

Tarpfrazinių pauzių trukmė skaitomoje kalboje

Asta Kazlauskienė

Vytauto Didžiojo universiteto Humanitarinių mokslų fakultetas
Vytautas Magnus University Faculty of Humanities
V. Putvinskio g. 23-204, LT-44243 Kaunas, Lietuva
E. paštas asta.kazlauskieni@vdu.lt
ORCID iD: <https://orcid.org/0000-0002-6135-1635>

Moksliniai interesai: fonetika, fonologija, fonotaktika, prozodija, akcentologija, taikomoji kalbotyra
Scientific interests: phonetics, phonology, phonotactics, prosody, accentology, applied linguistics

Eidmantė Kalašinskaitė

Vytauto Didžiojo universiteto Humanitarinių mokslų fakultetas
Vytautas Magnus University Faculty of Humanities
V. Putvinskio g. 23, LT-44243 Kaunas, Lietuva
E. paštas eidmante.kalasinskaite@stud.vdu.lt
ORCID iD: <https://orcid.org/0000-0002-8646-1654>

Moksliniai interesai: garsinės kalbos ypatybės
Scientific interests: features of spoken language

Santrauka. Šio tyrimo tikslas – nustatyti pauzių tarp pastraipų ir sakinių trukmę skaitomoje kalboje ir jų sąsają su šių kalbos vienetų ilgiu bei skaitymo greičiu. Tyrimo medžiagą sudaro 48 garso įrašai (aštuoni diktoriai skaitė šešis skirtingų funkcinį stilių tekstus). Tyrimo rezultatai rodo, kad pauzės tarp pastraipų paprastai daro visi skaitantieji balsu, o jų trukmė priklauso nuo skaitymo greičio: greičiau skaitančiųjų pauzės yra vidutiniškai nuo 400 ms iki 1340 ms, lėčiau skaitančiųjų – nuo 750 ms iki 1800 ms. Po trumpų pastraipų pauzės yra trumpiausios, o po ilgesnių pastraipų jos yra 1,2 karto ilgesnės. Pauzės tarp skaitomų sakinių taip pat beveik visada daromos. Greičiau skaitančiųjų pauzės vidutiniškai yra nuo 200 ms iki 980 ms, lėčiau skaitančiųjų – nuo 340 ms iki 1360 ms. Ilgejant sakiniams, ilgeja ir pauzės po jų (iki 1,4 karto). Tačiau po ilgų sakinių jos gali būti tokios pačios trukmės, kaip ir po vidutinių sakinių.

Reikšminiai žodžiai: pauzė, sakinytis, pastraipa, trukmė.

Tyrimas finansuotas Europos socialinio fondo lėšomis pagal priemonę Nr. 09.3.3-LMT-K-712 „Mokslininkų, kitų tyrėjų, studentų mokslinės kompetencijos ugdymas per praktinę mokslinę veiklą“.

The Duration of Pauses between Phrases in Reading Language

Summary. The aim of this research is to determine the duration of the pauses between paragraphs and sentences in reading language and their connection with the length and reading speed of said language units. The research material consists of 48 sound recordings (eight Lithuanian speakers read six texts representing different functional styles). The results show that all readers usually pause between paragraphs and the duration of those pauses depend on the reading speed: faster readers pause for approximately 400–1340 ms, while slower readers pause for 750–1800 ms. Pauses are the shortest after short paragraphs and they are 1.2 times longer after longer paragraphs. All the readers usually pause between sentences as well. Faster readers on average pause for 200–980 ms, slower readers pause for 340–1360 ms. The longer the sentence, the longer the pause after that sentence (up to 1.4 times). However, the pause can be the same duration after both long and medium length sentences.

Keywords: pause, sentence, paragraph, duration.

Įvadas

Pauzės yra daugelio mokslo sričių atstovų tyrimo objektas: psichologų, psicholingvistų, sociolingvistų, fonetikų, iškalbos meno tyrėjų ir praktikų ir kt. Jos padeda skaidyti tekstą į semantinius, gramatinius, intonacinius vienetus, siekti kalbos ekspresijos, pagaliau jos svarbios ir fiziologiškai – suteikia galimybę įkvėpti oro.

Užsienio mokslininkų darbuose pauzės dažnai nagrinėjamos tiriant kalbėjimo greitį (Benki et al. 2011; Trouvain et al. 2011), jos siejamos ir su emocijų raiška (Beller et al. 2006). Pauzės kaip vienas iš tyrimo objektų (kartu su kalbėjimo ir artikuliacijos greičiu) įtraukiamos į kai kurių fonetinių segmentų (pavyzdžiui, priebalsių skardumo) suvokimo tyrimus (Miller, Grosjean 1981). Jos gali padėti planuoti diskursą, reikšti emocijas (Viola, Madureira 2008). Į pauzes atsižvelgiama ir analizuojant užsienio kalbos sklandumą – greta kalbėjimo greičio ir pasitaisymų jos laikomos vienu iš sklandaus kalbėjimo svetimąja kalba rodiklių (Bosker et al. 2014). Pauzių tyrimai peržengia lingvistikos ribas, jos įdomios ir rinkodaros specialistams, stebintiems, kaip skirtingas reklamos greitis veikia vartotoją (Megehee et al. 2003). Be to, kitų šalių mokslininkai į pauzes kalboje atkreipia dėmesį tirdami įvairias ligas, tokias kaip depresija, Parkinsono, Alzheimerio ligos ir pan. (Mendiratta et al. 2018; Gosztolya et al. 2018; Cannizzaro et al. 2004). Pauzės aktualios ir siekiant užtikrinti sintezuotos kalbos natūralumą (Cahn 1990). Plačiausiai ir sistemiškiausiai pauzės yra tyrinėtos B. Zellner (1994) ir T. S. Kendallo (2009). B. Zellner (1994), lygindama spontaniinę ir skaitomąją kalbą, nustatė, kad sudėtingesnė komunikacinė situacija reikalauja daugiau pauzių, spontaniinėje kalboje jos dažnesnės nei skaitomoje, pauzės būna ilgesnės tarp nelabai susijusių žodžių ir žymiai trumpesnės bei retesnės tarp semantiškai ir gramatiškai susijusių žodžių. T. S. Kendallo (2009) darbe analizuojamas pauzių ir kalbėjimo greičio kitimas atsižvelgiant į kalbėtojo tautybę, lytį ir regioną, kuriame gyvena. Pastebėta, kad priklausymas tam tikrai etninei grupei daro įtaką pauzėms, pavyzdžiui, afroamerikiečiai daro trumpesnės pauzes kalboje nei kitų etninių grupių atstovai. Užsienio mokslininkų darbuose analizuojamos tiek spontaniinės, tiek skaitomosios kalbos pauzės, dažniausiai nagrinėjama jų trukmė, dažnumas, vieta ir funkcijos kalbos sraute.

Lietuvių kalbotyroje pauzės nedaug nagrinėtos, dažniausiai aptariamoms ar epizodiškai tiriamoms analizuojant intonacinę sistemą, kalbos vienetų ilgį ar kalbėjimo greitį. A. Kazlauskienė ir K. Veličkaitė (2003), tirdamos lietuvių kalbėjimo greitį iš dalies spontanišoje kalboje (diktoriai galėjo pasiruošti, ką kalbės, tačiau neskaitė užrašyto teksto), pastebėjo, kad, kalbant apie praeities įvykius, kai reikia kažką prisiminti, daromos ilgesnės pauzės ir dėl jų greitis lėtėja. G. Kundrotas (2010) tyrė lietuvių kalbos intonacinius kontūrus, o pauzės yra vienas iš intonacinės sistemos sudedamųjų elementų. I. Balčiūnienė ir L. Kamandulytė-Merfeldienė (2016) analizavo sakininės lietuvių kalbos pasakymų ilgį ir struktūrą. Sakytinės kalbos tyrimuose pauzės svarbios kaip pasakymo ribų nuorodos, tačiau autorės pastebėjo, kad pasakymu laikyti kalbėjimo atkarpą iki pauzės ne visada tikslinga dėl minties šuolių, kalbos nutrūkimo. Apie pauzes rašo ir oratorystės tyrėjai, nes jos padeda sutelkti klausytojo dėmesį, laikytis tinkamo ritmo, taisyklingai kvėpuoti (Koženiauskienė 1999). Reikia paminėti, kad visuose aptartuose lietuvių kalbos tyrimuose pauzė nėra apibrėžiama, nekalbama, kokios trukmės tylą reikėtų laikyti pauze, o kompleksiskai, kaip atskiras tyrimo objektas, pauzės lietuvių kalboje nenagrinėtos.

Šis straipsnis – tęstinio lietuvių kalbos pauzių tyrimo dalis. Tyrimu siekiama išanalizuoti pauzių atsiradimą kalboje ir jų trukmę bei padėti suformuluoti principus, kuriais remiantis būtų galima modeliuoti sintezuotos (dirbtinės kalbos, t. y. rašytinio teksto perskaitymo balsu, naudojant specialią programinę įrangą) kalbos pauzes, kurios suteiktų natūralumo. Pauzės aktualios ir atvirkštinuose šnekos technologijų produktuose – kalbos atpažinime (garsinės kalbos užrašymas tekstu): pauzės gali būti vienas iš signalų rašyti skyrybos ženklą. Pradiniame etape tirta pauzių trukmė skaitomuose trumpuose tekstuose ir nustatyta tendencija, kad pauzės beveik visada atsiranda ten, kur tekste yra aiški vizuali atskirtis (tarp pastraipų, sakinių), o kitais atvejais jos yra fakultatyvios (Kazlauskienė, Kalašinskaitė 2017). Tolesnį skaitomos kalbos pauzių tyrimą nutarta atlikti su ilgesniais tekstais ir jį suskaidyti į du etapus: a) nagrinėti pauzių tarp pastraipų ir sakinių trukmę (jos šiame straipsnyje vadinamos tarpfrazinėmis pauzėmis) ir b) tirti pauzes skaitomo sakinio viduje. Pirmuoju atveju pauzės daromos dėl teksto segmentavimo ir jos yra pagrindinis stambiųjų teksto segmentų ribų rodiklis. Antruoju atveju pauzės turi daugiau funkcijų (prozodinių, paralingvistinių ir kt.), be to, frazės viduje skaidymas į semantinius, sintaktinius, prozodinius vienetus dažnai gali būti paremtas ne tik pauzėmis, bet ir pagrindinio tono kaita. Todėl šiuos du svarbius garsinės kalbos elementus pravartu analizuoti drauge.

Šiame straipsnyje aprašomo tyrimo tikslas – nustatyti pauzių tarp skaitomo teksto pastraipų ir sakinių trukmę ir jų sąsają su šių kalbos vienetų ilgiu ir skaitymo greičiu. Tyrimo tikslui pasiekti kelti keli uždaviniai: a) nustatyti diktorių skaitymo greitį, b) nustatyti pauzių kiekį, c) išsiaiškinti pauzių trukmės ir pastraipų bei sakinių ilgio bei pauzių trukmės ir skaitymo greičio sąsajas, d) pateikti preliminarias rekomendacijas kalbos technologijų kūrėjams. Būtent dėl pastarojo uždavinio šis tyrimas skiriasi nuo kitų užsienio ir lietuvių lingvistų darbų.

Tyrimo metodologija

Tiriamąją medžiagą sudarė 48 garso įrašai, kurių bendra trukmė – 7 val. 36 min. Eksperimente dalyvavo aštuoni diktoriai: keturi vyrai ir keturios moterys, 20–25 m. amžiaus, septyni informantai yra studentai, vyrai studijuoja nehumanitarinius mokslus, moterys – humanitarinius, viena moteris turi aukštąjį nehumanitarinį išsilavinimą. Eksperimento dalyviai nėra profesionalūs diktoriai, jų skaitymo balsu įgūdžiai skiriasi. Būtina pabrėžti, kad ir neturėta tikslo įrašyti tik išraiškingai skaitančius profesionalius diktorius. Tokį pasirinkimą lemia tyrimo tikslas – kalbos atpažinimo programos orientuotos ne į pavienius kalbos vartotojus, o plačiąją kalbinę bendruomenę.

Diktoriai perskaitė šešis tekstus: ištrauką iš Balio Sruogos knygos „Dievų miškas“, Broniaus Radzevičiaus novelę „Laukais nubėga šešėliai“, dvi ištraukas iš Aistės Urbonienės mokslinio straipsnio „Moters mokslininkės įvaizdis Lietuvos moksleivių akimis“, du publicistinius straipsnius, paskelbtus naujienų portale 15min.lt: Laurynos Vireliūnaitės „Gydytojų rezidentų protestas: jaunieji medikai, bandydami išgyventi, griebiasi net egzotiškų papildomų darbų“ ir „Specialistai įspėja: bananai – ant išnykimo ribos“ (autorius nenurodytas). Bendra kiekvieno stiliaus tekstų apimtis yra apylygė. Šie tekstai pasirinkti dėl kelių priežasčių: 1) struktūros ypatumų: jie yra netrumpi, suskaidyti į nemažai pastraipų ir sakinių, 2) kalbinės įvairovės: reprezentuoja skirtingus funkcinis stilius su jiems būdinga sakinio struktūra ir leksika, jie gali turėti įtakos ir skaitymo greičiui, ir pauzių kiekiui bei jų trukmei, nors pauzacijos ir funkcinio stiliaus sąsaja nėra pagrindinis šio tyrimo objektas.

Remiantis diktorių įrašais, tekstai parengti automatiniam anotavimui (sukirčiuoti, sužymėtas žodžių šlijimas, frazių ribos, įvairūs paralingvistiniai reiškiniai, pvz., atskrenkštimas ir pan.). Tai padaryta vadovaujantis VDU mokslininkų taikoma metodologija ir jų pasirengtais įrankiais (plačiau žr. Kazlauskienė, Raškinis 2013: 104). Įrašai buvo automatiškai anotuoti, t. y. rašytinis tekstas susietas su garsiniu įrašu (nustatytos frazių, žodžių, garsų ribos ir rašytiniame tekste, ir įrašė)¹. Automatiškai anotuota tiriamoji medžiaga buvo rankomis tikrinama: tikslinamos ribos tarp žodžių, dirbta su garsų analizės programa PRAAT (Boersma, Weenink 2018). Sutvarkius medžiagą, buvo rankomis nustatyta pauzių trukmė: matuota tylos momento tarp žodžių, esančių abipus sakinio pabaigos skrybės ženklų, akustinė trukmė.

Tarpfrazinių pauzių trukmė skaitomoje kalboje

1. Diktorių skaitymo greitis

Skaitymo greičiui įtakos turi daug veiksnių: temperamentas, išsilavinimas (vidurinis ar aukštasis, humanitarinis ar gamtamokslinis), skaitymo įgūdžiai, teksto ypatumai (pasiviosios vartosenos žodžių gausa, retesnės morfologinės formos, sudėtingos sintaksinės konstrukcijos), pagaliau epizodiniai paralingvistiniai veiksniai (nuotaika, fiziologinė būseną) ir kt. Visiems šiems veiksniams aprėpti ir jų koreliacijai su skaitymo greičiu

¹ Autorės dėkoja kolegai prof. dr. Gailiui Raškiniui už pagalbą tvarkant empirinius duomenis.

nustatyti reikėtų atskiro plataus tyrimo su gerokai didesne tiriamųjų imtimi. Šiam tyrimui skaitymo greitis aktualus dėl to, kad jis tiesiogiai lemia kalbinių segmentų (taip pat ir pauzių) trukmę, o frazės viduje ir pauzių kiekį (Kazlauskienė, Kalašinskaitė 2017). Todėl svarbu išsiaiškinti diktorių skaitymo greitį ir juo remiantis analizuoti bei interpretuoti empirinius tarpfrazinių pauzių trukmės duomenis. Net ir skirtingų funkcinių stilių tekstų pasirinkimas šiuo atveju, kaip jau minėta, buvo tik pagalbinis aspektas siekiant nustatyti bendruosius tarpfrazinių pauzių trukmės dėsningumus. Todėl lingvistiniai tekstų skirtumai, paralingvistiniai ir ekstralingvistiniai veiksniai išsamiai šiame tyrime nenagrinėjami.

Skaitymo greitis nustatytas kiekvieno teksto žodžių kiekį padalinus iš garso įrašo trukmės. Iš 1 lentelėje pateiktų duomenų matyti, kad vidutinis diktorių skaitymo greitis – 117 žodžių per minutę.

1 lentelė. Diktorių skaitymo greitis (ž. / min)²

Tekstas	Žodžių kiekis	Pastraipų kiekis	Sakinių kiekis	Diktoriai								Vidurkis
				M1	M2	M3	M4	V1	V2	V3	V4	
G1	703	13	74	135	136	153	140	108	98	113	113	125
G2	1276	11	112	133	135	137	138	98	96	111	113	120
M1	987	12	41	120	119	121	122	98	85	104	103	109
M2	854	12	36	120	118	130	134	101	91	103	95	112
P1	1246	38	60	133	124	136	135	110	101	101	110	119
P2	912	20	61	125	124	130	129	100	96	109	112	117
Vidurkis				128	126	135	133	103	95	107	108	117

Iš 1 lentelės matyti, kad moterys diktorės perskaitė daugiau žodžių per minutę nei vyrai (moterys nuo 126 iki 135, vyrai nuo 95 iki 108). Vadinas, diktores moteris šiame tyrime galima būtų laikyti greitakalbėmis, o vyrus – lėtakalbiais, nors toks skirstymas, turint nedidelę imtį, yra labai sąlygiškas. Dėl šios priežasties negalima daryti kokių nors patikimesnių išvadų dėl individualių savybių, be to, jos ir nebuvo tyrimo objektas. Būtina paminėti, kad bendra įrašų trukmė ne visada yra vienareikšmis skaitymo greičio rodiklis: klausant įrašų girdėti, kad M1 skaito greičiau nei M2 ar M4. Ilgesnę jos įrašų trukmę ir mažesnį skaitymo greitį nulėmė gana dažni žodžio dalių pakartojimai, taip pat paralingvistiniai veiksniai (diktorė skaitydama neretai kosčiojo). Tokių reiškinių, žinoma, pasitaikė ir kitų diktorių įrašuose, tačiau gerokai mažiau. Imtis per maža daryti išvadas ir dėl išsilavinimo ir skaitymo greičio sąsajos. Galima tik paminėti, kad vyrai, studijuojantys nehumanitarinius mokslus, skaitė lėčiau, nors nehumanitarinį išsilavinimą turinti M4 skaitė gana greitai.

² Šioje lentelėje diktorių skiltyse M raide pažymėta moterys diktorės, V – vyrai, tekstų skiltyje G – grožiniai, M – moksliniai, P – publicistiniai tekstai.

Visi diktoriai mokslinius tekstus perskaitė šiek tiek lėčiau nei grožinius ar publicistinius. Tam, matyt, įtakos turėjo vyraujantys ilgesni mokslinių tekstų sakiniai (vidutiniai ir ilgi sakiniai sudarė 70 proc., kai grožiniuose tekstuose tik 10 proc., publicistiniuose – 34 proc.) ir retesni, o gana dažnai ir ilgesni žodžiai, nors parinktas mokslinis straipsnis skaitymui balsu nesudėtingas (nėra formuliu, grafinių iliustracijų ir pan., kurie šį tekstą labai išskirtų iš kitų skaitytų). Vis dėlto skaitymo greičio skirtumai nėra dideli ir daugeliu atvejų santykis tarp greičiausiai ir lėčiausiai perskaityto fragmento yra tik 1,1–1,2 karto.

2. Tarpfrazinių pauzių kiekis garso įrašuose

Skaitytuose šešiuose tekstuose vienas diktorius tarp pastraipų galėjo padaryti 100 pauzių, tarp sakinių – 275 pauzes (šiuo atveju pauzės po pastraipos neįskaičiuotos). Kaip jau minėta, pauzių ribos tikslintos rankomis ir medžiagos rankinis apdorojimas leido daryti prielaidą: kaip pauzė gali būti suvokiamas jau 20 ms (ir, žinoma, ilgesnis) tylos momentas, kuris ir psichoakustiškai girdimas kaip tyla, ir oscilogramoje bei spektrogramoje nefiksuo-jama akustinė energija. Kita vertus, reikia paminėti, kad kai kada rentasi ir pagalbinais kriterijais:riebalsių asimiliacijos reiškiniais, pagrindinio tono kaita. Tiesa, šie kriterijai aktualesni fiksuojant pauzes frazių viduje. Tarpfrazines pauzes nebuvo sudėtinga nustatyti ir matuoti, nes daugeliu atvejų jos gerai girdimos ir matomos garsų analizės programoje.

Iš 2 lentelės matyti, kad po visų pastraipų pauzės yra beveik visų diktorių įrašuose. Pauzės nebuvo padarytos tik po trijų pastraipų, visos jos yra iš publicistinių tekstų. M1 ir M2 nedarė pauzių prieš labai trumpas pastraipas. V4 nestabtelėjo prieš pastraipą, pradedančią jungtuku „tačiau“, rodančiu prieštaravimą ankstesnėje pastraipoje pradėtai minčiai. Tačiau tai tik 0,4 proc. visų pastraipų, todėl daryti kokius nors apibendrinimus gana sunku, nes šie atvejai galėjo būti atsitiktiniai (kokius nors paralingvistinių reiškinų šiose įrašo vietose neužfiksuota).

2 lentelė. Tarpfrazinių pauzių kiekis garso įrašuose

Pauzių kiekis	M1	M2	M3	M4	V1	V2	V3	V4
Tarp pastraipų	99	99	100	100	100	100	100	99
Tarp sakinių	264	275	260	272	273	273	275	275

Po kiekvieno sakinio visada stabtelėjo trys diktoriai, o ir kitų diktorių įrašuose tarpfrazinių pauzių šioje pozicijoje nebuvimas retas (sudaro tik 1,5 proc.). Pauzes tarp sakinių (kaip ir tarp pastraipų) skaitomoje kalboje „įpareigoja“ daryti sakinio pabaigos skyrybos ženklai. Skaitomas tekstas šiuo atveju daro labai didelį spaudimą: sakinio skyrybos ženklas paprastai yra jei ne pauzės, tai pagrindinio tono kaitos signalas. Taigi tokios pauzės paprastai sprendžia ne tik lingvistinius (skiria semantiškai ir gramatiškai savarankiškus vienetus), bet ir fiziologinius klausimus (suteikia progos įkvėpti oro). Jos yra gana lengvai prognozuojamos.

Greičiausiai skaičiusios diktorės tarp sakinių darė mažiau pauzių, vadinasi, galima įžvelgti nors ir nedidelę skaitymo greičio įtaką pauzėms atsirasti. Kai kuriais atvejais, kai sakiniai netrumpi, o diktoriai skaito gana lėtai (V1 ir V2), pauzė po vieno ar kito sakinio nedaroma. Taip yra, matyt, dėl to, kad tikimasi, jog užteks oro tolesniam tekstui skaityti. Klausant įrašų, iš kitų diktorių išsiskiria V1 ir V2 skaitymo balsu įgūdžiai: balse dažnai juntama abejonė, kai kurie žodžiai skaitomi itin lėtai, pakartojamos jų dalys, skaitoma labai monotoniškai, o tai galėjo nulemti kai kurių pauzių tarp sakinių nebuvimą. Tačiau šią prielaidą pagrįsti ar paneigti bus galima atlikus frazės viduje esančių pauzių išsamų tyrimą.

3. Pauzių tarp pastraipų trukmė

Remiantis užsienio mokslininkų tyrimais ir darbo su įgarsintais tekstais patirtimi (ne tik šio tyrimo metu), vidutinio ilgio pastraipomis laikytos tos pastraipos, kurios, rašant Times New Roman 12 pt šriftu, užima 7–11 eilučių (Alley 1996) arba yra sudarytos iš maždaug 150 žodžių (Connely, Forsyth 2011). Peržvelgus lietuviškus tekstus, paaiškėjo, kad į septynias eilutes telpa maždaug 90–100 žodžių, į 11 – apie 150–160 žodžių, todėl 90–160 žodžių pastraipos toliau laikomos vidutinėmis, mažiau žodžių turinčios – trumpomis,

3 lentelė. Pauzių tarp pastraipų trukmės duomenys (ms)*

Įrašai, pastraipos, statistiniai įverčiai		M1	M2	M3	M4	V1	V2	V3	V4	Bendras vidurkis
Grožinių tekstų įrašai										
Po trumpų (16)	Vidurkis	709	813	687	493	922	1006	1019	897	818
	Mediana	830	740	650	440	775	1035	895	865	740
Po vidutinių ir ilgų (6)	Vidurkis	1045	857	648	428	1178	965	1013	973	912
	Mediana	930	670	655	445	890	940	980	975	795
Mokslinių tekstų įrašai										
Po trumpų (13)	Vidurkis	1093	981	1004	679	982	1032	1386	1557	1148
	Mediana	935	845	945	525	805	965	1255	1565	985
Po vidutinių ir ilgų (9)	Vidurkis	773	994	1340	813	1284	1398	1609	1788	1217
	Mediana	730	880	1255	745	1285	1455	1685	1715	1125
Publicistinių tekstų įrašai										
Po trumpų (56)	Vidurkis	719	819	642	546	984	752	1328	1091	860
	Mediana	530	810	640	505	770	635	1345	980	760
Apibendrinti duomenys										
Po trumpų (85)	Vidurkis	779	845	411	402	987	840	1281	1120	833
	Mediana	725	805	420	400	780	780	1260	975	768
Po vidutinių ir ilgų (15)	Vidurkis	889	935	1026	626	1239	1212	1354	1439	1027
	Mediana	855	810	1065	535	1115	1075	1275	1275	859
Bendras vidurkis		795	858	754	567	1010	896	1290	1775	993
Bendra mediana		770	800	735	510	800	845	1255	1020	842

* Šioje ir 4 lentelėje greta nurodomo pastraipos ar sakinio ilgio parašytas tokių vienetų kiekis tiriamojame medžiagoje.

daugiau – ilgomis. Trumpų pastraipų tiriamuosiuose tekstuose rasta daugiausia – 85 proc. (publicistiniai tekstai buvo sudaryti tik iš trumpų pastraipų), vidutinių – 11 proc., ilgų – 4 proc. Dėl nedidelio vidutinių ir ilgų pastraipų kiekio jos nagrinėjamos ir aptariamos kartu.

Iš 3 lentelėje pateiktų apibendrintų duomenų matyti, kad pauzės tarp pastraipų gali būti vidutiniškai nuo 400 ms iki 1800 ms. Po vidutinių ir ilgų pastraipų pauzės, palyginti su trumpomis, pailgėja 1,2 karto, jų trukmės vidurkiai ir medianos didesni nei pausių po trumpų pastraipų, skirtumas statistiškai reikšmingas (apskaičiuota Studento kriterijaus reikšmė didesnė už kritinę $4 > 1,9$, pasikliaujamieji intervalai nepersidengia $854 \div 925$ ms ir $992 \div 1188$ ms). Tačiau konkrečių diktorių duomenys, ypač grožinių tekstų, yra įvairesni, plg. M3, M4, V2, V3 vidurkius ir medianas, kurie yra panašūs ir po trumpų, ir po ilgesnių pastraipų, o kai kurių net ir didesni po trumpų pastraipų. Taigi, perskaitę ilgą pastraipą, diktoriai gali stabelėti tiek pat laiko, kaip ir po trumpos pastraipos.

Apibendrinti duomenys rodo, kad esama sąsajos tarp skaitymo greičio ir pausių tarp pastraipų trukmės: lėtakalbių diktorių pauzės po trumpų pastraipų yra 1,7 karto, po ilgesnių – 1,5 karto ilgesnės nei greitakalbių, pausių trukmės vidurkiai ir medianos didesni nei greitakalbių, skirtumai statistiškai reikšmingi (lėtakalbių diktorių apskaičiuota Studento kriterijaus reikšmė didesnė už kritinę reikšmę atitinkamai po trumpų ir ilgesnių pastraipų $9,7 > 1,9$ ir $4,8 > 1,9$, pasikliaujamieji intervalai $687 \div 757$ ms ir $1000 \div 1114$ ms po trumpų pastraipų, $761 \div 978$ ms ir $1170 \div 1452$ ms po ilgesnių pastraipų). Peržvelgus konkrečius kiekvieno diktoriaus duomenis, matyti, kad yra bendrąją tendenciją neatitinkančių atvejų (plg. M1, M3, V1 pauzes moksliniuose tekstuose). Vadinas, lėčiau skaitančiųjų pauzės tarp pastraipų gali būti panašios trukmės, kaip ir greitai skaitančiųjų.

4. Pauzių tarp sakinių trukmė

Visi sakiniai suskirstyti į labai trumpus (sudarytus iš 1–8 žodžių), trumpus (9–17 žodžių), vidutinius (18–30 žodžių), ilgus (31–46 žodžių) ir labai ilgus (46 ir daugiau žodžių) (dėl tokio skirstymo žr. Bitinienė 2001; Leonavičienė 2010). Labai trumpų sakinių tekstuose pasitaikė 83, trumpų – 118, vidutinio ilgio – 62, ilgų – 9, labai ilgų – tik 3. Kadangi labai ilgų sakinių tekstuose pasitaikė mažai, jie priskirti prie ilgų sakinių.

Iš 4 lentelėje pateiktų pausių po skirtingo ilgio sakinių trukmės apibendrintų vidurkių matyti, kad trumpiausios pauzės yra po labai trumpų sakinių, po trumpų sakinių jos pailgėja 1,3 karto, po vidutinių ir ilgų – 1,4 karto. Taigi ilgėjant sakiniams daugeliu atvejų ilgėja ir pauzės po jų. Tačiau trukmės skirtumas tarp trumpų, vidutinių ir ilgų sakinių yra labai nedidelis, nors jis statistiškai nereikšmingas tik tarp vidutinių ir ilgų sakinių (žr. 5 lent.). Išanalizavus pausių medianas, gaunamos panašios tendencijos. Po trumpų sakinių pauzės 1,1 karto, po vidutinių 1,2 karto, po ilgų 1,4 karto ilgesnės nei po labai trumpų. Apskritai pauzės tarp sakinių gali būti nuo 200 ms iki 1300 ms.

4 lentelė. Pauzių tarp sakinių trukmės duomenys (ms)

Įrašai, sakiniai, statistiniai įverčiai		M1	M2	M3	M4	V1	V2	V3	V4	Bendras vidurkis
Grožinių tekstų įrašai										
Po labai trumpų (71)	Vidurkis	375	487	309	293	507	754	574	493	474
	Mediana	370	470	230	340	480	660	500	480	440
Po trumpų (73)	Vidurkis	510	559	432	407	687	867	811	645	615
	Mediana	410	550	460	400	590	780	700	610	540
Po vidutinių (13)	Vidurkis	432	526	517	450	640	897	1160	766	672
	Mediana	420	550	450	450	560	860	1290	740	550
Po ilgų (3)	Vidurkis	403	580	977	407	793	850	460	1190	597
	Mediana	370	590	740	390	840	870	440	1360	740
Mokslinių tekstų įrašai										
Po labai trumpų (1)	Vidurkis	690	880	940	440	40	540	1280	680	686
	Mediana	690	880	940	440	40	540	1280	680	685
Po trumpų (15)	Vidurkis	367	632	463	453	479	644	880	620	565
	Mediana	390	620	450	420	440	520	720	600	520
Po vidutinių (28)	Vidurkis	453	693	669	450	570	761	803	950	675
	Mediana	400	695	690	450	580	620	700	685	610
Po ilgų (8)	Vidurkis	378	631	648	463	675	799	971	819	685
	Mediana	315	660	685	415	650	850	795	770	685
Publicistinių tekstų įrašai										
Po labai trumpų (11)	Vidurkis	305	526	208	423	335	438	493	451	397
	Mediana	310	500	210	470	370	440	400	500	400
Po trumpų (30)	Vidurkis	334	570	334	365	426	615	855	567	508
	Mediana	320	560	340	395	400	660	720	545	475
Po vidutinių (21)	Vidurkis	412	619	470	441	553	728	800	573	574
	Mediana	340	630	450	450	480	630	720	530	530
Po ilgų (1)	Vidurkis	280	600	740	370	360	590	360	640	498
	Mediana	280	600	740	370	360	590	360	640	485
Apibendrinti duomenys										
Po labai trumpų (83)	Vidurkis	369	497	303	312	479	710	572	490	467
	Mediana	360	480	230	350	460	630	500	480	436
Po trumpų (118)	Vidurkis	447	571	411	402	594	774	831	622	582
	Mediana	390	560	420	400	520	690	710	560	531
Po vidutinių (62)	Vidurkis	435	633	569	471	579	778	877	784	641
	Mediana	400	630	530	470	560	640	710	620	570
Po ilgų (12)	Vidurkis	376	616	738	441	678	794	793	897	667
	Mediana	315	640	720	390	650	850	740	785	636
Bendras vidurkis		415	565	430	392	560	751	757	631	563
Bendra mediana		380	550	420	400	520	680	670	560	523

5 lentelė. Pauzių tarp sakinių trukmės statistinis vertinimas

Sakiniai	Apskaičiuota ir kritinė Studento kriterijaus reikšmės	Standartinis nuokrypis (ms)	Pasikliaujamieji intervalai (ms)
Labai trumpi sakiniai Trumpi sakiniai	7>2	308 338	443÷490 560÷603
Labai trumpi sakiniai Vidutiniai sakiniai	9>2	308 342	443÷490 611÷671
Labai trumpi sakiniai Ilgi sakiniai	5,9>1,9	308 312	443÷490 604÷729
Trumpi sakiniai Vidutiniai sakiniai	3>2	338 342	560÷603 611÷671
Trumpi sakiniai Ilgi sakiniai	2,3>1,9	338 312	560÷603 604÷729
Vidutiniai sakiniai Ilgi sakiniai	0,6<1,9	342 312	611÷671 604÷729

Jei būtų analizuojami konkrečių diktorių duomenys, paaiškėtų, kad bendroji tendencija ne visada išlieka. Pavyzdžiui, M1, V1, V2 pauzių po vidutinio ilgio sakinių trukmės vidurkiai yra šiek tiek mažesni nei pauzių po trumpų sakinių. M1, M2, M4, V3 pauzių po ilgų sakinių vidurkiai mažesni nei po vidutinių sakinių.

Pauzių trumpinimą po ilgesnių sakinių galėjo lemti pačių sakinių struktūra. Labai trumpi sakiniai paprastai yra vientisiniai, jų viduje stabtelima labai retai. Pauzės nėra reikalingos nei dėl fiziologinių, nei dėl gramatinių priežasčių. Žinoma, pauzių gali atsirasti dėl įvairių prozodinių ir paralingvistinių reiškinų. Trumpi sakiniai taip pat gana lengvai perskaitomi vienu atsikvėpimu arba reikalauja tik trumpo atsikvėpimo viduryje ir gale. Vidutinio ilgio ir ilgi sakiniai yra net vizualiai ilgesni, užima daugiau nei vieną eilutę, jie būna sudėtiniai arba turi nemažai išplėstinių sakinio dalių ar kitų elementų, išskiriamų skrybės ženklais, neretai implikuojančių ir gana aiškiais pauzėmis. Skaitant tokius sakinius, svarbus semantinis-gramatinis teksto skaidymas, todėl pauzių beveik neįmanoma išvengti, greičiausiai jų sąskaita diktoriai gali trumpinti pauzes po ilgesnių sakinių. Tačiau išsami frazės viduje daromų pauzė analizė yra jau kito tyrimo etapo objektas.

Išvados

Išanalizavus aštuonių diktorių skaitytus šešis tekstus (iš viso apie 6 tūkst. žodžių), galima daryti tokias išvadas:

1. Pauzes tarp pastraipų paprastai daro visi skaitantieji balsu. Jų trukmė priklauso nuo skaitymo greičio: greičiau skaitančiųjų pauzės yra vidutiniškai nuo 400 ms iki 1340 ms, lėčiau skaitančiųjų – nuo 750 ms iki 1800 ms. Bendrosios pauzių trukmės ir pastraipų ilgio sąsajos tendencijos yra panašios ir nepriklauso nuo skaitymo greičio: po trumpų pastraipų pauzės yra trumpiausios, o po ilgesnių pastraipų jos yra 1,2 karto ilgesnės.
2. Pauzės tarp skaitomų sakinių taip pat beveik visada daromos. Greičiau skaitančiųjų pauzės vidutiniškai yra nuo 200 ms iki 980 ms, lėčiau skaitančiųjų – nuo 340 ms

iki 1360 ms. Ilgėjant sakiniams, ilgėja ir pauzės po jų (iki 1,4 karto). Tačiau po ilgų sakinių jos gali būti tokios pačios trukmės, kaip ir po vidutinių sakinių.

3. Pauzės tarp pastraipų vidutiniškai yra 1,8 karto ilgesnės už pauzes tarp skaitomų sakinių. Vadinasi, siekiant sintezuotos kalbos natūralumo tarp pastraipų derėtų daryti beveik dvigubai ilgesnes pauzes nei tarp sakinių. Kalbos atpažinimo atveju beveik dvigubai ilgesnės pauzės gali būti signalas teksta užrašyti iš naujos eilutės.

Natūralioje kalboje yra daug veiksnių, lemiančių pauzių atsiradimą ir jų trukmę, pavyzdžiui, semantiniai ir gramatiniai žodžių ryšiai, prozodiniai reiškiniai ir pan. Tačiau kol kas jie nelabai paklūsta automatinei lietuvių kalbos analizei, jų negalima atpažinti garsinės kalbos signale, kompiuterinis prognozavimas yra labai sudėtingas, o kalbos technologijų reikmėms gal iš dalies ir per ankstyvas – reikėtų atlikti kur kas daugiau tyrimų. Vis dėlto kai kurie gautų rezultatų aspektai gali būti pravartūs, todėl ateityje reikėtų išsamiai tirti skaitomos kalbos frazės viduje daromų pauzių trukmę ir, žinoma, nagrinėti spontaniškos kalbos pauzes.³

Literatūra

Alley, M., 1996. *The Craft of Scientific Writing*. New York: Springer.

Balčiūnienė, I., Kamandulytė-Merfeldienė, L., 2016. Dabartinės sakininės lietuvių kalbos pasakymų ilgis ir struktūra. *Lietuvių kalba*, 10, p. 1–13. Prieiga: https://www.vdu.lt/cris/bitstream/20.500.12259/1378/1/ISSN1822-525X_2016_N_10.pdf [Žiūr. 2020 01 29]. <https://doi.org/10.15388/Lietkalb.2016.10.9914>

Beller, G., Hueber, T., Schwartz, D., Rodet, X., 2006. Speech rates in French expressive speech. In *Speech Prosody*, Dresden, May 2006, SproSig, ISCA. Prieiga: <http://articles.ircam.fr/textes/Beller06a/index.pdf> [Žiūr. 2020 01 29]

Benki, J., Broome, J., Conrad, F., Groves, R., Kreuter, F., 2011. Effects of speech rate, pitch and pausing on survey participation decisions. In: American Association for Public Opinion Research. Prieiga: <http://ww2.amstat.org/sections/SRMS/Proceedings/y2011/Files/400189.pdf> [Žiūr. 2018 04 28]

Bitinienė, A., 2001. Grožinio stiliaus prozos tekstų tiesioginės kalbos sakinio ilgis ir struktūra. *Kalbotyra*, 50 (1), p. 17–28. Prieiga: <http://etalpykla.lituanistikadb.lt/fedora/get/LT-LDB-0001:J.04~2001~1367161842738/DS.002.0.01.ARTIC> [Žiūr. 2018 04 28].

Boersma, P., Weenink, D., 2018. *Praat: doing phonetics by computer*, version 6.035. Prieiga: www.praat.org [Žiūr. 2020 01 29].

Bosker, H. R., Pinget, A. F., Quené, H., Sanders, T., De Jong, N. H., 2013. What makes speech sound fluent? The contributions of pauses, speed and repairs. *Language Testing*, 30 (2), 159–175. Prieiga: https://pure.mpg.de/rest/items/item_1900124_2/component/file_1900123/content [Žiūr. 2020 01 29]. <https://doi.org/10.1177/0265532212455394>

Cahn, J. E., 1990. The generation of affect in synthesized speech. *Journal of the American Voice I/O Society*, 8 (1), pp. 1–19. Prieiga: <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.435.9599&rep=rep1&type=pdf> [Žiūr. 2020 01 29]

Cannizzaro, M., Harel, B., Reilly, N., Chappell, Ph., Snyder, P. J., 2004. Voice acoustical measurement of the severity of major depression. *Brain and Cognition*, 56, pp. 30–35. Prieiga: <http://www.sci.brooklyn.cuny.edu/~levitan/nlp-psych/papers/cannizarro04.pdf> [Žiūr. 2019 01 04]. <http://dx.doi.org/10.1016/j.bandc.2004.05.003>

³ Autorės nuoširdžiai dėkoja anoniminiams recenzentams, kurių pastabos, o ypač klausimai atkreipė dėmesį į tyrimo aprašą, kai kurių teiginių (mums lyg ir savaime suprantamų, bet nelabai aiškių kiek atokesniuose nuo mūsų filologijos tyrimo laukuose dirbantiems mokslininkams) aktualizavimą.

Connelly, J., Forsyth, P., 2011. *Essay Writing Skills. Essential Techniques to Gain Top Grades*. London: Kogan Page.

Gosztolya, G., Vincze, V., Toth, L., Pakaski, M., Kalman, J., Hoffmann, I., 2018. Identifying Mild Cognitive Impairment and mild Alzheimer's disease based on spontaneous speech using ASR and linguistic features. *Computer Speech and Language*, 53, pp. 1–17. Prieiga: <http://www.inf.u-seged.hu/~ggabor/publist/2019-csl.pdf> [Žiūr. 2020 01 29]. <http://dx.doi.org/10.1016/j.csl.2018.07.007>

Kazlauskienė, A., Kalašinskaitė, E., 2017. Pauzių trukmė skaitomoje lietuvių kalboje. *Kalbų studijos*, 31, p. 71–82. Prieiga: https://etalpykla.vdu.lt/bitstream/handle/1/35611/ISSN2029-7203_2017_N_31.PG_71-82.pdf?sequence=1 [Žiūr. 2018 12 23]. <http://dx.doi.org/10.5755/j01.sal.0.31.18653>

Kazlauskienė, A., Raškinis, G., 2013. Intonuoto garsyno kūrimo principai. *Žmogus ir žodis*, 15 (1), p. 101–110. Prieiga: http://vddb.library.lt/fedora/get/LT-eLABa-0001:J.04-2013-ISSN_1392-8600.V_15.N_1.PG_101-110/DS.002.0.01.ARTIC [Žiūr. 2018 04 28]

Kazlauskienė, A., Veličkaitė, K., 2003. Pastabos dėl lietuvių kalbėjimo tempo. *Acta Linguistica Lithuanica*, 48, 49–57. Prieiga: <https://etalpykla.lituanistikadb.lt/object/LT-LDB-0001:J.04-2003-1367164520655/J.04-2003-1367164520655.pdf> [Žiūr. 2020 01 29]

Kendall, S. T., 2009. *Speech Rate, Pause, and Linguistic Variation: An Examination Through the Sociolinguistic Archive and Analysis Project*. Doctoral dissertation, Duke University, Durham, NC. Prieiga: <https://pdfs.semanticscholar.org/1f2c/d009a0a5cc77f4d676f1a6659a5936b147a3.pdf> [Žiūr. 2020 01 29]

Koženiauskienė, R., 1999. *Retorika: iškalbos stilistika*. Vilnius: Mokslo ir enciklopedijų leidybos institutas.

Kundrotas, G., 2010. Lietuvių kalbos intonacinių kontūrų funkcinės galimybės. *Žmogus ir žodis*, 12 (1), p. 13–19. Prieiga: <http://www.biblioteka.vpu.lt/zmogusirzodis/PDF/didaktinelingvistika/2010/kundr42-47.pdf> [Žiūr. 2019 01 03]

Leonavičienė, A., 2010. Sakinių ilgis – publicistinio ir šnekamojo stiliaus sandūros tekstuose požymis. *Kalbotyra*, 62 (1), p. 95–107. Prieiga: <http://www.zurnalai.vu.lt/kalbotyra/article/viewFile/7646/5523> [Žiūr. 2018 04 28]. <https://doi.org/10.15388/Klbt.2010.7646>

Megehee, C. M., Dobie, K., Grant, J., 2003. Time versus pause manipulation in communications directed to the young adult population: does it matter? *Journal of Advertising Research*, 43 (3), pp. 281–292. Prieiga: <http://data.adic.co.kr/lit/publication/tmp/A9001220/A9001220.pdf> [Žiūr. 2020 01 29]. <https://doi.org/10.1017/S0021849903030277>

Miller, J. L., Grosjean, F., 1981. How the components of speaking rate influence perception of phonetic segments. *Journal of Experimental Psychology: Human Perception and Performance*, 7 (1), pp. 208–215. Prieiga: <https://pdfs.semanticscholar.org/6c5f/520b1a6e4bfda506737470f7dc38afee2da9.pdf> [Žiūr. 2020 01 29]. <https://doi.org/10.1037/0096-1523.7.1.208>

Trouvain, J., Koreman, J., Erriquez, A., Braun, B., 2001. Articulation Rate Measures and Their Relation to Phone Classification in Spontaneous and Read German Speech. In: *Proc. of ITRW on Adaptation Methods for Speech Recognition*, Sophia Antipolis, France, pp. 155–158. Prieiga: http://ling.uni-konstanz.de/pages/home/braun/articles/trouvain_et_al_2001.pdf [Žiūr. 2018 04 28].

Viola, I. C., Madureira, S., 2008. The roles of pause in speech expression. *Fourth International Conference on Speech Prosody, Campinas, Brazil*, pp. 721–724. Prieiga: <https://pdfs.semanticscholar.org/1365/c596ccc881af3b76cfe76ad25bd0d54f7d8f.pdf> [Žiūr. 2019 01 04].

Zellner, B., 1994. Pauses and the temporal structure of speech. In: Ed. E. Keller. *Fundamentals of speech synthesis and speech recognition*, Chichester: John Wiley, pp. 41–62. Prieiga: <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.36.1414&rep=rep1&type=pdf> [Žiūr. 2018 04 28].