

Prozódiai jellemzők gépi feldolgozása és hasznosítása élőnyelvi korpuszok elemzésében

Szekrényes István

A disszertációm célja, hogy a beszélt nyelv kutatásához olyan újszerű módszertani eszközökkel és kísérletekkel járuljon hozzá, amelyek a beszéd prozódiaját nem csak a megnyilatkozások szintjén, hanem a társas interakciók mikro- és makrostruktúrája felől, mint önmagában is informatív aktusok mintázatát közelítik meg.

Az értekezés elején tágabb perspektívából elemzem, hogy a nyelvtudomány különféle irányzatai (generatív grammatika, beszédaktus-elmélet, konverzációs analízis) a prozódia milyen aspektusait világítják meg, majd bemutatom azokat a konkrét elméleti kérdéseket és hipotéziseket, amelyek a disszertáció keretében megvalósított kutatást leginkább motiválták. A saját fejlesztéseim és kísérleteim ismertetése előtt röviden körbejáróm a beszédtechnológia kapcsolódó területeit (beszéddetektálás és beszélő diarizálás) is, illetve ismertetem az intonáció modellezéséhez és számítógépes feldolgozásához elérhető megoldásokat és keretrendszereket (ToBI, Tilt, Prosogram).

A disszertáció legrészletesebben a prozódia automatikus annotációjára kidolgozott módszeremmel foglalkozik, amelynek célja, hogy a beszéd dallamának, intenzitásának és tempójának modulációit egymást követő, kategorikus címkékkel jellemzett események szekvenciájaként írja le. Technikai szempontból ez az adott jellemzőt leginkább meghatározó fizikai paraméter (a dallama esetén például a beszédjel alaphérfkvenciája) mérhető értékeinek stilizálását, illetve a stilizáció eredményeként kapott, a moduláció fő irányát kifejező trendek elemzését követelte meg. Hipotézisem szerint a kimenetként kapott esemény-szekvenciák olyan, a percepció számára is releváns modulációkat tárnak fel, amelyek egy megnyilatkozás attitűdtartalmának értelmezéséhez, a társalgás tagolásához, vagy hosszabb mintákban a beszédszituáció jellemzéséhez is támpontokat jelenthetnek. A beszéd dallamának elemzését végző algoritmus szöveges kimenetét webes környezetben megjeleníthető vizualizációkkal is kiegészítettem. Ez részben az algoritmus validálása érdekében történt, hogy a kimenetként kapott dallam-kontúrok könnyebben összevethetők legyenek a percepcióban megjelenő hangélménnyel.

A hipotézisek igazolására több percepciós kísérletet is lebonyolítottam. Ezekben azt kérdést vizsgáltam, hogy kizárólag a prozódiaira támaszkodva az alanyok (1) egy zárt halmazon belül meg tudják-e határozni a társalgási szituációt, (2) érzékelik-e a téma megváltozását, (3) tudják-e értékelni a az interakcióban résztvevők viszonyát, (4) tudják-e a társalgáson belül pozicionálni annak egy részletét, (5) be tudják-e azonosítani egy megnyilatkozás attitűdtartalmát. A témaváltással detektálásával és a beszédszituáció klasszifikálásával kapcsolatos feladatot a HuComTech korpuszból kinyert prozódiai jellemzőkön betanított mély neuronhálókkal (DNN) is kipróbáltam.