Kuptueshmëria e të folurit në gjuhën shqipe në përdorimin e aplikacioneve multimediale Skype dhe Viber

Blerta Prevalla[*]

Altin Shala[**]

Abstrakt

Ky punim ka për qëllim të analizojë matjet subjektive të kuptueshmërisë së të folurit në gjuhën shqipe, gjatë bisedës në mes të dy personave, duke përdorur aplikacione, të cilat në ditët e sotme janë shumë të përdorura për komunikim, siç janë Skype dhe Viber.

Kuptueshmëria e të folurit ka të bëjë me qartësinë e të folurit që dëgjohet, dmth në hyrje të sistemit kemi fjalët pa kuptim, kurse në dalje kemi atë që dëgjohet. Kuptueshmëria e të folurit paraqet përqindjen e fjalëve të pranuara saktë ndaj numrit të fjalëve të shqiptuara në hyrje. Matja bëhet në këtë mënyrë: në pjesën hyrëse të sistemit të transmetimit fliten fjalë ose fjali apo vetëm rrokje, ndërsa në anën pranuese regjistrohet ajo që dëgjohet. Përqindja e fjalëve, fjalive apo rrokjeve të pranuara saktë, në raport me numrin e atyre të shqiptuara në hyrje të sistemit, jep përqindjen e kuptueshmërisë.

Një shembull konkret, sepse nevojitet digjitalizimi i gjuhës shqipe dhe pse duhet të dizajnohen modele dhe algoritme për këtë, është rasti te Google, Search by Voice, ku edhe nëse flitet ndonjë fjalë jo qartë në gjuhën angleze, sistemi e gjen vetë fjalën më të afërt të mundshme dhe e gjeneron atë.

Metodat e matjeve bëhen me shpejtësi të ndryshme të internetit të realizuara në ambiente pa zhurmë dhe ato me zhurmë, me

[*] *Blerta Prevalla, Phd.Cand., Universiteti AAB, Fakulteti i Shkencave Kompjuterike, blerta.prevalla@aab-edu.net*

[**] *Altin Shala, MSc, Universiteti AAB, Fakulteti i Shkencës Kompjuterike, altin.shala@aab-edu.net*

qëllim për të parë ndikimin në kuptueshmëri të të folurit me parametra të ndryshëm objektiv.

Fjalë kyç: *Shkalla e gabimit të fjalëve WER (word error rate), kuptueshmëri, zhurmat, Skype, Viber, VoIP.*

Hyrje

Historikisht, të folurit dhe përpunimi i tij është trajtuar në mënyra të ndryshme si në shkencë kompjuterike, inxhinieri elektrike, gjuhësi, dhe psikologji. Fillet e para të zhvillimit të këtyre modeleve kanë lindur pas luftës së dytë botërore, kur edhe filloi zbulimi i kompjuterit, pra, në periudhën nga vitet 1940 deri në fund të viteve 1950, ku në mënyrë intensive u punua në zhvillimin e këtyre modeleve dhe në kuptueshmërinë e të folurit[1]. Njohja e të folurit, si ide është hasur disa dekada më herët tek filmat shkencor, ku kompjuterët e njihnin dhe identifikonin të folurit e personit, pa marrë parasysh se sa shpejt dhe çfarë gjuhe fliste ai. Por edhe sot e kësaj dite, në realitet nuk është arritur të dizajnohet një program për njohjen e të folurit, ashtu siç përshkruhet në filmat shkencor.

Skype dhe Viber janë aplikacione kompjuterike dhe mobile që përdoren si softuer për testim, të cilat ofrojnë komunikim përmes të folurit dhe në mënyrë të shkruar. Skype dhe Viber përdorin standardin *VoIP*[2] që do të thotë komunikim me zë përmes internetit. Këto aplikacione bëjnë pjesë në grupin e aplikacioneve të ashtuquajtura ueb aplikacione, të cilat funksionojnë duke

[1] Daniel Jurafsky, & James H. Martin, *Speech and Language Processing: An introduction to natural language processing, computational linguistics, and speech recognition*, 2006, Draft of June 25, 2007.

[2] VoIP –Do të thotë komunikimi me zë përmes Internet Protokollit siç janë aplikacionet MSN në facebok, Viber edhe shumë forume bisedash që funksionojnë me lidhje në internet.

pasur qasje në internet. Studimet rreth kuptimit të fjalës mund të kryhen në mënyra të ndryshme.

Që nga shpikja e Alexander Bell, inxhinierët dhe shkencëtarët e kanë studiuar fenomenin e komunikimit të të folurit si komunikim në mes njerëzve përmes pajisjeve telekomunikuese ose komunikimit në mes njeriut dhe makinës. Duke filluar që nga vitet 60-të, përpunimi digjital i sinjalit (DSP-Digital Signal Processing) dhe paraqitja e tyre si në formën vizuale dhe matematikore kanë qenë problem dhe njëherit diçka abstrakte për njerëzimin.[3]

Sot në teknologjinë moderne ende nuk është dizajnuar ndonjë softuer i tillë, që në bazën e të dhënave të kenë të vendosur edhe fjalët e gjuhës shqipe, përveç disa programeve që punojnë lokalisht.

Efekti i shpejtësisë së internetit në kuptueshmërinë e të folurit gjatë përdorimit të aplikacioneve që punojnë me platformën e VoIP

Disa shërbyes të internetit ofrojnë lidhje në internet me ruterin pa tela dhe në këtë rast është shumë praktike, e përshtatshme dhe komode biseda duke përdorur Skype dhe Viber. Duhet pasur parasysh se lidhja në internet me lidhje kabllore mund të jetë shumë më efikase për stabilitet gjatë bisedës në mes personave, sepse shpejtësia e internetit është konstante, ndërsa tek lidhjet pa tela kjo mund të jetë variabile. Gjithashtu tek lidhjet pa tela mund të ketë ndikim në jostabilitetin, sepse lidhjet pa tela ofrohen për qasje nga shumë përdorues dhe kjo ndikon në uljen e shpejtësisë së internetit, pastaj tek lidhjet pa tela ndikim më të madh mund të kenë efektet e tjera si interferencë dhe pengesa, të

[3] Adnan Maxhuni, *Gjetja e modelit për sintezën e të folurit nga tekstet e shkruara në gjuhën shqipe*, Prishtinë: Universiteti i Prishtinës "Hasan Prishtina",f. 10-11, 2014.

cilat shfaqen për shkak të brezave frekuencore dhe reflektimi i valëve.[4]

Efekti tjetër në lidhje me përdorimin e VoIP, është edhe humbja e paketave dhe si shkak i kësaj ndodh të humbë një pjesë e bisedës me ç'rast humb kuptueshmëria e të folurit. Humbja e paketave ndodh për shkak të ngarkesës në rrjet dhe si pasojë e kësaj, paketat me të dhëna audio mund të qëndrojnë në rrjet më shumë se sa është e caktuar për një kornizë dhe nëse kalohet koha e caktuar, atëherë paketa humbet, dmth nuk ka destinacion deri në atë moment. Shërbimet e internetit ndahen në klasa sipas kualitetit që ofrohet dmth nga klasa më e mirë që fillon nga A, B, C, e deri tek ajo më e ulëta në bazë të kualitetit D.

Jitter (dridhjet) - paraqesin devijimet standarde në mes të paketave apo kornizave të të dhënave. Madhësia e Jitter paraqet vlerësimin e valëzimit të të dhënave. Të gjitha vonesat që shkaktohen nga Jitter janë dëmet kryesore të rrjetit që ndikojnë në cilësinë e zërit [5]. Rritja e vonesave varet nga distanca fizike në mes komunikuesve dhe nga mediumi të cilin përdorin. Kur transmetimi realizohet me fije optike ose me mediume të tjera, paketat kanë një tolerancë për një vonesë në vlerë 5μs/km.[6]

ITU[7] rekomandon që vonesat e Jitter të mos jenë më të mëdha se sa 150 ms për shumicën e aplikacioneve dhe një kufizim për aplikacionet me komunikim me zë rreth 400ms.

Vonimi i përgjithshëm në sistem përbëhet prej komponentëve që vijojnë:

---

[4] Shërbim për testim online të shfletuesit. http://voiptest.8x8.com/. Hapur më: 15.04.2016.

[5] Lingfen Sun, *Spech Quality prediction for voice over internet protocol networks* , University of Plymouth, Ph.D. Thesis, January 2004.

[6] Jeff Rodman, *The Effect of Bandwidth on Speech Intelligibilty* ,Commonwealth Telecommunications Organisition, September 2006.

[7] ITU-International Telecommunication Union (Unioni Ndërkombëtar për Telekomunikacion).

1. Vonimi gjatë procesit të kodimit
2. Vonimi për shkak të pritjes
3. Vonimi i transmetimit
4. Vonimi për shkak të variacioneve të vonimit dhe masave të përmirësimit të variacioneve të vonimit
5. Vonimi gjatë procesit të dekodimit

Testimi dhe nxjerrja e rezultateve

Llogaritja e Word Error Rate - Parametri matës për matjen e kuptueshmërisë së të folurit është shkalla e gabimit të fjalëve (Word Error Rate). Për disa sisteme të thjeshta të njohjes (si p,sh fjalët e izoluara), performanca është përqindja e fjalëve të humbura ndaj totalit të fjalëve. Megjithatë ky parametër matës nuk është efikas, sepse sekuenca e fjalëve të njohura mund të përmbajë tri lloje të gabimeve. [8]

Ngjashëm me gabimin për njohjen e shifrave, gabimi i parë i njohur si ndërrimi i fjalëve, ndodh kur një fjalë e pasaktë është njohur në vend të fjalës së folur saktë. Gabimi i dytë, i njohur si shuarje e fjalëve, ndodh kur një fjalë e thënë nuk është e njohur (dmth, fjalitë e njohura nuk e kanë fjalën e folur në hyrje). Dhe në fund e kemi gabimin e tretë, i njohur si futja e fjalëve gjatë përpunimit. Ky rast ndodh kur fjalët e kyçura janë pranuar nga njohja e tyre (dmth, fjalia e njohur dhe e pranuar pranon më shumë fjalë se sa është dhënë në burim, siç janë zhurmat). Një shembull i tillë do të ishte :

Fjalia e folur në hyrje – Mirëmbrëma, si jeni, ka diçka të re nga ju?

Fjalia e kuptuar dhe e pranuar në dalje - Mirëmbrëma sa jeni, ka diçka nga ju!

[8] Myzafere Limani, *Elektroakustika*, Ligjerata të autorizuara, Universiteti i Prishtinës, 2005.

Shkalla e gabimit të fjalëve definohet si përqindja e fjalëve të pranuara gabim ndaj numrit të fjalëve të shqiptuara në hyrje.

$WER = 100\% * \left(\frac{S+D+I}{|W|}\right)$.....(1)

- Substitutions - Ndërrimi i fjalëve
- Deletions - Shuarja e fjalëve
- Insertions - Futja e fjalëve
- W - Numri total i fjalëve
-

$Shkalla\ e\ gabimit\ të\ fjalëve = 100\% * \left(\frac{\boldsymbol{Numri\ i\ fjalëve\ të\ gabuara}}{\boldsymbol{Numri\ total\ i\ fjalëve}}\right)$....(2)

Për të kuptuar më mirë matjen, ne do të japim një shembull të bisedës në mes një dhënësi dhe një marrësi duke përdorur Viber.

Fjala komplet e shqiptuar është: Në javën e ardhshme udhëtoj për në SHBA në trajnimet për Elektronikë VLSI.

Folësi në hyrje	Në	javën e	ardhshme	udhëtoj	për në	SHBA	në trajnimet	për	Elektronikë	VLSI
Pranuesi në dalje	Për	javën e	tashme	*****	për në	SHBA	në trajtimet	në	elektrikë	SI
Evaluimi	S		S	D			S	I	S	I

Duke zbatuar shprehjen e mësipërme

Shkalla e gabimit të fjalëve (WER)=$100\frac{4+2+1}{10}$=70%

WER=70%

Ndërsa kuptueshmëria llogaritet :

Shkalla e kuptueshmërisë së fjalëve definohet si përqindja e fjalëve të pranuara saktë ndaj numrit të fjalëve të shqiptuara në hyrje.

Kuptueshmëria e të folurit = $\frac{\boldsymbol{Numri\ i\ fjalëve\ të\ pranuara\ saktë}}{\boldsymbol{Numri\ total\ i\ fjalëve}} * \mathbf{100}\%$....(4.3)

Niveli i kënaqshmërisë shprehet me përqindje dhe kuptueshmëria ka disa ndarje në bazë të rezultateve të nxjerra.

Niveli performancës	E shkëlqyer	E mirë	E mjaftueshme	E dobët
Kuptueshmëria	76 - 100	66 - 75	61 - 65	30 - 60

Tabela 1. Shkalla e performancës së kuptueshmërisë së të folurit

Rezultatet e nxjerra janë bërë në bashkëbisedim në mes të dy personave meshkuj. Në dhënës, pra në hyrje të sistemit tek leximi i tekstit është Dardan Mehmeti mosha 25 vjeç me lexueshmëri dhe ton të qartë, kurse në marrës apo në dalje është Altin Shala, mosha 30 vjeç, i cili i ka vlerësuar rezultatet e këtyre matjeve. Sa i përket metodave dhe mënyrës së matjeve është zgjedhur mënyra për të testuar tekste të ndryshme, për të parë dhe krahasuar rezultatet e kuptueshmërisë së të folurit në gjuhën shqipe, varësisht nga tekstet që lexohen për testim duke u bazuar edhe në gjuhët e tjera botërore dhe modele të ndryshme. Si arritje tjetër është se nga rezultatet e konsultuara, gjuha jonë në aspektin e kuptueshmërisë dhe përshtatjes për komunikim në model është e ngjashme me gjuhën serbe dhe kroate[9].

Kur jemi tek tekstet e përdorura për lexim dhe komentim më të lehtë kemi përdorur ekuivalentimin e titujve me shkurtesa si në vijim:

F1	Teksti i njohur nga marrësi
F2	Teksti i panjohur për marrësin
F3	Shqiptimi i 50 fjalëve me bashkëtingëllore dy shkronjëshe
F4	Shqiptimi i 150 fjalëve më të përdorura në gjuhën shqipe
F5	Shqiptimi i 50 fjalëve më të gjata që përdoren më pak në gjuhën shqipe

[9] Nebi Caka & Ali Caka, *Lista e 100 fjalëve të para të radhitura sipas dendurisë së paraqitjes në korpusin njëmilionfjalësh të gjuhës shqipe*, Universiteti i Prishtinës, Prishtinë, 2006.

Matjet me aplikacionin Skype. Specifikacionet për kualitetin e internetit për matjet me aplikacionin Skype.

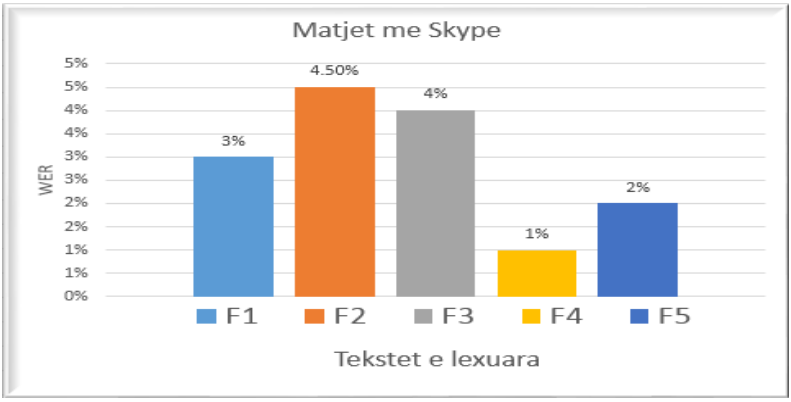
- Shpejtësia e shkarkimit -100 Mbps
- Shpejtësia e ngarkimit - 50 Mbps
- Ping - 30 ms
- Ping-u në mes serverit të Skype dhe PC-së që i kemi bërë matjet është Ping - 66ms
- Jitter - 21ms
- Linja e kualitetit të internetit është B
- MOS[10]- 4.29
- Distanca e matjes nga hosti lokal i ofruesit të Internetit Star Link në serverin në Prishtinë Art Motion

Siç e treguam edhe më lart për përdorimin e shkurtesave të teksteve, në fillim është bërë matja për tekstin F1 me titull: ''Qarqet amplifikuese me shumë stade'', gjithsej 477 fjalë dhe pastaj me rradhë për tekstet e tjera.

Tekstet e lexuara	Kuptueshmëria e të folurit %	WER (Word Error Rate) Shkalla e gabimit të fjalëve%
F1	97%	3%
F2	95.5%	4.5%
F3	96%	4%
F4	99%	1%
F5	98%	2%

Në diagramin më poshtë e kemi paraqitur varshmërinë e gabimit të fjalëve në raport me tekstin e lexuar.

[10] MOS-Mean Opinion Score (parametër matës objektiv për kualitetin e shërbimeve)



Matjet me aplikacionin Viber

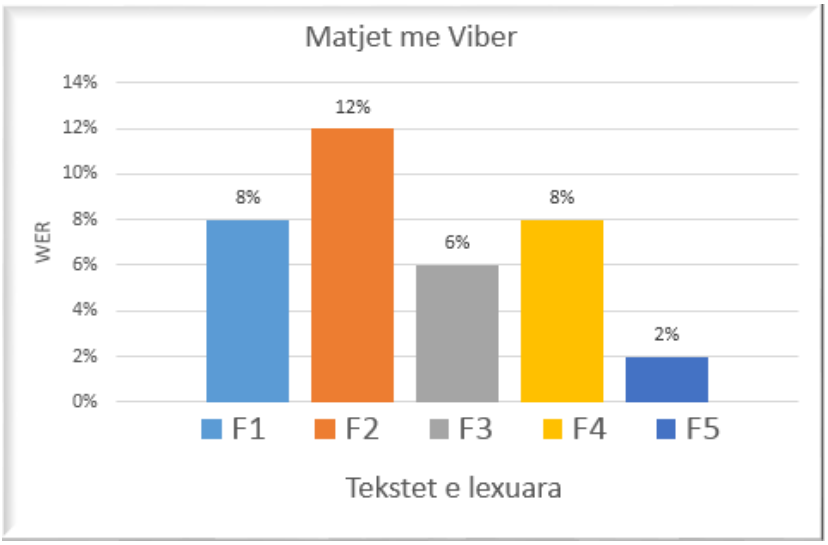
Të dhënat për kualitetin e internetit për matjet me aplikacionin Viber në ambient të pastër pa zhurma.

- Shpejtësia e shkarkimit 4Mbps
- Shpejtësia e ngarkimit 1.8 Mbps
- Ping 48 ms
- Jitter 1ms
- Linja e kualitetit të internetit është B
- MOS 4.1
- Distanca e matjes nga hosti lokal Ipko në serverin në Frankfurt të Gjermanisë

Njësoj sikurse tek Skype edhe në Viber kemi vepruar me metoda të njëjta të matjeve, vetëm se këtu kemi përdorur Viber në Smart Phone me specifikacione të cekura më lart.

Tekstet e lexuara	Kuptueshmëria e të folurit %	WER (Word Error Rate) Shkalla e gabimit të fjalëve%
F1	92%	8%
F2	88%	12%
F3	94%	6%
F4	92%	8%
F5	98%	2%

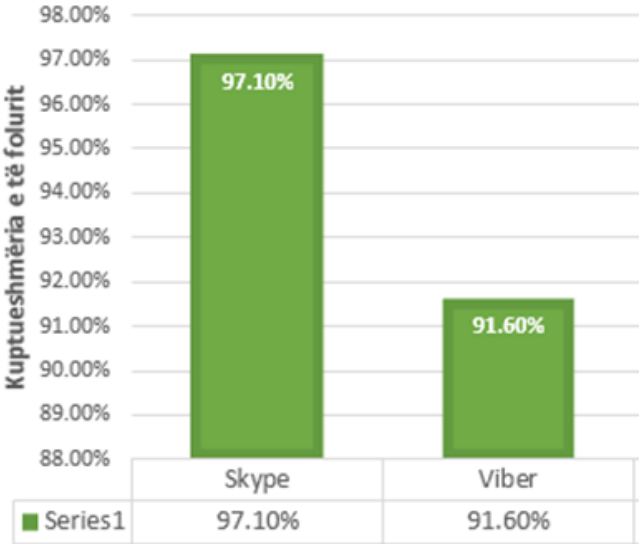
Nga këto të dhëna paraqesim diagramin si më poshtë.



Në vazhdim do të paraqesim dallimin në mes të aplikacioneve Skype, Viber dhe Zoiper në aspektin e kuptueshmërisë së të folurit në ambientet pa zhurma, për të parë si përmbledhje e të gjithë teksteve totale, dmth nga të gjitha llojet e teksteve duke gjetur vlerën mesatare

Aplikacionet e përdorura	Kuptueshmëria e të folurit % (Vlera mesatare)
Skype	97.10%
Viber	91.60%

Kurse diagrami i paraqitjes ka këtë formë.



Përfundim

Rezultatet e arritura të kuptueshmërisë së të folurit janë të shkëlqyeshme për të gjitha aplikacionet e përdorura, që do të thotë se janë me vlerën mbi 76%, ku konsiderohen që të jenë të shkallës më të lartë. Nga kjo përfundojmë se, gjuha shqipe është një gjuhë që kuptohet shumë lehtë dhe arsyeja kryesore sepse janë këto rezultate pozitive, janë zanoret, të cilat i japin kuptimin fjalëve. Mbetemi me rezervë për këto rezultate, pasi që matjet e bëra në këtë punim janë bërë vetëm në mes të dy personave dhe

duke u njohur në mes tyre. Nëse bëhen testime me më shumë bashkëbisedues të panjohur në mes tyre, pritet të kemi vlera më të ulëta të kuptueshmërisë së të folurit. Një ulje e kuptueshmërisë mund të jetë tek matjet kur personat që komunikojnë nuk njihen midis tyre, gjë që në rastin tonë matjet janë realizuar në mes personave, të cilët njihen dhe kjo është një nga arsyet për rezultate me vlera më të larta.

Rekomandime

Për të ardhmen në këtë fushë mund të themi se ka shumë për të bërë dhe kërkohet që të punohet në grup me fusha të ndryshme si inxhinieri, programim, gjuhësi, sepse sa sfiduese që është, po aq është edhe e nevojshme për kohën e sotme. Realisht, duhet që të bashkohen forcat gjithë kombëtare të ekspertëve përkatës dhe të bëhet digjitalizimi i gjuhës shqipe, sepse kjo do të ketë përparësi jo vetëm në kuptueshmërinë e të folurit, por edhe në shumë sisteme automatike si robotikë, mjekësi, tek makinat që punojnë me sinjalizim të zërit, të cilat edhe përdoren në vendin tonë. Në të ardhmen, në industrinë e prodhimit të makinave të teknologjisë së informacionit është e pakuptimtë të mos integrohet gjuha shqipe në aplikacione dhe makina që do të zhvillohen.

Bibliografia

Beci, Bahri. *Fonetika e gjuhës shqipe*. Tiranë: 2004.

Caka, Nebi. & Caka, Ali. *Lista e 100 fjalëve të para të radhitura sipas dendurisë së paraqitjes në korpusin njëmilionfjalësh të gjuhës shqipe.* Universiteti i Prishtinës. Prishtinë: 2006.

Limani, Myzafere. *Ligjerata të autorizuar,* Prishtinë 2005.

Jurafsky. Daniel and James H**.** Marti :*Speech and Language Processing: An introduction to natural language processing,*

computational linguistics, and speech recognition Copyright 2006, draft of June 25, 2007.

Maxhuni, Adnan. *Gjetja e modelit për sintezën e të folurit nga tekstet e shkruara në gjuhën shqipe.* Prishtinë: Universiteti i Prishtinës "Hasan Prishtina". 2014.

Rodman, Jeff. *The Effect of Bandwidth on Speech Intelligibility.* Commonwealth Telecommunications Organization, September 2006.

Veton Këpuska. "VOn Wake-Up-Word Speech Recognition Task, Technology, and Evaluation Results against HTK and Microsoft SDK 5.1. "Invited Paper: World Congress on Nonlinear Analysts, Orlando 2008, *Journal of Nonlinear Analysis, Theory, Methods & Applications. & Klein*, T. 2008.

Vipperla, Ravichander. *Automatic Speech Recognition for ageing voices*, Institute for Language, Cognition and Computation School of Informatics University of Edinburgh 2011.

Sun, Lingfen. *Spech Quality prediction for voice over internet protocol networks*, University of Plymouth, Ph.D. Thesis, January 2004