

Szintetizált beszéd természetesebbé tétele

Csapó Tamás Gábor IV. Info., csapszi@sch.bme.hu

**Konzulens: Dr. Németh Géza, Távközlési és Médiainformatikai Tanszék,
nemeth@tmit.bme.hu**

A mai életben egyre fontosabb szerepet játszik az ember-gép kapcsolat, hiszen a világ az információs társadalom létrejötte felé halad. Egyre több mindent várunk el a számítógépektől, hogy egyszerűbbé tegyék mindennapjainkat. Ebben a folyamatban a beszédtechnológia, ezen belül a beszéd-szintézis is fontos szerepet játszik.

Napjainkban a jó minőségű gépi beszéd-keltés egyik kulcskérdése a megfelelő, de egyben változatos prozódia (intonáció, hangsúlyozás, ritmus) megvalósítása. Ma már nem elég az, ha egy szabályrendszer segítségével a nyelvi elvárásoknak megfelelő, adott szöveghez mindig azonos prozódia-t rendelünk. A cél az, hogy a gépi megoldás abból a szempontból is hasonlítson a természeteshez, hogy ugyanazt a mondatot ne mindig ugyanúgy mondja, hanem legyen benne némi (sokszor az adott személyre jellemző) változatosság.

A dolgozat keretében először áttekintem a témával foglalkozó szakirodalmat. Saját kísérleteim keretében első lépésben azonos bemondó hasonló jellegű bemondásait elemzem, ezekből próbálok a mondatok dallamszerkezetére jellemző információt kinyerni. A későbbiekben több bemondó azonos szöveghez tartozó hamgmintáit is megvizsgálom. A munka célja mondattípusokhoz és személyekhez rendelhető olyan prozódiai minták kialakítása, melyek gépi felolvasó rendszerekbe beépíthetők.

Az ily módon természetesebb hangzásúvá tett szövegfelolvasó rendszer számos gyakorlati alkalmazásban használható, mint például SMS, email-felolvasó, tudakozó. A jobb minőséget főleg olyan helyzetekben lehetne kihasználni, ahol hosszú szövegeket kell felolvasni, hiszen ekkor jobban számít a prozódia.

Feldolgozott irodalom:

1. Olasz Gábor – Kovács Magdolna – Nikléczy Péter – Gósy Mária: Magyar nyelvi beszédtechnológiai alapismeretek. (600 oldal CD-ROM-on). Szerk.: Olasz Gábor Nikol Kiadó 2002. Budapest <http://alpha.tmit.bme.hu/pub/beszinf/start.html>
2. Gósy Mária: Fonetika, a beszéd tudománya, Osiris Kiadó, Budapest, 2004
3. Puppel, S. and Demenko, G., editors, Prosody 2000: Speech Recognition and Synthesis, Krakow, Poland. Adam Mickiewicz University