

Összefoglaló a spontánbeszéd- kutató műhely 2010. január 22-i találkozásáról

A találkozó témája: *Spontán beszéd géppel?*

A találkozón Zainkó Csaba és Csapó Tamás mutatta be a kísérleti eredményeit a gépi spontán beszéd előállításról, és megvitatták a műhely tagjaival az eredményeket. Az alábbi összefoglaló az előadásról és a beszélgetésről készült hangfelvétel tartalmi kivonata.

A találkozót megelőzően a műhely tagjainak egy webes teszt állt rendelkezésre, ahol meghallgathatták és értékelhették az elkészült szintetizált mondatokat.

A beszélgetés első részében egy kis előadás keretében megismerhették a résztvevők a használt szintézistechnológia főbb elemeit, majd meghallgattuk a tesztben hallható mondatokat. A beszélgetésen a következő gondolatok fogalmazódtak meg:

- Spontán beszéd és a felolvasott beszéd közötti különbség személyfüggő. Egyeseknek a felolvasott beszéde rosszabb „minőségű”, mint másoknak a spontán beszéde. Például olvasott beszéd esetén is hangok redukciója jellemző egyes adatközlőkre. A felolvasás nehézséget jelenthet egyeseknek, megakadások és egyéb spontán attribútumok jelennek meg a felolvasott beszédben.
- Miért kell spontánnak lennie a gépi beszédnek?
 - Bizonyos helyzetben ne legyen spontán, idegesítő lehet.
 - Nem elvárt a géptől a spontán reakció.
 - Milyen helyzetekben van létjogosultsága a spontán beszédnek?
 - Ö-zés kerülendő
 - Autó, gép beszélgetni fog velünk egyszer. A mai korosztály nem biztos, hogy elfogadja ezt, de a következő generációk már lehet, hogy igen.
 - Egyedülállóknak fontos lehet a beszélgető gépi partner.
- Hosszabb felolvasáskor a gépi rendszerek nehezen tolerálhatóak. Spontán jegyek segíthetnek az elviselhetőbbé tenni a hallgatást. Emberi vagy művészi felolvasást is fárasztó hallgatni, a gép még rosszabb lehet. Színészek nem spontán, de érdekes módon olvasnak fel, így tartják ébren a figyelmet. Gépi rendszer akkor jó, ha élénk, dinamikus. Felolvasott beszédben a hangredukció többnyire elfogadható.
- A spontán beszéd és a természetes beszéd különböző lehet. Külön kell választani a két fogalmat.
- A korpuszos rendszerrel kapcsolatos észrevételek:
 - prozódiaileg más helyre került szavak kiugró prozódija zavaró,
 - olyan az érzetet kelt, mintha akcentusa lenne,
 - időszerkezete nem megfelelő, az időzítések fontosak, ritmus kell,
 - az időszerkezet-módosítás akusztikus minőségromlást okozhat,
 - adatbázis-bővítés javasolt, a hibatípusok megállapításával,
 - szavakat, mondatrészeket több prozódiai pozícióban kell rögzíteni.
- Az angol nyelvű rendszerek mások, a meghallgatott minták jónak tűnnek, de mivel nem vagyunk anyanyelvűek, igazából nem tudjuk megítélni a minőséget.
- Tesztmondatok értékelése:
 - a szintetizált ö-k rövidek, nagyon zavaróak,
 - az érthetőséget nehéz értékelni, mert a mondatok megegyeztek, így a tartalom a második mondattól már ismert volt,
 - a dialógusrendszerben való használata kérdés nehezen értelmezhető,
 - az emberi spontán beszéd meghallgatása után a „mennyire spontán” kérdés nem értelmezhető,

- o a tartalom is számít az értékelésben, nem csak a spontán jegyek.
- Továbbfejlesztési lehetőségek:
 - o fordított teszt: funkcióhoz választani mondatot. Pl.: adott szituációban gépi, spontán vagy köztes jellegű szintetizált beszéd az elfogadható,
 - o következő tesztben több mondat kell, variálni kell a tartalmat,
 - o nagy számú adatközlőre van szükség,
 - o lehetséges, hogy fiatalokat kell kérdezni,
 - o funkció-mondat párosítás jellegű teszt is kell,
 - o természetes – spontán fogalom szétválasztása,
 - o csak 1,5 – 2 mondat hosszú hangminták elegendőek,
 - o paraméterek variálása, pl. csak F0, csak szünetek, csak megakadások megvalósítása, illetve ezek kombinálása.