

*Csapó Tamás Gábor* (BME TMIT és MTA–ELTE Lendület Lingvális Artikuláció Kutatócsoport): Az élő helyett jó az online? Valós idejű hang- és képmódosító eljárások online platformokon

Márciusban villám sebességgel váltottunk át az élő kapcsolatok helyett az online platformokra. A Skype, a Zoom, a Teams és a hasonló szoftverek az elmúlt években nagy fejlődésen mentek át, és az a céljuk, hogy az élő kommunikáció során olyan audiovizuális élményt nyújtsanak, mintha élőben, a valós fizikai világban beszélgetnénk valakivel. Az internet mint hálózat velejárója azonban, hogy különböző késleltetések és torzítások léphetnek fel, például ha gyenge WiFi-n keresztül csatlakozunk, vagy nagy a háttérzaj, vagy sokan használják egyszerre ugyanazt a hálózatot. A szoftverek ezeket megpróbálják automatikusan javítani: a kiesett beszédszakaszt automatikusan helyettesítik, felgyorsítják vagy lelassítják; esetleg más beszédsegmentummal helyettesítik. A végfelhasználóknak azonban nincs ezen módosítások felett kontrolljuk. Ugyanígy lehetséges az adó oldalán a mikrofonban felvett beszéd manipulálása (pl. valós időben megoldható a beszéd alaphangfrekvenciájának emelése), így nem lehetünk biztosak abban, hogy amit a képernyőn látunk és hallunk, valóban úgy hangzott el, és abban a környezetben készült, aminek tűnik. Az előadásban áttekintjük a kapcsolódó beszédtechnológiai módszereket az élő beszéd automatikus módosítására, és a manipulációk felismerési lehetőségeire.