

FJALORI THEMELOR I GJUHËS SHQIPE NË PËRQASJE ME FJALORIN THEMELOR TË ANGLISHTES

ALI CAKA; VJOLLCA BELEGU-CAKA

AKTET V, 4: 605-!Syntax Error, !, 2012

PËRMBLEDHJE

Në këtë punim do të bëhet krahasimi i tre fjalorëve themelorë të gjuhës shqipe, me 1500, 3000 dhe 5000 fjalë, hartuar nga autori i parë i këtij punimi (A. Caka) bazuar në korpusin njëmilionfjalësh të gjuhës shqipe, me fjalorët përkatës (për nga numri i fjalëve) themelorë të gjuhës angleze me qëllim të përcaktimit të mospërputhjeve të fjalëve të përfshira në këta fjalorë. Në bazë të tyre do të nxirren përfundime se a janë këto mospërputhje pasqyrë e kulturës së popujve që i flasin këto gjuhë apo të metodologjisë së hartimit të fjalorëve themelorë përkatës dhe do të jepen sugjerime se çfarë duhet bërë që fjalori themelor i gjuhës shqipe të jetë sa më përfaqësues.

Fjalët kyçe: fjalori themelor, mësimi-i-shqipes, mësimi-i-anglishtes, gjuhë-e-huaj

SUMMARY

This paper will compare three basic vocabularies of Albanian Language, with 1500, 3000 and 5000 words, compiled by the first author of this paper (A. Caka) based on one million Albanian Language corpus, with corresponding basic vocabularies (for the number of words) of English Language in order to determine the variance of the words included on these vocabularies. Based on them will be drawn conclusions on whether these discrepancies are reelection of the culture of peoples who speak these languages or methodology underlying the compilation of vocabularies and will be given suggestions on what should be done that the basic vocabulary of the Albanian Language to be more representative.

Key words: basic-vocabulary, teaching-of-Albanian, teaching-of-English, foreign-language

HYRJE

Përzgjedhja e numrit minimal të fjalëve që mund të përdoren së bashku në konstrukte gramatikore në kontekste të ndryshme është njëri ndër problemet kryesore dhe ndër më të vështirit gjatë mësimin të një gjuhe të huaj, dhe ngërthen në vete hartimin e fjalorëve të posaçëm bërthamë shpjegues njëgjuhësh dhe dygjuhësh.

Në komunikimin e përditshëm në gjuhë amtare, njerëzit rëndom shfrytëzojnë një fond prej 1500 deri 2000 fjalë, të cilat i përdorin dhe i ripërdorin në pjesën më të madhe të kohës pothuajse në mënyrë automatike.

Sot (sipas Jean Pruvost-it) merret se shumica e francezëve përdorin më pak se 5000 fjalë.

Bazuar në “Metodën Pimsleur” e nxënies së gjuhëve të huaja të Dr. Pimsleur është me rëndësi jo sa fjalë të një gjuhe i mësojmë, por që ato fjalë të jenë fjalë të përdorimit të dendur, që përbëjnë leksikon bërthamë të kësaj gjuhe dhe që ato të dimë t’i përdorim me strukturat përkatëse.

Përkufizimi i fjalorit bërthamë

Ekzistojnë mënyra të ndryshme për të përfutur fjalorin bërthamë të një gjuhe, si p.sh. [Lehman 1989]:

1. Fjalori bërthamë i një gjuhe përbëhet prej n fjalëve më të shpeshta (më të dendura) të asaj gjuhe.

2. Fjalori bërthamë i një gjuhe është fjalori i cili është i përbashkët për të gjithë folësit që atë gjuhë e kanë gjuhë amtare.

3. Fjalori bërthamë semantik i një gjuhe përbëhet prej atyre fjalëve që mjaftojnë për të përkufizuar të tërë pjesën tjetër të fjalorit (leksikut) të asaj gjuhe.

Siç mund të vërehet, dy përkufizimet e para bazohen në metoda statistikore, ndërsa përkufizimi i tretë kërkon një qasje semantike.

Metodat statistikore na japin përndarjen (distribucionin) e fjalëve në një korpus të dhënë, si p.sh. në “korporin e Braunit” prej rreth 1 milion fjalëve (paraqitjesh) kemi këtë përndarje të fjalëve [Kucera-Francis 1967]:

10 fjalët më të shpeshta: 24,26%

100 fjalët më të shpeshta: 47,43%

1000 fjalët më të shpeshta: 68,86%

Këto shifra variojnë (ndryshojnë) pak nga madhësia e korpusit. Vlera të përafërta janë përftuar edhe gjatë analizës së korpusit të gjuhës gjermane. Por, duhet vërejtur se përderisa fjalët më të shpeshta në një korpus janë të qëndrueshme (stabël), denduria e paraqitjes së tyre ndryshon nga korpusi në korpus apo nga një nënkorpus te tjetri nënkorpus brenda një korpusi të caktuar. Kjo varet nga:

1. Tema
2. Funkzioni komunikues.

Kështu, p.sh. në një nënkorpus për kulturë do të hasen më dendur (më shpesh) fjalët si: muzikë, teatër, regjisor etj., ndërsa në një nënkorpus për sport do të hasen më dendur (më shpesh) fjalët si: gol, futbollist, trajner etj.

Karakteristikat e paraqitjes së fjalëve: - Ka shumë pak fjalë me denduri të madhe dhe shumë fjalë me denduri të vogël (“Përndarja e Zipf-it”) [Zipf 19..].

Fjalët më të shpeshta pothuajse në të gjitha gjuhët janë fjalët gramatikore. Në listën e fjalëve më të shpeshta mund të hasen edhe foljet ndihmëse e modale, si p.sh.: jam, kam, duhet etj.

Kështu, p.sh., fjala (nyja) të përfshin rreth 7.9 % të një teksti në shqip. Njëqind (100) fjalët më të shpeshta përfshijnë mesatarisht rreth 56 % të një teksti, ndërsa 1000 fjalët më të shpeshta përfshijnë mesatarisht rreth 75 % deri 82 % të një teksti. Njëmijë (1000) fjalët e tjera marrin pjesë në tekst me rreth 7 %, e 1000 fjalët e tjera pas tyre

marrin pjesë me vetëm 3.7 %. Pra vetëm një numër i vogël fjalësh paraqitet me denduri të madhe; shumica e fjalëve kanë një denduri të vogël të paraqitjes, e disa madje nuk paraqiten më shumë se një a dy herë në një korpus prej një milion (1.000.000) fjalësh. Shumica e 100 fjalëve më të shpeshta në anglishte janë fjalë gramatikore. Ndër 100 fjalët më të shpeshta të shqipes [Korpusi CAKA] hasen edhe këto fjalë leksikore: një, bëj, them, dua, mund, vit, punë, kohë, madh(i), mirë, shqiptar, vend, jetë, di, ditë, gjuhë, fund, fjalë, shkollë.

Siç u tha, fjalori bërthamë nuk mund të ndërtohet vetëm në bazë të dendurisë së paraqitjes së fjalëve në një korpus, por duhen marrë parasysh edhe kriteret semantike, si ai i përmbylljes linguistike (angl. linguistic closure) – mundësisë së formimit të fjalive natyrore me vetëm fjalët e bashkësisë (listës) së caktuar. Kështu kemi rastin, bie fjala, kur në një fjalor të bazuar vetëm në kriterin statistik (sipas dendurisë) del fjala (mbiemri) femëror, por jo dhe antonimi i saj: mashkullor etj.

Gjatë përcaktimit të fjalëve që duhen përfshirë në fjalorin themelor (bërthamë) duhen marrë parasysh edhe dy parametra statistikorë:

1) rangu (diapazoni) a shtrirja e fjalës: numri i teksteve të ndryshme në të cilat paraqitet fjala, goftë edhe vetëm një herë dhe

2) rrafshësia e përndarjes (distribucionit) të fjalës, d.m.th. fakti se një fjalë paraqitet me një denduri të konsiderueshme dhe relativisht të njëjtë në një diapazon të gjerë të tekstit.

Të kombinuar, këta dy parametra na japin masën statistikore të neutralitetit pragmatik, meqë na tregojnë se a është një fjalë e kufizuar në përdorime të veçanta diatipike. [Stubbs 1986].

Çështja e zgjedhjes së fjalëve që do të përfshihen në një fjalor, sidomos në fjalorët shkollorë, ka një shekull që është duke marrë rëndësi gjithnjë e më të madhe, dhe është aktuale edhe sot e kësaj dite. Synim i po thuajse të gjithë këtyre fjalorëve ishte shkurtimi sa më shumë i afateve të nxënies (së kënaqshme) së gjuhëve të huaja.

Duhet theksuar se kriteri i dendurisë për përcaktimin e një fjalori themelor, çfarë janë edhe shumica e fjalorëve shkollorë, nuk është i

mjaftueshëm, ngase me fjalët e një fjalori themelor të hartuar vetëm sipas kriterit të dendurisë së paraqitjes së tyre në një korpus të marrë si “përfaqësues” të një gjuhe nuk do të mund të formoheshin “fjali natyrale” apo fjali të një sistemi natyral të gjuhës. Me një fjalor themelor të tillë, nuk do të mund të përkufizoheshin (shpjegoheshin a sqaroheshin) fjalët e listuara në të. Kriteri i dendurisë do të duhej të shoqërohej edhe me kriteret semantike, pra me fjalët relevante për të kuptuar fjalët e përfaqësuar në këtë fjalor themelor, i cili në fakt do të ishte një fjalor bërthamë semantik.

Gjatë përcaktimit të fjalëve që duhen përfshirë në fjalorin themelor (bërthamë) duhen marrë parasysh edhe dy parametra statistikore:

1) rangu (diapazoni) a shtrirja e fjalës: numri i teksteve të ndryshme në të cilat paraqitet fjala, qoftë edhe vetëm një herë dhe

2) rrafshësia e përndarjes (distribucionit) të fjalës, d.m.th. fakti se një fjalë paraqitet me një denduri të konsiderueshme dhe relativisht të njëjtë në një diapazon të gjerë të teksti.

Të kombinuar, këta dy parametra na japin masën statistikore të neutralitetit pragmatik, meqë na tregojnë se a është një fjalë e kufizuar në përdorime të veçanta diatipike. [Stubbs 1986].

Çështja e zgjedhjes së fjalëve që do të përfshihen në një fjalor, sidomos në fjalorët shkollorë, ka një shekull që është duke marrë rëndësi gjithnjë e më të madhe, dhe është aktuale edhe sot e kësaj dite. Synim i po thujtës së gjithë këtyre fjalorëve ishte shkurtimi sa më shumë i afateve të nxënies (së kënaqshme) së gjuhëve të huaja.

Duhet theksuar se kriteri i dendurisë për përcaktimin e një fjalori themelor, çfarë janë edhe shumica e fjalorëve shkollorë, nuk është i mjaftueshëm, ngase me fjalët e një fjalori themelor të hartuar vetëm sipas kriterit të dendurisë së paraqitjes së tyre në një korpus të marrë si “përfaqësues” të një gjuhe nuk do të mund të formoheshin “fjali natyrale” apo fjali të një sistemi natyral të gjuhës. Me një fjalor themelor të tillë, nuk do të mund të përkufizoheshin (shpjegoheshin a sqaroheshin) fjalët e listuara në të. Kriteri i dendurisë do të duhej të shoqërohej edhe me kriteret semantike,

pra me fjalët relevante për të kuptuar fjalët e përfaqësuar në këtë fjalor themelor, i cili në fakt do të ishte një fjalor bërthamë semantik.

Ky fakt i detyroi metodistët e sotëm të heqin dorë nga mendimi për t'i zgjedhur fjalët, duke u nisur vetëm nga denduria dhe filluan të përpunojnë edhe kriteret tjera për seleksionimin e fjalëve, siç janë: tematika, mundësia e një fjale për t'u lidhur me fjalë të tjera, vlera semantike e fjalëve, polisemia e tyre dhe kriteret tjera ndihmëse [Basha 1979].

Një analizë krahasimtare e semantikës leksikore bazuar në tekste në anglishte, gjermanishte dhe rusishte jepet në tekstin: Sharoff, Serge, (2002). “Meaning as use: exploitation of aligned corpora for the contrastive study of lexical semantics”. Proc. of Language Resources and Evaluation Conference (LREC02). May, 2002, Las Palmas, Spain.

Analiza e fjalorit të një autori mund të shihet nga teksti: Nebi Caka, Analizë statistike dhe stilistike e gjuhës së Lasgush Poradecit, punim magjistrature, Prishtinë, 1986.

Për analizën e fjalëve në kontekst mund të shërbejë teksti: Davies, Mark, Corpus del Español, 2002.

LESIKU (VOKABULARI) BËRTHAMË

Anglishtja speciale

“Zëri i Amerikës” (VOA) që nga viti 1959 transmeton në radio, e së voni edhe në televizionin satelitor, programin e vet edhe në të ashtuquajturën “anglishte speciale”, dedikuar njerëzve që nuk e kanë gjuhë amtare anglishten. Anglishtja speciale ka një fjalor bërthamë prej 1500 fjalëve. Ato janë kryesisht fjalë të thjeshta që përshkruajnë gjësende, veprime dhe emocione (ndjenja). Fjalitë e shkruara me këtë anglishte janë të thjeshta dhe shqiptohen (lexohen) qartë dhe me një temp më të ngadaltë. Qëllimi i përdorimit të këtij ‘varianti’ të anglishtes amerikane është që dëgjuesit anë e mbanë botës të cilëve u dedikohen këto emisione të mësojnë më lehtë anglishten amerikane nëpërmjet informimit për jetën amerikane, për të rejtat e fundit në botë, si dhe për të arriturat më të reja në shkencë, e posaçërisht në mjekësi.

[<http://www.voanews.com/specialenglish/>].

Fjalori themelor Oxford 3000

Fjalët-titull të fjalorit themelor të Oxford 3000 janë përzgjedhur me kujdes nga një grup specialistësh të gjuhës dhe mësuesish me përvojë, si fjalë të cilat duhet fituar prioritetin në studimin e fjalësit (vokabularit) për shkak të rëndësisë dhe të dobishmërisë që kanë. Përzgjedhja është e bazuar në tri kritere.

Janë përfshirë të gjitha fjalët që paraqiten më së shpeshti në anglishte, bazuar në informacionin nga Korpusi Kombëtar Britanik BNC (angl. British National Corpus) dhe nga Koleksionin e Korpusit të Oksfordit (angl. Oxford Corpus Collection). Sidoqoftë, të jetë më e shpeshta në korpus nuk është e mjaftueshme që fjala të kualifikohet si fjalë-titull (angl. keyword): mund të jetë që fjala ka përdorim të dendur, por vetëm në fusha të caktuara, siç janë gazeta apo artikujt shkencorë. Që të mund të mënjanohej gjatë përfshirjes këto fjalë të kufizuara, na i fusim si fjalë-titull të gjitha fjalët të cilat kanë denduri të lartë në një rang (diapazon) të teksteve të ndryshme. Me fjalë të tjera, fjalët-titull janë fjalët të cilat kanë denduri të lartë dhe përdoren në kontekste të ndryshme. Mes tjerash, lista përfshin edhe disa fjalë që nuk kanë denduri të lartë, ndonëse u janë të njohura shumicës së anglishtfolësve. Këtu përfshihen, p.sh., fjalët të cilat emërtojnë pjesët e trupit, fjalët e përdorura në udhëtim, dhe fjalët të cilat janë të dobishme për të shpjeguar se çfarë kemi parasysh (çfarë mendojmë) atëherë kur nuk e dimë fjalën e saktë për diçka. Këto fjalë janë identifikuar duke konsultuar një panel prej më se shtatëdhjetë specialistëve nga fusha e mësimdhënies dhe e studimeve të gjuhës.

Me interes është edhe Fjalori bazë i anglishtes i Odgen-it me 850 fjalë [“Odgen's Basic English 850 words in his sequence”].

Lista e fjalëve akademike (të anglishtes) AWL

Një nga fushat më të rëndësishme të anglishtes në të cilën duhet të përqendrohen studentët ndërkombëtarë që planifikojnë të studiojnë në Universitetet e Amerikës Veriore është zgjerimi i vokabularit (leksikut). Ja disa statistika interesante me peshë të veçantë për studentët e huaj:

- Baza e vokabularit prej 2000 fjalësh të rëndomta në anglishte do t'u mundësojë atyre të kuptojnë afërsisht 80% të fjalëve në një tekst akademik.

- Baza e vokabularit prej 3000 fjalësh vetëm sa do ta rritë këtë në 85%.

- Meqë në anglishte janë afër 54.000 familje fjalësh, të kuptuarit e afër 3000-5000 prej tyre merret si një minimum për sukses në nivelin universitar.

“Lista e fjalëve akademike” (angl. Academic Word List, AWL) është një listë prej 3000-3500 fjalësh, që konsiderohen si të nevojshme për t'u mësuar nëse studentët e ardhshëm duhet t'i nënshtrohen testit TOEFL® iBT për të hyrë në një Universitet të Amerikës Veriore.

Edhe pse disa nga fjalët e përfshira në fjalorë nuk kanë pasur denduri të duhur për t'u përfshirë në fjalorin themelor, ose nuk kanë pasur shtrirje të duhur në të gjitha nënkorpuset (fjalët e nënvizuara në shembujt e mëposhtëm), ato megjithatë janë përfshirë në të sipas kriterit të funksionalitetit ose sipas kriterit të kompletimit të vargut kuptimor, si në shembujt e mëposhtëm: moti dhe koha: [ditë, natë, javë, muaj, vit, shekull; pranverë, verë, vjeshtë, dimër; e hënë, e martë, e mërkure, e enjte, e premte, e shtunë, e diel; janar, shkurt, mars, prill, maj, qershor, korrik, gusht, shtator, tetor, nëntor, dhjetor; ylber, shi, borë, vesë, erë, furtunë], përmasat: [gjatësi, gjerësi, saktësi, thellësi, trashësi, madhësi, sasi, pak, shumë, mjaft, plot, tepër], fjalët afirmative dhe përsëritëse [po, jo, mirë, keq, ndoshta, kënaqësi, faleminderit, besoj, kështu, sigurt, di], bimët dhe vendet ku rriten: [pyll, misër, oriz, pambuk, grurë, farë, li, tërshërë, bimë, fushë, kodër], krijimtaria: [bibliotekë, librari, autor, kritik, poet, gazetë, revistë, letërsi, roman, novelë, dramë, përrallë, poemë, strofë, tregim], veprimet [rris, korr, mbjell, shkul, bluaj, ngre, ul, tund, bëj], mjetet e punës [lopatë, kazmë, sëpatë, kosë], veshjet dhe tesha gjumi [bluzë, fund, pantallona, këmbishë, pallto, pizhame, xhaketë, këpucë, sandale, pantofla, fanellë, pulovër, shall, dorëza, çarçaf, jorgan, bataniye, jastëk, dyshek], ngjyrat themelore: [e bardhë, e kuqe, e verdhë, e zeze, bojëkafe, e hirtë, e gjelbër, e kaltër], mjetet e shkrimit: [laps, fletore, lodër], relacionet familjare:

[babë, nënë, motër, vëlla, gjysh, gjyshe, vajzë, bir, djalë, bijë, grua, burrë, bashkëshort, bashkëshorte,, xhaxha, teze, kushëri, fëmijë, nip, mbesë, vjehërr, dhëndër, nuse, kunat, foshnje, rini, pleqëri, fëmijëri], profesione dhe emërtime: [punëtor, shtëpiak, zdrukthëtar, murator, shofer, mekanik, bujk, fshatar, bari, ushtar, pilot, rrobaqepës, berber, kasap, mësues, teknik, inxhinier, aktor, piktor, ekonomist, doktor, dentist, këngëtar, deputet, sportist, shkrimtar, gazetar, artist, kryetar], mjete lundrimi dhe veprimet me to: [anije, barkë, lundër, aeroplan, biçikletë, autobus, automobil, lundrim, fluturim, notim, lundroj, fluturoj,, notoj, port, aeroport, pishinë, det, qiell, lumë, detar, pilot, notar], mjete bartëse: [valixhe, çantë, shporttë], hapësirat e banimit: [hotel, dhomë, banjë, krevat, shtrat, karrige, kat, korridor], ushqimet: [bukë, gjalpë, qumësht, mjaltë, djathë, reçel, vezë, supë, mish, gjellë, sallatë, peshk], orenditë, mjetet dhe enët e kuzhinës: [tryezë, pecetë, thikë, pirun, lugë, gotë, shishe, filxhan], frytet dhe pijet: [arrë, lajthi, gështenjë, kikirik, akullore, kafe, çaj, kakao, ujë, birra, limonadë, verë, raki], frutat dhe bimët ushqimore, mëlmesa: [kumbull, pjeshkë, kajsja, banane, rrush, qershi, pjepër, shalqi, uthull, kripë, sheqer, piper, mollë, dardhë, portokall, limon], perimet dhe bimët barishtore ushqyese: [perime, patate, fasule, lakër, qepë, hudhër, karotë, ulli, spinaq, trangull, domate], veglat muzikore: [fyell, daulle, lahutë, kitarë], pjesët e brendshme dhe të jashtme të trupit të njeriut: [kokë, fytyrë, flok, ballë, sy, vetull, qepallë, qerpik, bel, vithe, gjymtyrë, krah, vesh, hundë, buzë, dhëmb, faqe, mjekër, zverk, grykë, mollëz, shpatull, sup, llërë, sqetull, bërryl, dorë, grusht, gisht, thua, kofshë, gju, shputë, zemër, mushkëri, stomak, zorrë, mëlçi, tëmth, veshkë, skelet, nofull, shprekë, gjak, venë, arterie, qelizë, mish, muskul, dell]

Jemi munduar që kualifikimin e fjalëve për fjalësin e fjalorit themelor 5000-fjalësh dhe të fjalorit 2500-fjalësh ta bëjmë sipas Fjalorit drejtshkrimor të gjuhës sotme shqipe [FDGJSH,1976]. Rregullat të cilave ju kemi përmbajtur po i japim më poshtë:

1. I kemi marrë homonimet që përputhen fonetikisht në trajtën përfaqësuese, por që dallohen në trajtat e tjera (homofomat), jepen si fjalë të veçanta edhe atëherë kur i përkasin së njëjtës pesë të ligjëratës, p.sh.: pres (preva), pres (prita), verë (stinë), verë (pije)
2. Ndryshe nga fjalori janë marrë si një zë emrat dhe mbiemrat homonimik: plak m/mb
plakë f/mb
3. Si fjalë të veçanta nuk janë marrë: homonimet që i përkasin së njëjtë pjesë të ligjëratës dhe që përputhen si në trajtën përfaqësuese, ashtu edhe në trajta tjera, jepen vetëm një herë, pavarësisht nga kuptimet e ndryshme që kanë (p.sh. pjek fol. (poqa, ~ur), pjek fol.(takoj).
4. Homonimet që u përkasin pjesëve të pandryshueshme të ligjëratës jepen vetëm një herë (p.sh.: drejt (ndajf), drejt (parafj), dhe (lidh), dhe (pj) por dhe (kore e tokës) dhe dhe (jap) jepen si zë i veçantë.
5. Ndryshe nga fjalori drejtshkrimor kemi marrë si zë të veçantë ndajfoljet që kanë të njëjtën temë me mbiemrat e nyjshëm përkatës: keq - i keq, mirë – i mirë, bukur – i bukur.
6. Trajtat e diatezës pësore ose vetvetore nuk jepen më vete por së bashku me trajtat verore përkatëse (p.sh.: afro [afrohem], lag [lagem]).
7. Jepen si zë më vete vetëm foljet që kanë vetëm diatezë vetvetore (endem, grindem, kollem, përgjigjem).
8. Nuk jepen si zë të veçantë trajtat pavetore të tipit: ecet, flihet, rrihet,
9. Trajtat supletive të pjesëve të ndryshme të ligjëratës dhe trajtat e lakimit të përemrave nuk jepen më vete por s’ bashku me trajtën përfaqësuese.
10. Po ashtu nuk janë dhënë veçantë as trajtat e parregullta që ndryshojnë shumë nga trajtat përfaqësuese të fjalës: kalë/kuaj, dua/doja, ka/qe.
11. Pjesët e ndryshueshme të ligjëratës jepen në trajtën e tyre përfaqësuese: emrat – në emëroren e pashquar të njëjësit ose të shumësit (kur përdoren vetëm në shumës); mbiemrat, numërorët rreshtorë e përemrat - në trajtën e gjinisë mashkullore të njëjësit ose në të dyja gjinitë, kur tema nuk ndryshon sipas gjinisë; foljet – në vetën e parë njëjës të së tashmes së dëftores

(ose në vetën e tretë njëjës, kur janë pavetore a njëvetore).

12. Janë marrë si zë më vete fjalët e përbëra që shkruhen me vizë lidhëse, p.sh.: organizatë-bazë, vinç-urë, bujqësor-industrial.

13. Shënohen te fjalët që gjatë lakimit a zgjedhimit ndodhin ndryshime tingullore që në rrokjen e parë të temës së fjalës, atëherë këto trajta jepen në kllapa, p.sh.: natë (net), pres (prita), cirk (cirqe),

Në fjalor nuk janë përfshirë numërorët: njëmbëdhjetë, dymbëdhjetë, trembëdhjetë, katërbëdhjetë, pesëmbëdhjetë, gjashtëmbëdhjetë, tetëmbëdhjetë, nëntëmbëdhjetë, tridhjetë, pesëdhjetë,

gjashtëdhjetë, shtatëdhjetë, tetëdhjetë, nëntëdhjetë, njëqind, dyqind, treqind etj., të cilët mund të ndërtohen në bazë të analogjisë, siç mund të formohen numërorët e tipit mijë nga numërorët themelorë dhe fjala mijë. Po ashtu nuk janë përfshirë numërorët rreshtor që kanë trajtë të njëjtë.

Janë marrë vetëm si emra apo si mbiemra edhe pse hasën në të dy përdorimet fjalët si: armik, fqinj, mysliman, mik, (i) panjohur, atdhetar, bujar, fatos, përgjegjës, ortodoks, zyrtar etj.

Po ashtu nuk janë marrë fjalë të veçanta fjalët si: afër, andej që përdoren si ndajfolje dhe si parafjalë.

	Fjalori 2500-fjalësh		Fjalori 5000-fjalësh	
	Shuma	Përqindja	Shuma	Përqindja
Emra	1374	54.96%	2864	57.28%
maskullorë	657	47.81%	1491	52.06%
femërorë	717	52.19%	1373	47.94%
asnjanës	0	0%	0	0%
Mbiemra	332	13.28%	785	15.7%
të nyjshëm	266	80.12%	427	54.39%
të panyjshëm	66	19.88%	358	45.60%
Folje	465	18.6%	839	16.78%
kalimtare	382	82.36%	709	84.62%
jokalimtare	82	17.64%	129	15.38%
Ndajfolje	187	7.48%	285	5.7%
Të tjera	142	5.68%	227	4.54%
Gjithsej	2500	100%	5000	100%

Tabela 1 Struktura morfologjike e fjalorëve themelorë 2500-fjalësh dhe 5000-fjalësh

Kur një fjalë mund të zëvendësojë plotësisht një fjalë tjetër me denduri më të vogël, atëherë kemi marrë vetëm fjalën me denduri më të madhe si p.sh. janë marrë: ende (422), prandaj (426), baba (269), madje (358) - e nuk janë marrë fjalët: akoma (163), andaj(35), babë (72), babi (44), bile

(102) që kanë denduri më të vogël, si dhe për arsye stilistike e etimologjike: akoma – greqizëm, bile – turqizëm. Për të njëjtat arsye janë marrë shkronjë (72) e jo edhe germë (22), çfarë (45) dhe çka (404) por jo dhe ç'(577).

Janë marrë çiftet emër mashkullor – emër femëror: gjysh – gjyshe, plak – plakë, im- ime, or – ore.

Fjalori themelor 2500 – fjalësh përmban edhe mjaft mbiemra prejpsesor, shumicën e të cilëve mund të mos i përfshijmë në një fjalor me më pak se 2500 fjalë, meqë lehtë nxirren nga foljet (pjesoret) përkatëse, siç mund të shihet nga tabela.

ANALIZA E REZULTATEVE

Është me interes të shihet se cila është struktura leksikore, po dhe morfologjike e semantike e këtyre fjalorëve.

Në tabelën 1 është paraqitur struktura morfologjike e këtyre dy fjalorëve.

Krahasimi i 850 fjalëve të para të fjalorit 2500-fjalësh me fjalorin 850-fjalësh të anglishtes (C.G.Ogden) tregon një përputhshmëri të lartë ndërmjet tyre, dhe shumica e tyre paraqitet në krye të listës 10000 fjalët e para të renditura sipas dendurisë.

Mund të konstatojmë se fjalorët janë mjaft të baraspeshuar edhe për sa u përket antonimeve të përfshira në ta. Kështu hasim çiftet e antonimeve, si:

i ardhshëm i shkuar
i bardhë i zi
i drejtë i shtrembër
i egër i butë
i ëmbël i hidhur
i ftohtë i ngohtë
i gjallë i vdekur
i gjatë i shkurtër
i hollë i trashë
i majtë i djathtë
i mençuri marrë
i mirë i keq
i ngushtë i gjërë
i pastër i ndotur
i ri i vjetër
i saktë i gabuar
i shëndoshë i sëmurë
i thjeshtë i përbërë
i ulët i lartë
i zgjuar i fjetur
femëror mashkullor

fort dobët
herët vonë
lartposht
pak shumë
përpara mbrapa
thatë lagësht

dashuri urretje
dritë terr
gëzim hidhërim
gazvaj
lindje vdekje
liri robëri
etj.

LITERATURA

- [Agalliu etj. 2002], Agalliu, Fatmir etj, Gramatika e gjuhës shqipe I, ASHSH, IGJL, Tiranë, 2002.
- [Altenberg 2002], Altenberg B. and Granger S. (eds) (2002) Lexis in Contrast. Corpus-based Approaches. Studies in Corpus Linguistics 7. Amsterdam & Philadelphia: Benjamins (337 pp.).
- [Ball 1996], Ball, N. Catherine (1996), Concordances and Corpora, Corpus linguistics course, Department of Linguistics, Georgetown University, Washington DC, <http://www.georgetown.edu/ball/corpora/tutorial.html>].
- [Basha 1979], Basha, Ylvi, Fjalor themelor Rusisht – Shqip; Shtëpia botuese e Librit shkollor, Tiranë, 1979.
- [Buk 2004], Buk, Solomiya, Lexical base as a compressed language model of the world (on the material of the Ukrainian language), 2004.
- [Caka 1986], Caka, Nebi, Analizë statistike dhe stilistike e gjuhës së Lasgush Poradecit, punim magjistrature, mentor Rexhep Ismajli, Prishtinë, 1986.
- [Cherata 1997], Cherata, Sanda, CONCORD - Software System for Concordances of Romanian Poetical Texts- Tufis, Dan and Andersen Poul, 1997, Recent Advances in Romanian Language Technology.

8. [Church 1995], Church, K. and Gale, W. (1995), "Poisson Mixtures", *Journal of Natural Language Engineering* 1:2
9. [Davies 2002], Davies, Mark, *Corpus del Español*, 2002, www.corpusdelespanol.org
10. [Davies 2004], Davies, Mark, *Vocabulary Range and Text Coverage: Insights from the Forthcoming – Routledge Frequency Dictionary of Spanish*.
11. [Davies 2005], Davies, Mark, *Selected Proceedings of the 7th Hispanic Linguistics Symposium*, ed. David Eddington, 106-115. Somerville, MA: Cascadilla Proceedings Project, 2005.
12. [Francic-Kucera 1964], Francis, W. Nelson, Henry Kucera, *Brown corpus manual of information (Manual to accompany A Standard Corpus of Present-Day, Edited American English, for use with Digital Computers)*, Brown University, 1964
13. [Granger 2002], Granger S., Dagneaux E. and Meunier F. (2002) *The International Corpus of Learner English. Handbook and CD-ROM*. Louvain-la-Neuve: Presses Universitaires de Louvain (162 pp.).
14. [Granger 2002a], Granger S., Hung J. and Petch-Tyson S. (eds) (2002) *Computer Learner Corpora, Second Language Acquisition and Foreign Language Teaching. Language Learning and Language Teaching 6*. Amsterdam & Philadelphia: Benjamins
15. [Granger 2002b], Granger, Sylviane, *International Corpus of Learner English – ICLE, Centre for English Corpus Linguistics – CECL,*
16. [Granger 2003], Granger S. and Petch-Tyson S. (eds) (2003) *Extending the Scope of Corpus-based Research: New Applications, New Challenges*. Amsterdam & Atlanta: Rodopi (261 pp.).
17. [Granger 2003b], Granger S., Lerot J. and Petch-Tyson S. (eds) (2003) *Corpus-based Approaches to Contrastive Linguistics and Translation Studies*. Amsterdam & Atlanta: Rodopi (219 pp.).
18. [Hysa 1970], Hysa, Enver (1970): *Formime me prapashtesa, në shqipen e sotme letrare*, *Studime filologjike*, 1, 1970.
19. [Hysa 1973], Hysa, Enver (1973): *Disa çështje të prapashtesimit në gjuhën shqipe*, *Studime filologjike*, 3, 1973.
20. [Hysa 1985], Hysa, Enver. (1985): *Dukuri e prirje të formimit të fjalëve në shqipen e sotme letrare*, *Gjuha jonë*, 1, 1985.
21. [Ismajli 1987], Ismajli, Rexhep, *Artikuj mbi gjuhën shqipe*, Rilindja, Prishtinë, 1987.
22. [lusi 2004], Francisco, lusi, *Konkordancat elektronike të teksteve letrare, leksikografia shqipe – trashëgimi dhe perspektivë*, Konferencë shkencore, Tiranë, 3 dhjetor 2004
23. [Kokona 1972], Kokona, Vedat, *Fjalor themelor Frëngjisht – Shqip; Shtëpia botuese e Librit shkollor*, Tiranë, 1972.
24. [Kostallari 1961], Kostallari, Androkli. (1961): *Mbi disa veçori të fjalës së përbërë në gjuhën shqipe*, *BUSHt*, 3, 1961.
25. [Kostallari 1972], Kostallari, Androkli (1972): *Në rrugën e hartimit të fjalorit normativ të shqipes së sotme, "Studime mbi leksikon" II*, Tiranë, 1972, f. 81.
26. [Kostallari 1972a] Kostallari, Androkli (1972): *Parimet themelore për hartimin e "Fjalorit të gjuhës së sotme shqipe"*, në *"Studime mbi leksikon dhe mbi formimin e fjalëve në gjuhën shqipe" II*, Tiranë, 1972, f. 3.
27. [Kostallari etj. 1980], Grup autorësh, *Fjalori i gjuhës së sotme shqipe, I, II*, ASHSH. IGjL, Tiranë, 1980.
28. [Lehman 1989], Lehman Hubert, *Towards a core vocabulary for natyral language system*, 1989.
29. [Lönngren 1993], Lönngren, Lennart (ed.), 1993. *Chastotnyj slovar' sovremennogo russkogo jazyka. (A Frequency Dictionary of Modern Russian. With a Summary in English.) Acta Universitatis Upsaliensis, Studia Slavica Upsaliensia*
30. [McEnery 1996], McEnery, Tony and Andrew Wilson, *Corpus Linguistics*, Edinburgh: Edinburgh University Press, 1996
31. [Meyer 2002], Meyer, Charles F., *English Corpus Linguistics – An Introduction*, Cambridge University Press, 2002.
32. [Philippe 2006], Philippe Blache, Marie-Laure Guenot & Cristel Portes *Laboratoire Parole et Langage – CNRS /Universite de Provence*, 29 av.

Robert Schuman 1361 Aix en Provence cedex, {pb; mug; portes}@lpl.univ-aix.fr

33. [Samara 1998], M. Samara, Fjalori i antonimeve në gjuhën shqipe, IGJL, Tiranë, 1998.

34. [Sharoff 2002], Sharoff, Serge, (2002). "Meaning as use: exploitation of aligned corpora for the contrastive study of lexical semantics". Proc. of Language Resources and Evaluation Conference (LREC02). May, 2002, Las Palmas, Spain.

35. [Sinclair 1996], Sinclair, Jones, EAGLES Preliminary recommendations on Corpus Typology, EAG--TCWG--CTYP/P, School of English, University of Birmingham, 1996

36. [Stubs 1986], Michael Stubbs, Language development, lexical competence and nuclear vocabulary, Kevin Durkin, ed (1986) Language Development in the School Years. Croom Helm..

37. [Thomai 1981], Thomai, Jani (1981): Tiparet themelore të derivacionit semantik në gjuhën shqipe, Studime filologjike, 3, 1981.

38. [Thomai 1983], Thomai, Jani (1983): Formimi i fjalëve dhe prejardhja kuptimore në gjuhën shqipe, Studime filologjike, 3, 1983.

39. [Thomai 1989], Thomai, Jani (1989): Prejardhja semantike në gjuhën shqipe, Tiranë, 1989.

40. [Thomai 1992], Thomai, Jani (1992): Mbi mënyrën e pasqyimit të sistemit të kuptimeve dhe të përdorimit të fjalëve në fjalorin shpjegues të gjuhës shqipe, "Studime mbi leksikon" II, Tiranë, 1992, f. 129.

41. [Thomai 1992a], Thomai, Jani (1992): Teksti dhe gjuha, Tiranë, 1992.

42. [Thomai 1999], Thomai, Jani (1999): Leksikologjia e gjuhës shqipe, SHBLU, Tiranë, 1999.

43. [Thomai etj.], J. Thomai, M. Samara, H. Shehu, Th. Feka....Fjalori sinonimik i gjuhës shqipe, ASHAK. IGJL, Tiranë..

44. [Thomaj 2005], Leksikografija shqipe, Jani Thomaj, Trashëgimi dhe perspektivë.

45. [Xhuvani 1980], Xhuvani, Aleksandër (1980): Vepra, I, 1980. LITERATURA