

Válasz Dr. Gosztolya Gábor tudományos főmunkatárs bírálatára

Köszönöm Dr. Gosztolya Gábor tudományos főmunkatárs sokoldalú bírálatát, javaslatait és támogató összegzését. Válaszaimat a bírálat szerkezetéhez illeszkedve fogalmazom meg. A bíráló megállapításait *félkövér betűtípussal* idézem meg. A saját válaszaimat pedig normál betűtípussal adom meg.

1. A dolgozat szerkezete

1.1 *„A dolgozat struktúrája átgondolt, .. könnyen olvasható.... fontos ..., hogy a technológiák nem ugyanazon alapötlet .. továbbfejlesztéseként születnek meg, hanem a hardver- és szoftverváltozásokat lekövetve, sokszor alapvetően más elven nyugszanak. .. A bemutatott megoldások nem azonos tesztelési körülmények között ... készültek el.”*

Köszönöm bírálóm elismerő sorait. Valóban arra törekedtem, hogy végig kövessem, ahogy az alkalmazási céltól és a rendelkezésre álló műszaki feltételektől függően újabb és újabb kutatási célkitűzések merültek fel, majd azokat jól követhetően bemutassam. Az összehasonlítás nemcsak azért nehéz, mert a megoldások gyakran más-más adatbázison nyugszanak, hanem a felhasználók szempontjai is merőben eltérhetnek. Például a látássérült emberek számára az a legfontosabb, hogy a lehető legnagyobb beszédsebességű, de érthető beszédet használhassanak. Több olyan vak felhasználóval találkoztam, akik olyan gyors beszédtempóval használták a ProfiVox diád/triád rendszert, hogy én, a fejlesztő már nem tudtam megérteni. A vasúti utastájékoztatásnál pedig az az alapvető, hogy zajos környezetben is kellemes, jól érthető legyen a beszédhang.

1.2 *„Bár Németh Géza dolgozata nem emelte ki az adatbázisokat mint saját tudományos eredményeit ..., én azért szeretném ezt megtenni.”*

Köszönöm, hogy bírálóm is elismerte a beszédatadattal kapcsolatos fontosságát és az ezzel kapcsolatos eredményeimet. Az adatbázisok összessége (ld. disszertáció 2. táblázat) csapatmunka eredménye, ezért noha jelentősen hozzájárultam létrejöttükhöz, azokat nem tekinthetem saját eredménynek. Viszont vannak olyan részterületek, amelyek meg is jelennek a disszertációban saját eredményként. Például az I.2 altézis kimondottam a beszédatadattal tervezéséről szól. A II.2 és a III. 1. altézisnek része az adatbázis tervezési feladat.

1.3 *„Az eljárások esetében jól követhető, hogyan bővültek a hardver lehetőségek, melyek egyre összetettebb, egyúttal egyre absztraktabb algoritmusokat tettek lehetővé. ... hiányolom a ... mélytanulás-alapú rendszerek eredményeit.”*

Mint bírálóm is jelezte, a dolgozat sok éves (1997-2019) időszakot fog át. Ezalatt a gépi beszédalkalmazás négy alapvetően különböző technológiájának kutatás-fejlesztése zajlott. Az adott területi korlátok mellett az ezekhez kapcsolódó minden lényeges kutatási eredményem megfelelő színvonalú ismertetése nem megoldható. Ezért a már megállapodottnak tekinthető három technológiával kapcsolatos kutatási eredményeimre összpontosítottam. Így is több bírálóm hiányolta egyes tézisek részletesebb diszkusszióját. A DNN témakörben a rövidebb átfutási idejű konferencia és folyóirat cikkekben publikáltam eredményeimet.

1.4 „A IV. téziscsoport témája (multimodális beszédinformációs rendszerek) eltér a korábbiaktól... messze esik saját szakterületemtől, így ezen pontok relevanciáját, ... nehezebben is tudom megítélni.”

Valóban, a IV.1 és a IV.3 altézis alkalmazás jellegű, azonban szorosan kötődik a kötetlen szókészletű gépi beszédkeltés (és beszédfelismerés) témaköréhez. Az alátámasztó publikációk bírált nemzetközi konferenciákra (ill. az abból készült könyvhöz) készültek. A IV.2 altézis eredményeit pedig egy ma már 26. alkalomnál tartó nemzetközi konferencia sorozat¹ Budapesten tartott rendezvényén a beszéd-szintézis technológiájából objektív módszerekkel származtatott akusztikus jelkészlet eredeti megoldásaként mutattam be.

1.5 „A 9. fejezet az eredmények gyakorlati alkalmazásait, a pályázóhoz kapcsolódó ... műszaki alkotásokat mutatja be. ... Némelyik bemutatott megoldás (pl. a NAO robot kórházi alkalmazása) nem is igazán kapcsolódik a ... beszédgeneráláshoz. ... Ugyanakkor a gyakorlati alkalmazások ... növelik a(z) ... ismertett tudományos eredmények értékét... az alkalmazott eljárással szemben támasztott követelmények felismerése ... is tanúskodik a pályázó tudományos kvalitásairól.”

Köszönöm, hogy bírálóm is megállapította, hogy a „konkrét gyakorlati esetekben az alkalmazott eljárással szemben támasztott követelmények fölismerése és ezek alapján a megfelelő eljárás kiválasztása” tudományos kvalitást jelez. Éppen ez az indok, ami miatt jelentős terjedelmet szenteltem az alkalmazások körének. Tudniillik az alkalmazások maguk vetnek fel kutatási problémákat, melyek megoldása újabb alkalmazásokat tesz lehetővé, amik újabb kutatási témákat generálnak, és így tovább. Számomra ez a ciklikusság a tudományos mérnöki tevékenység meghatározó eleme. Noha első látásra nem világos, mi köze van a NAO robot kórházi alkalmazásának a célorientált gépi beszédkeltéshez, remélem, érhető lesz az alábbi kicsit részletesebb magyarázattal. A NAO robot ún. humanoid robot, és csak beszéddel vagy gesztussal tud kommunikálni, nincs rajta képernyő. Magyar nyelvű felhasználásához a benne levő beágyazott Linux operációs rendszerű mikroszámítógépen a sok egyéb valós-idejű feladat mellett kellett megoldani a magyar nyelvű gépi beszédkeltés és beszédfelismerés feladatát. Ezt a nem triviális beszéd-szintézis kihívást a ProfiVox diád technológiával tudtuk megoldani. A robot „robotos” hangját nagyon jól fogadták a kísérletben részt vevő gyerekek. Maga az alkalmazási ötlet is annyira innovatív, hogy még 2021-ben is kaptunk rá független nemzetközi hivatkozást (összesen 12 független nemzetközi hivatkozást ért el a kapcsolódó, CoginfoCom konferencián megjelent cikkünk). Sajnos szélesebb körű kutatáshoz vagy gyakorlati felhasználáshoz nem sikerült támogatást szereznünk.

2. A pályázó tudományos publikációi

2.1 „A pályázó tudományos publikációt illetően, az I-III. téziscsoportot alátámasztó cikkek alapvetően hozzák, vagy túl is teljesítik az adott időszak (2000-2016) magyarországi kutatóinak teljesítményét. ... Elvárás lett vezető nemzetközi szakmai folyóiratokban publikálni... meglepő..., hogy bár a pályázónak vannak ilyen... cikkei, ... csak egyetlen ilyen alátámasztó cikk szerepel.”

¹ <https://icad.org/>

Valóban igaz, hogy hosszú ideig alapvetően a Eurospeech/Interspeech konferencia sorozatra fókuszáltam, ahol az intenzív nemzetközi szakmai élet zajlik jelenleg is. Büszke vagyok arra, hogy a sorozat indulása (1987, Edinburgh) óta minden Európában tartott ilyen konferencián volt kutatócsoportunknak publikációja. 1999-ben pedig tudományos titkárként játszhattam szerepet a Budapesti Eurospeech rendezvényen. Ha jelentős mennyiségű eredmény gyűlt össze, akkor igyekeztem folyóirat cikket megjelentetni (pl. az első Speech Communication cikkünk 5 ember 3 éves munkájának eredménye volt és 24 független hivatkozást kaptunk rá). A tudományos minősítési rendszer változása miatt az utóbbi időben valóban előtérbe kerültek a folyóirat publikációk. Például a III.2 altézis eredményei vezettek olyan újabb kutatásokhoz, melyek az IEEE Journal in Selected Topics in Signal Processing folyóiratban publikálásra kerültek. Ezt a disszertációban a 48. oldalon jeleztem is ([79] hivatkozás).

2.2 „A pályázó ... szerkesztőként jegyzi A magyar beszéd c. könyvet, azonban ez több tézispont esetén ... is alátámasztó publikációként van jelölve – itt ... érdemes lett volna a konkrét vonatkozó fejezeteket föltüntetni.”

Köszönöm bírálóm megjegyzését. Valóban szerencsésebb lett volna annak a 15 könyvrészletnek a pontos feltüntetése, ahol szerzőként vagy társszerzőként szerepelek az Akadémia publikációs díját elnyert könyvben. Ez azért maradt el, mert nem találtam hatékony, helytakarékos módszert ennek megoldására. A teljes könyv letölthető az Internetről², és annak xxiii oldalán megtalálható a pontos szerzői azonosítás.

2.3 „Összességében a pályázó publikációs teljesítményét elegendőnek tartom az MTA doktora címhez.”

Köszönöm az elismerő, pozitív publikációs összegzést.

3. Formai megjegyzések

3.1 „Sajnos elég sok az egybe-különírási hiba”

A magyar helyesírás még sok publikációs tapasztalattal bíró kutatóként is kihívásokat tartogat számomra. Igyekeztem helyesírás ellenőrzést futtatni. Lehet, hogy ez is vitt be hibákat.

3.2 „... a feliratok időnként angolul szerepelnek..., ami egy magyar nyelven írt dolgozat esetén nem a legszerencsésebb.”

Sok ábrát angol nyelvű publikációimból vettem át és nem mindig dolgoztam át az angol feliratokat. Ezért a felületességért elnézést kérek.

3.3 „Az irodalomjegyzékben előfordulnak hiányos/túlrövidített szerzőnevek...”

Köszönöm bírálómnak, hogy felhívta a figyelmemet ezekre a hiányosságokra. Ezek részben a Microsoft Word irodalomjegyzék generáló algoritmusából adódnak. Az alábbiakban korigálom az adatokat:

[40] Zhizheng Wu, Oliver Watts, Simon King³

[97] Stavros Garzonis, Simon Jones, TimJay, Eamonn O’Neill⁴

² <http://smartlab.tmit.bme.hu/downloads/pdf/A-magyar-beszed.pdf>

³ https://www.isca-speech.org/archive/SSW_2016/abstracts/ssw9_PS2-13_Wu.html

⁴ <https://dl.acm.org/doi/10.1145/1518701.1518932>

[102] Péter Mihajlik, Zoltán Tobler, Zoltán Tüske and Géza Gordos⁵

[114] Gellért Sárosi, Balázs Tarján, András Balog, Tamás Mozsolics, Péter Mihajlik, Tibor Fegyó⁶

[122] Directorate General for Regional Policy, European Commission⁷.

Gondolom, hogy a [112] a bírálóban elírás

[45] Magyar Elektronika folyóirat, 6. évfolyam, 12. szám, 1989

4. Összefoglaló értékelés

„... úgy vélem, hogy a pályázó tudományos teljesítménye megfelelő ahhoz, hogy részére az MTA doktora címet odaítéljék.”

Köszönöm Dr. Gosztolya Gábor tudományos főmunkatárs elismerő véleményét. Külön köszönöm, hogy kiemeli a kutatás előfeltételeinek megteremtése és a kutatási eredmények gyakorlati alkalmazása érdekében végzett erőfeszítéseimet.

Budafok, 2021. 07. 10.

dr. Németh Géza

⁵ https://www.isca-speech.org/archive/archive_papers/interspeech_2005/i05_2677.pdf

⁶ <https://ieeexplore.ieee.org/document/6421932/authors#authors>

⁷ https://ec.europa.eu/regional_policy/sources/docgener/presenta/projectbook/dg_regio_project_book_en.pdf