

A SZAKMAI MEGVALÓSÍTÁS RÉSZLETES ÜTEMEZÉSE

Részprojekt neve:		Az ember-gép kommunikáció multimodális adatbázisai: adatok és interpretáció	Időszak
1. feladat megnevezése:		Megfelelő nagyságú és eloszlású spontán beszéd adatbázis gyűjtése és annotálása	1. év
Ssz.	Részfeladat megnevezése	Szakmai tartalom az adott munkaszakaszban	Negyedév
1.1	Nyelvi adatgyűjtést és feldolgozást támogató hardver- és szoftverkörnyezet összeállítása	A megvalósításhoz szükséges környezet összeállítása: beszédatadabázis rögzítéséhez, tárolásához, feldolgozásához, valamint a beszédfelismerő rendszer betanításához, futtatásához megfelelő hardver- és szoftverhátter kiépítése. Az annotáció sok kézi munkát igényel – a manuális tevékenységet segítő alkalmazások kifejlesztése.	1-2.
1.2	Adatbázis-kialakítási szempontok kidolgozása	A felismerő betanításához olyan beszéd- és szöveganyagot kell összeállítani, amely alkalmas az akusztikai-fonetikai és intonációs modellek, valamint a nyelvi modellek betanításához. Tehát a hanganyagoknak az adott nyelv hangkapcsolatait jól kell reprezentálni a bemondók és az adott feladat esetében, tartalmaznia kell az előfordulható ejtés- és intonációs variációkat. A beszédatadabázis tartalmának kialakításában figyelembe vesszük a vonatkozó európai szabványokat.	1.
1.3	Egységes modellépítési koncepció kidolgozása	Olyan annotálási szabályrendszer és adatfeldolgozási koncepció kialakítása, amely megoldást ad (a feliratozáson, az előforduló zajok annotálásán és a beszélők adatainak rögzítésén túl) az intonációs frázisok és egyéb vizsgálandó jelenségek (érzelmekek, diskurzusmarkerek) kezelésére.	2.
1.4	Audió- és videófelvevételek gyűjtése és annotálása	Az 1.2. pontban előkészített adatbázis kialakítási szempontok alapján felvételek készítése megfelelő akusztikai és hardver körülmények között, magyar nyelven. A rögzítés digitálisan wav és mts formátumban történik. A felvételek annotálásakor pontosan rögzítésre kerül a nyelv, a bemondó azonosítója, neme, a felvétel körülményei, valamint az adott nyelv ortografikus karaktereivel lejegyzendő az elhangzott szöveg, a szokásostól eltérő akusztikai esemény/zaj, valamint az intonációs, érzelmi és diskurzus-markerek. Eredménye min. 50 órányi annotált, magyar nyelvű hangfelvétel.	1-4.
	Konzultáció	A részprojekt folyamatos szakértői tanácsokkal való ellátása, a megfelelő adatáramlás biztosítása.	1-4.

2. feladat megnevezése:		Az adatbázis felhasználása a prozódiai jellemzők vizsgálatához-kutatásához	0,5-1,5. év
Ssz.	Részfeladat megnevezése	Szakmai tartalom az adott munkaszakaszban	Negyedév
2.1	Vizsgálatokat támogató szoftverkörnyezet összeállítása	Ebben a munkaszakaszban a (Praat) hangelemző program intonációs vizsgálatokra történő előkészítése és a statisztikai vizsgálatokhoz szükséges szoftverkörnyezet felállításának történetje meg.	3.
2.2	Prozódiai jellemzők vizsgálata az annotált audió és videó anyagon	A beszéd alaphangfrekvencia-, energia- és időviszonyainak vizsgálata statisztikai módszerekkel végezhető magyar nyelvű, legalább 50 órányi hanganyagot tartalmazó spontán beszéd adatbázison. A vizsgálatok a beágyazások, beékelések, iterációk, hezitálások, újrakezdések és backchannel-ek, illetve az érzelmekek és diskurzusjelölők prozódiai jellegzetességeire terjedhetnek ki, különös tekintettel a tonális folytonosság, a nem-folytonos tonális összeállítás és a tonális rekonstrukció elvére.	3-6.

2.3	Szabályalkotás, statisztikai modellezés	A statisztikai modellezéskor a betanítás során az adatbázis hangfájlaiból az előfeldolgozással nyert szupraszegmentális jellemző vektorok, valamint az adatbázis szegmentálási és címkézési adatai használhatók fel a prozódiai modellek felépítéséhez. A szabályalkotás beágyazott részek lokalizálásával magában foglalja a szintaxis pontosabb feltérképezésének kísérletét is. Lehetőség nyílik az audio és videó címkék közötti összefüggések vizsgálatára.	5-6.
	Konzultáció	A részprojekt folyamatos szakértői tanácsokkal való ellátása, a megfelelő adatáramlás biztosítása.	3-6.

3. feladat megnevezése:		Vizsgálati eredmények beépítése a beszédfelismerőbe, tesztelés	2. év
Ssz.	Részfeladat megnevezése	Szakmai tartalom az adott munkaszakaszban	Negyedév
3.1	Beszédfelismerő keretrendszer tervezése	Beszédfelismerő rendszerek tanulmányozása, tesztelése, a kutatás céljainak legmegfelelőbb keretrendszer kiválasztása, paraméterezése.	5-6.
3.2	Az intonációs modellek bekapcsolása a beszédfelismerő rendszerbe	Az új modul a rendszerbe illesztve annak vizsgálata következik, hogy milyen mértékben javítható a beszédfelismerés hatékonysága. A prozódiai szegmentálás ismeretében a hipotézis gráfok újraszűzhetőek, így a végeredmény kiértékelését már a prozódia alapján nyert információ is befolyásolja. Érzelemfelismerő és társalgási forduló felismerő modul tervezése.	6-7.
3.3	Integrálás, tesztelés	A részprojekt záró tevékenysége, a kapott eredmények folyamatos integrálása és tesztelése.	6-8.
	Konzultáció	A részprojekt folyamatos szakértői tanácsokkal való ellátása, a megfelelő adatáramlás biztosítása.	5-8.