

Audió interfészek

Beszéd: beszédhang sorozatokból áll. A zöngés hangok – zöngétlen hangok. Zöngés hangoknak van alapfrekvenciája, férfiaknál kb. 100 Hz, nőknél, kb. 200 Hz, gyerekek 300-400 Hz is lehet. A beszéd kb 10kHz-ig tart, a lényegi energia ebbe a részben található. Az alap telefonos átvitel csak az érthetőség szempontjából leglényegesebb részeket viszi át, max 4 kHz-ig. Modern beszédkódolók már magasabb frekvenciát is átvisznek. Tipikus a 16KHz-es mintavételi frekvencia a beszédfeldolgozásban.

Hallás: Egyenlő hangosságú görbék megadják, milyen energiájú jeleket hallunk azonosnak, ha a frekvenciájuk különbözik. Léteznek elfedési jelenségek, időbeli és frekvenciabeli. Az időbelinél az maszkoló hang előtt és utáni hangokat bizonyos ideig nem halljuk meg. Frekvenciabeli elfedésnél egy erős adott frekvenciájú hang mellett közeli gyengébb jeleket nem hallunk meg. A hallás irányfüggő, fej mozzgatása esetén az iránymeghatározás pontosabb az embernél.

Beszédszintetizátor: Szövegből beszédet állít elő.

Beszédfelismerő: Beszédből szöveget állít elő. Általában beszélőfüggetlen, tehát bárki beszédét képes felismerni. Lehet kötött szótáras, ahol adott, hogy mit ismerhet fel, vagy kötetlen, ahol szinte bármit felismerhet. A felismerést a témakör meghatározza, a nyelvi modell adja meg mit tud felismerni. Mindig a felhasználáshoz kell igazítani a beszédfelismerőt.

Akusztikai arculat: Hasonló a vizuális arculathoz, meghatározza, hogy mikor milyen zene, hanghatás vagy pl. köszönés szükséges. Pl. a telefonos tartászene kiválasztása.