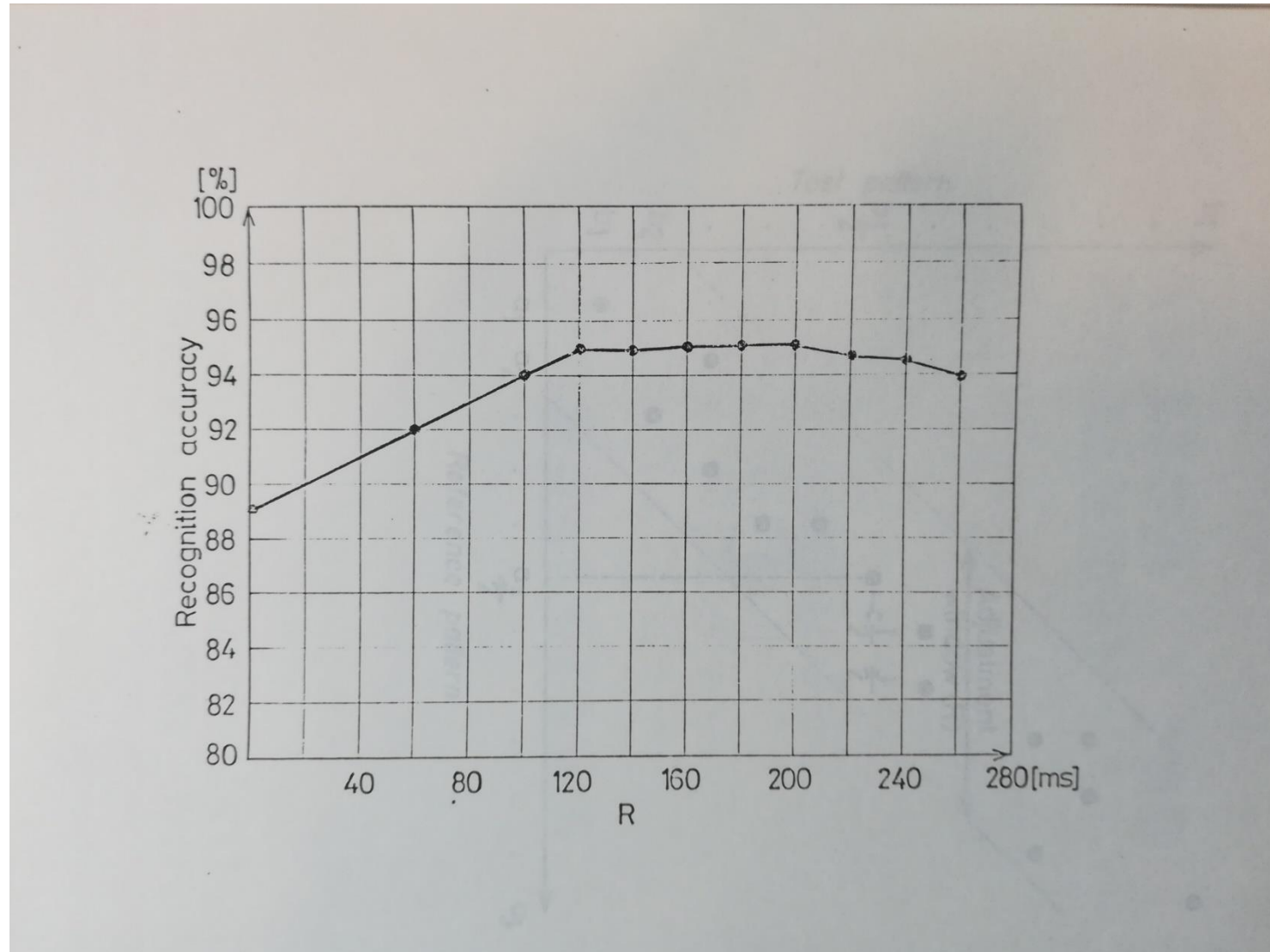


## A beépített számítógép

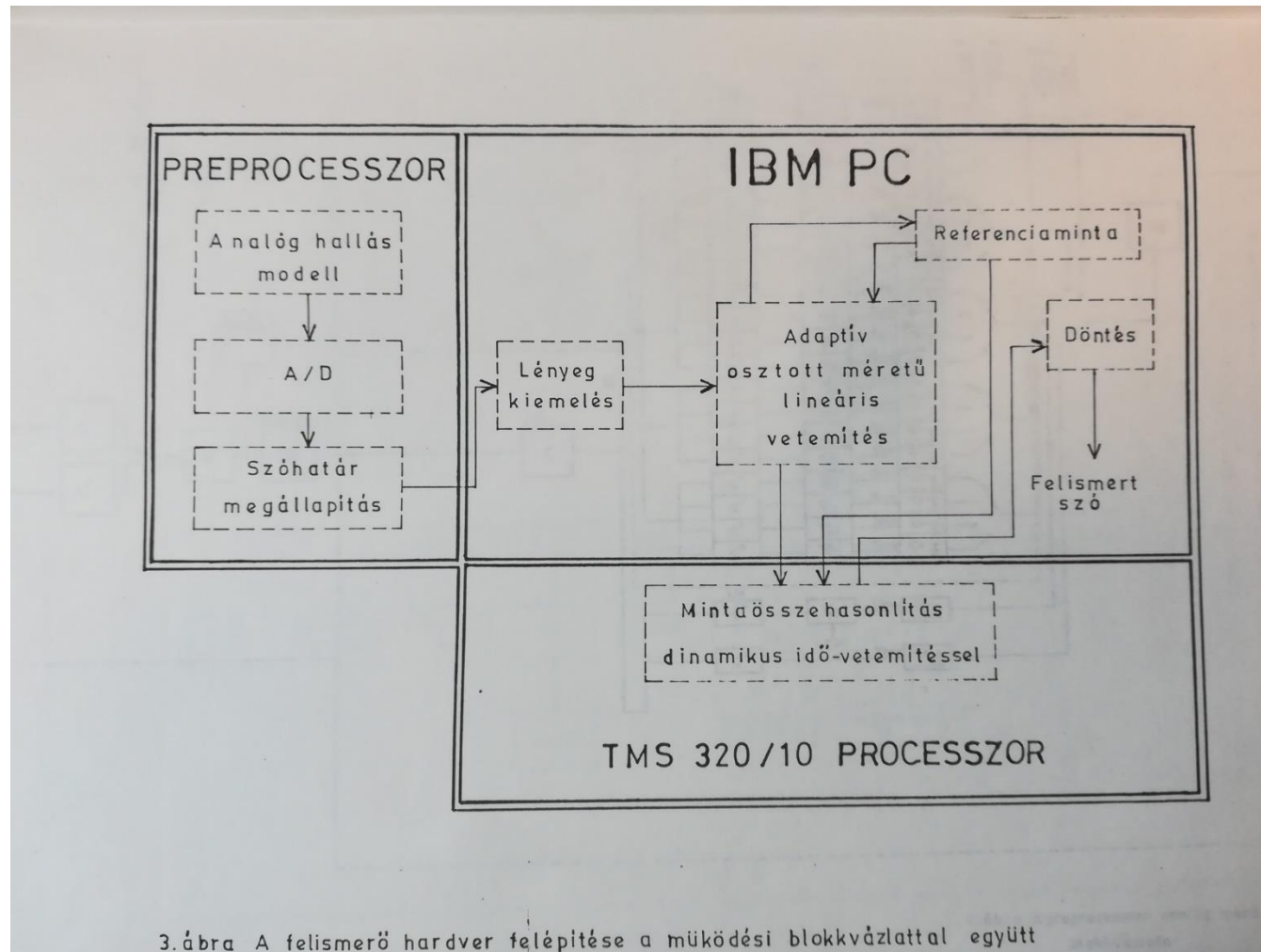




A készülék szófelismerési pontossága a szavak hosszának a függvényében



# 1986- ra elkészült a beszédfelismerő IBM XT/AT kompatibilis számítógépeken futó változata



3. ábra A felismerő hardver felépítése a működési blokkvázlattal együtt



MTA MŰSZERŰGYI ÉS  
MÉRÉSTECHNIKAI SZOLGÁLAT

Békésy György Akusztikai Kutatólaboratórium

## BESZÉDFELISMERŐ RENDSZEREK IBM XT/AT KOMPATIBILIS SZÁMÍTÓGÉPEKEN

Az MTA MMSZ Akusztikai Kutatólaboratóriumában kifejlesztett beszédfelismerő rendszerek lehetővé teszik, hogy személyfüggően tetszőlegesen választott szavakkal és egyszerű mondatokkal számítógépeket, robotokat vagy egyéb berendezéseket lehessen vezérelni. Minden olyan területen igen hatékonyan alkalmazhatóak, ahol a berendezések szóval történő vezérése a kezek felszabadítását teszi szükségessé más tevékenység elvégzésére. Testi fogyatékosok, betegek, akik kezükkel nem tudják a megszokott módon a környezetükben lévő berendezéseket irányítani, szóval történő vezérlés alkalmazásával több hasznos tevékenységek elvégzésére is képessé válnak.

### Típusos felhasználási területek:

#### GYÁRAK, ÜZEMEK:

- automaták, robotok, egyéb berendezések szóban történő irányítása
- Funkciók: adatgyűjtés, folyamatszabályozás, vezérlés, minőségellenőrzés

#### HIVATALOK:

- Billentyűzet részben vagy egészben történő kiváltása.
- CAD/CAM rendszereknél, térképkészítésnél

#### POSTA:

- Levelek és csomagok szortírozása

#### KUTATÓ-FEJLESZTŐ INTÉZETEK:

- magasszintű ember-gép kapcsolat

#### TESTI FOGYATÉKOSOK:

- környezetben lévő berendezések vezérése, pl. telefon, rádió

#### EGYÉB:

- leltárkészítés, megrendelésfelvétel, cím és üzenettrójtás, stb.

### Kisszótár kezelő rendszer (ST-03) /specifikáció/

Szótárkészlet: szabadon választható, változtatható szavak és tömodulok, nyelvfüggetlen

Megengedett bemondási hossz: 2 s

Bemondások közötti minimális szünet: 0,5 s

Felismerési pontosság: 95-99% (szótármérettől és jel-zaj viszonytól függ; pontosság > 95% ha a szótárméret < 80 és a jel-zaj viszony > 20 dB)

Felismerési idő: < 500 ms

Többfelhasználós rendszer: felhasználók száma egyidőben max. 10

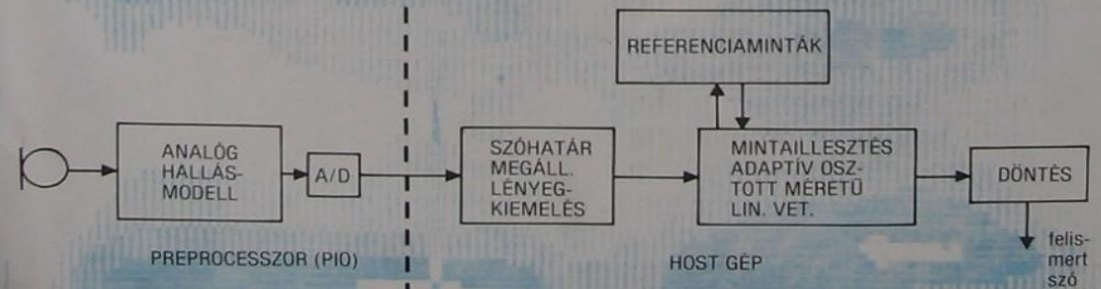
Előfeldolgozó egység: mikrofon bemenettel egy IBM kártyán

Felismerő rendszer: működik minden IBM XT/AT kompatibilis számítógépen

VEDELEM TEVESZTÉSEK ESETÉN!

A preprocesszor a mikrofonnal felvett jel akusztikus elemzését 19 párhuzamos szűrőcsatornára alapuló hallásmodell segítségével végzi. A 10 ms időintervallumban mintavett paraméterek A/D átalakítás után a „párhuzamos porton” keresztül a HOST gépbe kerülnek, ahol a további feldolgozás és a felismerés történik. A mintaillesztés adaptív időcsoportokban lineáris idővetítéssel történik.

### A RENSZER FELÉPÍTÉSE:



A kiállításokra készült  
„korabeli” prospektus

# Beszéd felismerő bemutatása a Budapesti Akusztikai Konferencián

1986 –ban a Központi Fizikai Kutatóintézet által készített robot vezérlésére





# Azóta, sok szép eredmény született Magyarországon

A BME TTT-n izolált szavas felismerő készült,

- 
- 

Takács György stokholmi KTH béli tartózkodása során neurális háló alapú folyamatos szöveget lejegyző rendszert készített.

- 
- 

BME TMIT-n Fegyó Tíbor-Mihajlik Péter HMM alapú folyamatos szöveget lejegyző rendszereket állított és állítanak elő  
a mai napig is,

- 
- 
- 

stb. (a lista nem teljes).

## 32 év után ma

Google által fejlesztett kereső-, diktálórendszerek  
okostelefonokon magyar nyelven is jól működnek.

Igaz, hogy még a mondathatárokkal bajok vannak,

de

Szaszák György a Beszédakusztikai laboratórium munkatársa magyar nyelvű prozódia felismerő rendszerével a mondathatárok felismerése már lehetséges, és beilleszthetővé vált az általános tartalomfelismerő rendszerekbe.

Köszönöm a figyelmet

[viczi@tmit.bme.hu](mailto:viczi@tmit.bme.hu)

<http://alpha.tmit.bme.hu/speech/>