COMPUTER SIMULATION EXPERIMENTS IN PHONETICS AND PHONOLOGY

კომპიუტერული სიმულაციური ექსპერიმენტები ფონეტიკასა და ფონოლოგიაში

Tinatin Mshvidobadze
Doctor of Technical Sciences
Associate Professor of Gori State University
Gori, Chavchavadze st. =53, 1400, Georgia
+995555118379 tinikomshvidobadze@gmail.com
https://orcid.org/0000-0003-3721-9252

Abstract. Speech is a continuous, highly variable acoustic signal that is perceived as a sequence of sounds and words. Phonetics and phonology are concerned with the production of these units in the human articulatory apparatus, their properties as acoustic phenomena and perception, as well as with the speech units used in a given language and their possible combinations. Experimental phonetics and laboratory phonology use corpora with large numbers of speech recordings. Alternatively, certain types of phenomena are studied specifically with humans in a laboratory setting. The paper deals with computer simulation experiments as a "third kind" approach to phonetics and phonology. Their content emphasizes the broader context of simulation technology in the process of linguistic studies of human speech. Computer simulation experiments offer the highest degree of control over data. In this article, we discuss the effectiveness of these experiments in phonetics and phonology.

Keywords: phonetics, phonology, computer technology, simulation experiments, speech technology.

თინათინ მშვიდობაძე
ტექნიკურ მეცნიერებათა დოქტორი
გორის სახელმწიფო უნივერსიტეტის ასოცირებული პროფესორი
ქ. გორი, ჭავჭავაძის ქ. = 53, 1400, საქართველო
+995555118379 tinikomshvidobadze@gmail.com
https://orcid.org/0000-0003-3721-9252

აბსტრაქტი. მეტყველება არის უწყვეტი, უაღრესად ცვალებადი აკუსტიკური სიგნალი, რომელიც აღიქმება, როგორც ბგერებისა და სიტყვების თანმიმდევრობა. ფონეტიკა და ფონოლოგია ეხება ამ ერთეულების წარმოქმნას ადამიანის არტიკულაციურ აპარატში, მათი თვისებებით, როგორც აკუსტიკური ფენომენებით და აღქმით, ასევე მოცემულ ენაში გამოყენებული სამეტყველო ერთეულებით და მათი შესაძლო კომბინაციებით. ექსპერიმენტულ ფონეტიკასა და ლაბორატორიულ ფონოლოგიაში გამოიყენება კორპუსები დიდი რაოდენობის სამეტყველო ჩანაწერებით. ალტერნატიულად, გარკვეული სახის ფენომენები შესწავლილია სპეციალურად ადამიანებთან ერთად ლაბორატორიულ გარემოში. ნაშრომი ეხება კომპიუტერულ სიმულაციურ ექსპერიმენტებს, როგორც „მესამე სახის“ მიდგომას ფონეტიკასა და ფონოლოგიაში. მათი შინაარსი ხაზს უსვამს სიმულაციური ტექნოლოგიის უფრო ფართო კონტექსტს ადამიანის მეტყველების ლინგვისტური კვლევების პროცესში. კომპიუტერული სიმულაციური ექსპერიმენტები გვთავაზობენ მონაცემებზე კონტროლის უმაღლეს ხარისხს. სტატიაში განვიხილავთ ამ ექსპერიმენტების ეფექტურობას ფონეტიკასა და ფონოლოგიაში.

საკვანძო სიტყვები: ფონეტიკა, ფონოლოგია, კომპიუტერული ტექნოლოგია, სიმულაციური ექსპერიმენტები, მეტყველების ტექნოლოგია.

შესავალი. წარმოდგენილი კვლევა ითვალისწინებს სიმულაციური ექსპერიმენტების დეტალურ ანალიზს ფონეტიკასა და ფონოლოგიაში, რომელიც ავსებს როგორც ემპირიულ, ისე თეორიულ კვლევებს. ნაშრომი, ერთის მხრივ, მოიცავს ფონეტიკისა და ფონოლოგიის კვლევის სხვადასხვა ქვედისციპლინებს, ასევე ადამიანის მეტყველების აღქმისა და წარმოების კვლევის სფეროებს მიმდებარე დისციპლინებიდან, როგორიცაა ფსიქოლოგია, შემეცნებითი მეცნიერება, ფსიქოლინგვისტიკა და სხვა ლინგვისტური დისციპლინები. მეორეს მხრივ, კომპიუტერულ მეცნიერებებს, როგორიცაა მანქანათმცოდნეობა, ხელოვნური ინტელექტი ან მეტყველების ტექნოლოგია. ნაშრომში წარმოდგენილი კვლევის მთავრი მიზანია აჩვენოს კომპიუტერული სიმულაციური ექსპერიმენტების გამოყენების შესაძლებლობები ფონეტიკასა და ფონოლოგიაში. განხილული კომპიუტერული სიმულაციური კვლევები შეიძლება ჩაითვალოს გამოთვლითი ლინგვისტიკის აპლიკაციებად.

ხელოვნური ინტელექტისა და მანქანური სწავლების შესახებ და მეტყველების ტექნოლოგიებისა და ბუნებრივი ენის დამუშავების (NLP) კვლევის ხანგრძლივი ისტორიის მანძილზე, არსებობს ცოდნის უზარმაზარი საცავი, მეთოდოლოგიური პარადიგმები და ინსტრუმენტები, საიდანაც ფონეტიკოსებს და ფონოლოგებს შეუძლიათ აირჩიონ დიზაინის ჩარჩო.

მკვლევარი ფანტი (Fant, 2004: 24) მიმოიხილავს ფონეტიკისა და ფონოლოგიის განვითარებას მე-20 საუკუნის მეორე ნახევრიდან. ის აღნიშნავს, რომ ” 1965 წლიდან 1980 წლამდე პერიოდში კომპიუტერები იქცა სტანდარტულ ლაბორატორიულ ინსტრუმენტად, რაც ნიშნავს ანალოგურიდან ციფრულ დამუშავებაზე გადასვლას მეტყველების კვლევის ყველა ასპექტში. ექსპერიმენტულ ფონეტიკაში ჩვეულებრივი პრაქტიკაა ზოგიერთი საცდელი სუბიექტის მეტყველების ნიმუშების ჩაწერა ლაბორატორიულ პირობებში, რაც უზრუნველყოფს მონაცემებზე კონტროლის მაღალ ხარისხს. კიდევ ერთი გავრცელებული პრაქტიკაა მეტყველების კორპუსების ფონეტიკური ანალიზი, ანუ ჩაწერილი და ანოტირებული მეტყველების ჩანაწერების მონაცემთა ნაკრები. ასეთ სამეტყველო კორპუსებს შეუძლიათ უზრუნველყონ ბუნებრიობის უფრო მაღალი ხარისხი, ხოლო მკვლევარის კონტროლი მონაცემებზე ბევრად უფრო შეზღუდულია.

რამდენადაც არ არსებობს ყოვლისმომცველი განხილვა სიმულაციური ტექნოლოგიის შესახებ ფონეტიკასა და ფონოლოგიაში, აუცილებელია გავცდეთ ამ სფეროების საზღვრებს, რათა უკეთ გავიგოთ კომპიუტერული სიმულაციური ექსპერიმენტების მნიშვნელობა მეტყველებისა და ენის ლინგვისტური შესწავლისთვის.

მკვლევარი სანი (Sun, 2008: 6) აღნიშნავს, რომ: „გამოთვლითი მოდელები უზრუნველყოფენ ალგორითმულ სპეციფიკას [...]. ამიტომ ისინი უზრუნველყოფენ როგორც კონცეპტუალურ სიცხადეს, ასევე სიზუსტეს“.

ბეიკერი (Baker, 2008: 289–307) აღნიშნავს: „გამოთვლითი კვლევები ხშირად ავლენს თეორიების ფარულ შედეგებს [...]“. ის განიხილავს აშკარა პარადოქსს, რომლის მიხედვითაც მოდელების მოლოდინი გაფართოვდება მაშინ, როცა ისინი მხოლოდ იმ ცოდნის კოდირებას მოახდენენ, რაც უკვე აქვს მოდელს. ამრიგად, სიმულაციური ექსპერიმენტის მიზანია ცოდნის გაფართოება სისტემის და მისი მოდელის ცვლადების ურთიერთობისა და ურთიერთქმედების საფუძველზე.

ბორსმა და ჰამანმა გამოიყენეს სიმულაციური ექსპერიმენტები ოპტიმალურობის თეორიის ფარგლებში, რათა ეჩვენებინათ, რომ ხმის სისტემებში „დისპერსია“ ავტომატურად წარმოიქმნება რამდენიმე თაობის შემდეგ, მისი კონტროლისთვის. ექსპერიმენტების ჩატარების ძირითადი იდეაა ინტერესის საგნის მაქსიმალურად იზოლირება, რათა შესაძლებელი იყოს მისი დეტალების გამოკვლევა (Boersma & Hamann, 2008: 217–270). ექსპერიმენტული ფონეტიკის ან ლაბორატორიული ფონოლოგიის პრობლემა ისაა, რომ შეუძლებელია ენობრივი სისტემის კონკრეტული ნაწილების

გამოყოფა მომხსენებლებისაგან. კომპიუტერული სიმულაციების დახმარებით მკვლევარს შესაძლებლობა აქვს შეიმუშაოს ექსპერიმენტები, რომლებიც ეხება ფონეტიკისა და ფონოლოგიის აბსტრაქტულ თეორიულ ასპექტებს, დამოუკიდებლად რომელიმე კონკრეტული ენისა და მოსაუბრისგან.

კომპიუტერული სიმულაციები ხშირად იყენებენ ფსევდო-შემთხვევით რიცხვებს, რომლებიც ეფექტურად იქმნება ალგორითმულად და ასევე ხელახლა გენერირდება იმავე თანმიმდევრობით, საჭიროების შემთხვევაში. ეს უკანასკნელი ფაქტი მნიშვნელოვანია ექსპერიმენტის ზუსტი განმეორებისთვის და მისი შედეგების რეპროდუცირებისათვის. და ბოლოს, ტერმინი „კომპიუტერული სიმულაცია“ ეხება სიმულაციას, რომელიც შეიძლება შესრულდეს როგორც პროგრამა კომპიუტერზე, ძირითადი მოდელის პარამეტრიზაციის გათვალისწინებით.

პირველი კომპიუტერული სიმულაციური ექსპერიმენტები ჩატარდა 1947 წელს სტანისლავ უილიამის და ჯონ ნეუმანის მიერ *ENIAC*-ზე, ერთ-ერთ პირველ ციფრულ კომპიუტერზე. საინტერესოა, რომ ეს პირველი სიმულაციური ექსპერიმენტები შეიძლება ჩაითვალოს აპარატურის სიმულაციად, რამდენადაც ისინი ძირითადად იმეორებდნენ „MONTE CARLO TROLLEY“-ს, ან „FERMIAC“-ის ფუნქციონირებას. ეს იყო ანალოგური კომპიუტერი, რომელიც შეიქმნა მონტე კარლოს გამოთვლების შესასრულებლად ბირთვულ ფიზიკაში (Metropolis, 1987: 125–130).

ტურინგმა შემოგვთავაზა ცნობილი იმიტაციის თამაში, რომელსაც ხშირად მოიხსენიებენ როგორც ტურინგის ტესტს, რათა შეემოწმებინა, შეძლებდა თუ არა მანქანა სიტყვიერად გადმოცემის უნარს. ეს კომპიუტერი-ადამიანის ტესტი სრულდება ორი მოთამაშიდან ერთის კომპიუტერით ჩანაცვლებით. (Turing, 1950: 456).

ფონეტიკასა და ფონოლოგიაში კომპიუტერული სიმულაციების თანამედროვე ეპოქის დასაწყისია ლილენკრანტსისა და ლინდბლომის ნაშრომი ხმოვანთა სისტემების შესახებ. მათ გამოიკვლიეს ფაქტორები, რომლებიც გავლენას ახდენს ხმოვან სისტემებზე კომპიუტერული სიმულაციების გამოყენებით, როგორიცაა თვითორგანიზაციის პრინციპები. მათი სიმულაციებით მიღებული სტრუქტურები შედარებულია სხვადასხვა ბუნებრივ ენაზე არსებული ხმოვანთა სისტემების ემპირიულ მონაცემებთან. ეს მკვლევარები განიხილავენ ფონეტიკასა და ფონოლოგიას (და ზოგადად ლინგვისტიკას) შორის ურთიერთობას და მათ განსხვავებას. ფონეტიკა, როგორც ისინი ამტკიცებენ, უპირველეს ყოვლისა ეხება მეტყველების სტრუქტურას და ენობრივი ფორმის ინტერპრეტაციას. მათი სიმულაციური ექსპერიმენტული დასკვნიდან გამომდინარე, ფონოლოგიური სტრუქტურა შეიძლება წარმოიშვას დაბალი დონის ურთიერთქმედებიდან. ამიტომ, ისინი გვთავაზობენ ფონეტიკისა და ლინგვისტიკის ინტეგრირებას. (Liljencrants & Lindblom, 1972: 839-862).

მეთოდები. მოცემულ სტატიაში გამოყენებულია აღწერითი, ანალიზისა და განმარტების მეთოდები, რის საფუძველზეც გამოკვეთილია აღნიშნული კვლევის მნიშვნელოვანი საკითხები. კონცეფციის ჩამოსაყალიბებლად დავიმოწმეთ სხვადასხვა მკვლევარების შეხედულებები, მათ საფუძველზე მოვახდინეთ მსჯელობისა და დასკვნების ილუსტრირება.

ნაშრომში მოცემულია ყველაზე ხშირად გამოყენებული მეთოდებისა და ფორმალური ჩარჩოების მიმოხილვა კომპიუტერული სიმულაციური ექსპერიმენტების შესახებ ფონეტიკისა და ფონოლოგიის სფეროებში, როგორიცაა ძირითადად გამოთვლითი მოდელირებისა და მანქანათმცოდნეობის მეთოდები. აღნიშნული მეთოდები, ძირითადად გამოიყენება სინთეზურ მონაცემებთან, სადაც წარმოქმნიან უამრავ გადაწყვეტილებას.

შედეგები და მსჯელობა. ლილენკრანტსი და ლინდბლომი იკვლევენ ხმოვანთა სისტემების წარმოქმნილ სტრუქტურებს, რომელიც შედგება 3-დან 11 ხმოვნისაგან. ხმოვნები მათ მოდელში წარმოდგენილია წერტილებით ორგანზომილებიან, გამოყოფილ სივრცეში. თითოეული სიმულაცია იწყება ხმოვანთა მითითებული რაოდენობით, რომლებიც განთავსებულია სივრცის ცენტრიდან თანაბარ მანძილზე. პროგრამა განმეორებით ცდილობს მაქსიმალურად გაზარდოს ორმხრივი კონტრასტი ხმოვანთა შორის. ეს მიიღწევა „მთლიანი ენერგიის საზომის“ განსაზღვრით:

$$E = \sum_{i=1}^{n-1} \sum_{j=0}^{n-1} \frac{1}{r_{i,j}^2} \quad (1)$$

სადაც $r_{i,} = (x_i - x_j)^2 + (y_i - y_j)^2$, არის მანძილი ორ ხმოვან წერტილს შორის *i* და *j,* შესაბამისად წარმოდგენილია x და y კოორდინატებით. ხმოვნები მოძრაობენ ხმოვანთა სივრცის შიგნით, ვიდრე არ მოიძებნება E-ს მინიმალური რაოდენობა. აქ მკვლევარები ამტკიცებენ ლინგვისტიკის მიდგომას, რომელიც არის „სუბსტანციაზე დაფუძნებული“ ძირითადი სალაპარაკო ენისა და მისი ყველა ასპექტის გათვალისწინება.

გოლდსმიტი და სხვები ყურადღებას ამახვილებენ ფონოტაქტიკაზე და იყენებენ სწავლის უკონტროლო მეთოდებს, რათა გამოიკვლიონ, თუ როგორ შეიძლება ვისწავლოთ განსხვავება ორი ხმის კატეგორიის ხმოვანსა და თანხმოვანს შორის კორპუსიდან აღებული ფონემური სიმბოლოების განაწილების მიხედვით. ისინი განიხილავენ სუხოტინის ალგორითმს, რომელმაც პირველმა შემოგვთავაზა სიმბოლურ წარმოდგენაში ხმოვანთა და თანხმოვანთა გამოყოფა. ალგორითმი ემყარება კლების და თანხვედრის სიხშირეებს და იმ ვარაუდს, რომ ხმოვნები ყველაზე ხშირია და რომ ისინი ჩვეულებრივ უფრო ხშირად ხვდებიან თანხმოვანებთან. მკვლევარები ყურადღებას ამახვილებენ ყველაზე ფუნდამენტურ ცნებებზე: “ფონემების ხმოვანებად და თანხმოვანებად დაყოფაზე და მათი „აღმოჩენის“ შემდეგ, ხმოვანთა ჰარმონიასა და მარცვლების სტრუქტურაზე”. (Goldsmith et al, 2009: 4–38).

სიმულაციური ექსპერიმენტები სინთეზურ მონაცემთან

აბსტრაქტული სისტემების სიმულაციები მოქმედებს მონაცემთა წარმოდგენაზე აბსტრაქციის მაღალ დონეზე. თუმცა, სიმულაციებს, რომლებიც უფრო უშუალოდ ეხება რეალურ ფიზიკურ მოვლენებს, ასევე შეუძლიათ გამოიყენონ შედარებით აბსტრაქტული მონაცემები. სიმულაციურ ექსპერიმენტებში სინთეზური მონაცემების გამოყენების მოტივი ორგვარია: რეალური მონაცემები შეიძლება იყოს ძალიან რთული/ხმაურიანი ექსპერიმენტის მიზნებისთვის ან შეიძლება საერთოდ არ იყოს ხელმისაწვდომი. მიტრა და სხვებმა, შექმნეს სინთეზურ მონაცემთა ბაზა არტიკულაციური და აკუსტიკური ინფორმაციით ANN-ზე დაფუძნებული მეტყველების ამომცნობის მოსამზადებლად. ჟესტური ინფორმაცია აშენდა რენტგენის მიკროსხივის მეტყველების მონაცემთა ბაზაზე და შესაბამისი აკუსტიკური სიგნალის სინთეზირებით. (Mitra et al, 2012: 2270–2287).

შეფასების ღონისძიებები

თუ კომპიუტერული სიმულაციები გაგებულია, როგორც სისტემის ქცევის იმიტაცია მისი მოდელის მიერ, ჩნდება კითხვა, რამდენად ეფექტურია ეს იმიტაცია რეალურ სისტემასთან შედარებით. კომპიუტერული სიმულაციური ექსპერიმენტების შეფასებები ეხება სიმულაციის დაკვირვების ქცევის ვალიდობას ან ხარისხს, ან მის მიერ წარმოებულ გამომავალ მონაცემებს.

სიზუსტე, სისრულე და F-ზომა

ტურინგის მიხედვით სიზუსტე (P) და სისრულე (R) არის ორი ფართოდ გამოყენებული შეფასების ღონისძიება ბუნებრივი ენის (NLP) დამუშავებისას, ან

ინფორმაციის მოძიებისას. სიზუსტე არის სწორი გადაწყვეტილების თანაფარდობა ყველა ჰიპოთეზულ გადაწყვეტილებასთან, ხოლო სისრულე არის სწორი გადაწყვეტილების თანაფარდობა ყველა შესაძლო გადაწყვეტილების საცნობარო მონაცემებთან. ისინი შეიძლება განისაზღვროს შემდეგნაირად:

$$P = \frac{Tp}{Tp + Fp} \quad (2)$$

$$R = \frac{TP}{Np + Fn} \quad (3)$$

სადაც TP არის ჭეშმარიტი დადებითი, FP მცდარი დადებითი, FN მცდარი უარყოფითი, TN ჭეშმარიტი უარყოფითი.

რამდენადაც უფრო მოსახერხებელია წარმოქმნილი ამოხსნის ხარისხის რაოდენობრივი დადგენა ორის ნაცვლად ერთი რიცხვით, ხშირად მგამოიყენება F-ზომა (Van Rijsbergen, 1979: 134). სიზუსტისა და სისრულის მიხედვით იგი განისაზღვრება როგორც:

$$F_\beta = \frac{(\beta^2+1)*P*R}{\beta^2*P+R} \quad (4)$$

მეტყველების სეგმენტაცია

ტერმინი „მეტყველების სეგმენტაცია“ ეხება სხვადასხვა ერთმანეთთან დაკავშირებულ პრობლემებს, რომლებიც განსხვავდებიან იმ ზომით, რაც რეალურად ითვლება „სეგმენტად“. მეტყველებისა და ბუნებრივი ენის დამუშავებისას ენობრივი ცოდნის გამოთვლითი შეძენის უახლესი მიდგომებია სტატისტიკური და ზედამხედველობითი მეთოდები. ეს მეთოდები დამყარებულია ანოტირებულ მონაცემებზე.

ჩვენ, აღვიქვამთ მეტყველებას, როგორც დისკრეტულ ერთეულთა თანმიმდევრობას, რომლებიც ინდივიდუალურად მეორდება ან იცვლება, მაგ: სიტყვები ან სიმბოლოები. თუმცა, აღქმის სამეტყველო ერთეულები უნდა განვასხვავოთ, როგორც სმენითი დაბალი დონის და უფრო მაღალი დონის ფონოლოგიური ერთეულებისგან (Massaro, 1996: 87).

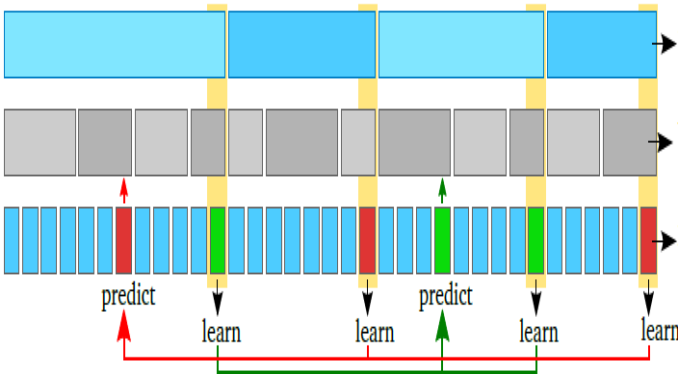
ამრიგად, მეტყველების ნაკადის სეგმენტირება ენობრივად შესაბამის სეგმენტებად, როგორიცაა ასოები, სიტყვები ან ფრაზები, შეიძლება ჩაითვალოს ადამიანის ერთ-ერთ ფუნდამენტურ შესაძლებლობად.

სეგმენტაციის სტრატეგიები

მეტყველების სეგმენტაციის შემოთავაზებული მიდგომები იყენებს სხვადასხვა ტიპის ძირითად სტრატეგიებს. ბრენტი, რომელიც მიმოიხილავს სხვადასხვა გამოთვლითი მეტყველების სეგმენტაციის მოდელებს, განასხვავებს სამ ძირითად კლასს: გამოთქმა-საზღვრის, დიაგნოსტიკურ და სიტყვების ამოცნობის სტრატეგიას. (Brent, 1999: 294–301).

მეტყველების სეგმენტაციის გამოთქმა-საზღვრის სტრატეგიის მოდელები ადარებენ სეგმენტური წარმოდგენის სხვადასხვა დონეებს. ისინი ემყარება დაშვებას, რომ გამოთქმის საზღვრები ასევე არის სიტყვების საზღვრები და რომ თანმიმდევრობები, რომლებიც უშუალოდ წინ უსწრებს გამოთქმის საზღვრებს. ამგვარად, საზღვრები ერთ მოცემულ დონეზე (მაგ., სიტყვის დონეზე) წარმოდგენილია ქვედა დონის ერთეულების (მაგ. ფონემების) თანმიმდევრობის საფუძველზე, რომლებიც დამახასიათებელია ამოსაცნობი სეგმენტების დაბოლოებებისათვის.

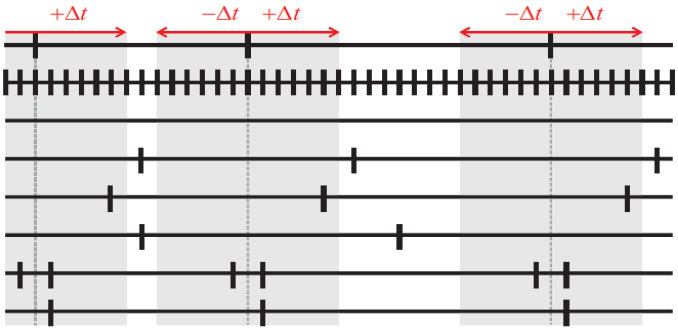
მაღალი დონის სეგმენტებზე დაყრდნობით (მაგ. გამოთქმის დონე). სურათი 1 გვიჩვენებს გამოთქმა-საზღვრის სტრატეგიისა და წარმოდგენის დონეების ილუსტრაციას. ალგორითმის მიერ წარმოებული სეგმენტაცია ფასდება კორპუსის გამოთვლის საწყისი სეგმენტაციის მიხედვით.



სურათი 1.: მეტყველების სეგმენტაციის გამოთქმა-საზღვრის სტრატეგიის პრინციპის ილუსტრაცია სიმბოლური თანმიმდევრობის მიდგომებში.

ქსანთოსი აღნიშნავს, რომ გამოთქმა-საზღვრის სტრატეგიის უპირატესობა არის ის, რომ იგი ეყრდნობა „მონაცემებისგამორჩეულ აღქმად მახასიათებლებს და არა აშკარა სტატისტიკურ თვისებებს, რომლებიც გამოიყენება პროგნოზირებისას დაფუძნებული სტრატეგიებით“. (Xanthos, 2003: 178).

დიაგნოსტიკური სეგმენტაცია ეს არის სეგმენტაციის სტრატეგიის იმპლემენტაცია, რომელიც ასევე მოიცავს მორფოლოგიურ კომპონენტს რეალური სეგმენტატორის მიერ წარმოებული სეგმენტაციის „დასაფიქსირებლად“. (სურ. 2).



სურათი 2. დიაგნოსტიკური სეგმენტატორების სქემატური ილუსტრაცია: ჰორიზონტალური ხაზები წარმოადგენს მონაცემთა თანმიმდევრობის დროის ღერძს სეგმენტის საზღვრებით.

სურათის მიხედვით პარამეტრი Δ t უდრის შესაბამისი ყოველი მომდევნო ფანჯრის პარამეტრს, რომელიც გამოიყენება შეფასებისათვის.

სიტყვების ამოცნობის სტრატეგიის მოდელები ერთდროულად ცდილობენ ისწავლონ მეტყველების სეგმენტების ინვენტარი (მაგ. სიტყვების ლექსიკა) და ქვედა დონის მეტყველების სტრიქონის გაყოფა ამ ინვენტარის შინაარსის ამოცნობის საფუძველზე. იმ შეზღუდვით, რომელიც სეგმენტებს წარმოდგენის ერთ დონეზეა შეუძლებელი, სადაც მთავრდება ერთი ცნობილი სეგმენტი (ანუ სიტყვა) იწყება შემდეგი. ჩვეულებრივ, უცნობი მეტყველების მონაკვეთები ცნობილ სეგმენტებს შორის ემატება როგორც ახალი სიტყვები ლექსიკონში.

პერუჩეტის და სხვების მიერ, წარმოდგენილია კოგნიტურად მოტივირებული სეგმენტაციის მეთოდი სახელწოდებით *PARSER*. ეს მეთოდი განმეორებით აყალიბებს შეწონილ გონებრივ ლექსიკას („აღქმის ფორმირება“) განმტკიცების პროცესების მეშვეობით. ყოველ საფეხურზე თავდაპირველი ანალიზი ხორციელდება შემთხვევითი აღქმის აღებით შეყვანილი მონაცემთა ნაკადიდან. აღქმის ფორმირებისთვის ახალი ერთეულები ემატება ლექსიკონს PS საწყისი წონით. (Perruchet & Vinter, 2019: 246–263).

გოლდვოტერმა და სხვებმა, შეიმუშავეს ბაიესის სწავლებაზე დაფუძნებული ციტირებული სიტყვების ამოცნობის სტრატეგიის მოდელები მეტყველების სეგმენტაციის საფუძველზე. მათი აზრით შესაძლო სიტყვების რაოდენობა შეუზღუდავია ნებისმიერ ბუნებრივ ენაში. მაგალითად, ბაიესის სასწავლო ჩარჩოში მუშაობა გულისხმობს უსასრულო ლექსიკას, ფონემების (ან მარცვლების) რაოდენობა კი შედარებით მცირეა. (Goldwater et al., 2009: 44).

ექსპერიმენტი: მუდმივი ცვლის კორპუსები

განვიხილავთ ექსპერიმენტს, რომელიც იყენებს ხელოვნურად აგებულ საწყის მონაცემებს, რომლებიც მოიხსენიება როგორც მუდმივი ცვლის კორპუსები. ისინი შექმნილია იმისათვის, რომ გაადვილდეს სეგმენტაცია და უზრუნველყოს სასარგებლო ინფორმაცია. მისი მიზანია სეგმენტაციის მეთოდის გააზრება კონტროლირებადი მონაცემების პირობებში. საწყისი თანმიმდევრობა წარმოდგენილია ექსპერიმენტების ფარგლებში. (Duran et al, 2020: 133–168).

მეტყველების სეგმენტაციის პრობლემა ფორმალურად განისაზღვრება, როგორც: მოცემული ინტერვალის X სეგმენტაცია არის l_p საზღვრების სიგრძის თანმიმდევრობა:

$$0 < t_1 < t_2 < < t_n < l_p \text{ , n} \geq 1 \quad (5)$$

აქ მეტყველება არ არის წარმოდგენილი დისკრეტული ანბანით. ლინგვისტური სტრუქტურა გათვალისწინებულია ცალკეული სამეტყველო ერთეულებისთვის (ინტერვალი X) ან მეხსიერებისთვის. ამ სეგმენტში წარმოდგენილი ექსპერიმენტების მიზანია გამოიკვლიოს, თუ როგორ შეიძლება განისაზღვროს სეგმენტური სტრუქტურა კონტროლის გარეშე. ეს არის ფუნდამენტური პრობლემა, რომელიც თავდაპირველად უნდა გადაწყდეს ენის შერჩევისას.

მეთოდი და განხორციელება

სეგმენტაციის ექსპერიმენტი განხორციელებულია და შესრულებულია პროგრამულ უზრუნველყოფაში“ MATLAB“.

ეს შემთხვევით გენერირებული კორპუსები წარმოდგენილია, როგორც მეხსიერების თანმიმდევრობა კონკრეტული ექსპერიმენტის გაშვებისათვის. ექსპერიმენტისთვის სრულმეტრაჟიანი ინტერვალი X იქმნება სამიზნე სეგმენტზე დაყრდნობით, რომელიც ავსებს მას დამატებითი შემთხვევითი ვექტორებით ჩარჩოების სტანდარტულ სიგრძემდე. ფუნქცია *createProbe* წარმოქმნის ასეთ სრულმეტრაჟიან ინტერვალს.

მსგავსების ზომების რაოდენობა მოცემული მეხსიერების თანმიმდევრობისთვის და ინტერვალი გამოითვლება ფუნქციით *getSimilarity*. ეს ფუნქცირბი დანერგილია C პროგრამულ გარემოში და ჩართულია როგორც ბინარული mex ფაილი MATLAB-ში.

MATLAB-ის ინტერფეისი არის ფუნქცია *mexFunction*. იგი იღებს საწყის მონაცემებს გაშვებული სკრიპტებიდან, იწყებს ფაქტობრივ გამოთვლებს და ადგენს შედეგებს. C-ში მსგავსების ფუნქციის განხორციელება გამოთვლითი OPENMP26-ის გამოყენებით, პარალელური დამუშავების საშუალებას იძლევა.

ექსპერიმენტში მეტყველების რეალური მონაცემების გამოყენებისას, კორპუსის ქვემიმდევრობათა უმეტესობა არ არის მოცემული ინტერვალის მსგავსი და ინტუიტიურადაც არ არის შესაბამისი მისი სეგმენტაციისთვის.

ყველა ექსპერიმენტის შესრულებისათვის აგებულია ახალი ინტერვალი და ახალი მეხსიერების თანმიმდევრობა. “მელის სიხშირის ცეპსტრალური კოეფიციენტის” (MFCC)[1] გამოსახულება ეფექტური კადრების სიხშირით 500 ჰ, გამოიყენება მეტყველების მონაცემების დასაშიფვრად, პარამეტრები აღებულია *POLISH BOSS CORPUS* მეტყველების მონაცემებიდან .

დასკვნა. წინამდებარე ნაშრომი იძლევა პერსპექტივას მომავალი მუშაობისათვის ფონეტიკასა და ფონოლოგიაში. კომპიუტერული სიმულაციური ექსპერიმენტები მოდელების გამოკვლევისა და ჰიპოთეზების ტესტირების საშუალებას იძლევა. ამ ექსპერიმენტებს შეუძლიათ მიმართონ ემპირიულად მიუწვდომელ ფენომენებს. მკვლევარი, რომელიც შეიმუშავებს კომპიუტერული სიმულაციურ ექსპერიმენტს მეტყველებისა და ენის შესასწავლად, უნდა იცნობდეს იმ გადაწყვეტილებებს, რომლებსაც ხელოვნური ინტელექტისა და მანქანური სწავლების ორმოცდაათ წელზე მეტი ხნის კვლევა გვთავაზობს.

კელო და პლაუტი განიხილავენ *DIVA მოდელს* (მიმართულებები არტიკულატორების სიჩქარეებში) და მეტყველების შეძენისა და წარმოების ერთ-ერთ ადრინდელ სიმულაციას და აღნიშნავენ, რომ ეს „თეორიაზე ორიენტირებული მოდელები“ ეფუძნება გამარტივებულ დაშვებებს აკუსტიკური და არტიკულაციური მეტყველების მონაცემების წარმოდგენის შესახებ. (Kello & Plaut, 2004: 2356). ისინი ამტკიცებენ, რომ წარმოდგენილი მიმდევრობის სეგმენტაციის მიდგომა შეიძლება შემოწმდეს ხელოვნურ მონაცემებზე მეტყველების მსგავსი თვისებებით, როგორიცაა შემთხვევითი ხმაური, რომელიც წარმოადგენს კორპუსის აგებულებას განმეორებითი ფსევდოსეგმენტებით.

წარმოდგენილი ხმოვანთა დაჯგუფების ექსპერიმენტის შედეგები აჩვენებს, რომ ფონეტიკურ სეგმენტებში კოარტიკულაციური ინფორმაცია განთავსებულია შორ მანძილზე. ეს დასკვნა მნიშვნელოვანია მთლიანი თანმიმდევრობითი სეგმენტაციის მიდგომისათვის.

ჯონსონის კვლევამ (Johnson., 1997: 107) აჩვენა რომ სწორი კონტექსტური ინფორმაცია არსებობს მრავალსილაბური სიტყვების საწყის მარცვლებში.

ამგვარად, ფონეტიკა და ფონოლოგია უნდა განიხილებოდეს, როგორც აღქმისა და წარმოების ცვალებადობის დანერგვა ენის მომხმარებელთა განსხვავებული ინდივიდუალური თვისებების გამო.

რამდენადაც, ენა არის თვითორგანიზებული ფენომენი, რომელიც განსახიერებულია ადამიანის ანალოგურ ტვინში, ეს უკანასკნელი მუშაობს უაღრესად პარალელურად, ასინქრონულად და ასოციაციურად. ამიტომ ენის აღწერისას მიზანშეწონილია რეალური რიცხვებისა და წესების გამოყენება, ნაცვლად ალბათობებისა, რათა მივიღოთ რეალისტური შედეგი.

[1] (MFCC - გამოიყენება ხმის სპექტრული მახასიათებლების წარმოსაჩენად ისე, რომ კარგად შეეფერება მანქანური სწავლების სხვადასხვა ამოცანებს, როგორიცაა მეტყველების ამოცნობა).

გამოყენებული ლიტერატურა

Brent, M. (1999). Speech segmentation and word discovery: a computational perspective. *Trends in Cognitive Sciences*, 3(8):294–301;

Boersma, P., & Hamann, S. (2008). The evolution of auditory dispersion in bidirectional constraint grammars. *Phonology*, 25:217–270;

Baker, A. (2008). Computational approaches to the study of language change. *Language and Linguistics Compass*, 2(2):289–307;

Duran, D., Schütze, H., Möbius, B., & Walsh, M. (2020). A computational model of unsupervised speech segmentation for correspondence learning. *Research on Language & Computation*, 8(2-3):133–168, 2020;

Frank, M., Goldwater, S., Mansinghka, V., Griffiths, T., & Tenenbaum, J. (2007). Modeling human performance in statistical word segmentation. In *Proceedings of the 29th Annual Meeting of the Cognitive Science Society*, pages 281–286, Nashville, Tennessee, USA.

Fant, G. (2004). Phonetics and phonology in the last 50 years. In Janet Slifka, Sharon Manuel, and Melanie Matthies, editors, *From Sound to Sense: 50+ Years of Discoveries in Speech Communication*;

Goldsmith, J., & Xanthos, A. (2009). Learning phonological categories. *Language*, 85(1):4–38.

Johnson, K. (1997). The auditory/perceptual basis for speech segmentation. *OSU Working Papers in Linguistics*, 50:101–113.

Kello, C., & Plaut, D. (2004). A neural network model of the articulatory-acoustic forward mapping trained on recordings of articulatory parameters. *Journal of the Acoustical Society of America*, 116:2354–2364;

Liljencrants, J., & Lindblom B. (1972). Language, Cited by 1395 — Vol. 48, No. 4, 839-862.

Mitra, V., Nam, H., Espy-Wilson, C., Saltzman, E., & Goldstein, L. (2012). Recognizing articulatory gestures from speech for robust speech recognition. *Journal of the Acoustical Society of America*, 131(3):2270–2287.

Metropolis, N. (1987). The beginning of the Monte Carlo method. *Los Alamos Science*, 15:125–130.

Massaro, D. W. (1996). Integration of multiple sources of information in language processing. In T. Inui & J. L. McClelland (Eds.), *Attention and performance 16: Information integration in perception and communication* (pp. 397–432);

Perruchet, P., & Vinter, A. (2019). Parser: A model for word segmentation. *Journal of Memory and Language*, 39:246–263;

Sun, R. (2008). Introduction to computational cognitive modeling. In *the Cambridge Handbook of Computational, Psychology*, chapter 1, pages 3–19.

Turing, A. (1950). Computing machinery and intelligence. *Mind*, 59:433–460, New Series.

Van Rijsbergen, C. (1979). *Information Retrieval*. 2nd ed. London Butterworths. 208 pp.

Xanthos, A. (2003). An incremental implementation of the utterance-boundary approach to speech segmentation, In *Proceedings of the 14th Meeting of Computational Linguistics in the Netherlands*, pages 171–180.

REFERENCE

Brent, M. (1999). Speech segmentation and word discovery: a computational perspective. *Trends in Cognitive Sciences*, 3(8):294–301;

Boersma, P., & Hamann, S. (2008). The evolution of auditory dispersion in bidirectional constraint grammars. *Phonology*, 25:217–270;

Baker, A. (2008). Computational approaches to the study of language change. *Language and Linguistics Compass*, 2(2):289–307;

Duran, D., Schütze, H., Möbius, B., & Walsh, M. (2020). A computational model of unsupervised speech segmentation for correspondence learning. *Research on Language & Computation*, 8(2-3):133–168, 2020;

Frank, M., Goldwater, S., Mansinghka, V., Griffiths, T., & Tenenbaum, J. (2007). Modeling human performance in statistical word segmentation. In *Proceedings of the 29th Annual Meeting of the Cognitive Science Society*, pages 281–286, Nashville, Tennessee, USA.

Fant, G. (2004). Phonetics and phonology in the last 50 years. In Janet Slifka, Sharon Manuel, and Melanie Matthies, editors, *From Sound to Sense: 50+ Years of Discoveries in Speech Communication*;

Goldsmith, J., & Xanthos, A. (2009). Learning phonological categories. *Language*, 85(1):4–38.

Johnson, K. (1997). The auditory/perceptual basis for speech segmentation. *OSU Working Papers in Linguistics*, 50:101–113.

Kello, C., & Plaut, D. (2004). A neural network model of the articulatory-acoustic forward mapping trained on recordings of articulatory parameters. *Journal of the Acoustical Society of America*, 116:2354–2364;

Liljencrants, J., & Lindblom B. (1972). Language, Cited by 1395 — Vol. 48, No. 4, 839-862.

Mitra, V., Nam, H., Espy-Wilson, C., Saltzman, E., & Goldstein, L. (2012). Recognizing articulatory gestures from speech for robust speech recognition. *Journal of the Acoustical Society of America*, 131(3):2270–2287.

Metropolis, N. (1987). The beginning of the Monte Carlo method. *Los Alamos Science*, 15:125–130.

Massaro, D. W. (1996). Integration of multiple sources of information in language processing. In T. Inui & J. L. McClelland (Eds.), *Attention and performance 16: Information integration in perception and communication* (pp. 397–432);

Perruchet, P., & Vinter, A. (2019). Parser: A model for word segmentation. *Journal of Memory and Language*, 39:246–263;

Sun, R. (2008). Introduction to computational cognitive modeling. In *the Cambridge Handbook of Computational, Psychology*, chapter 1, pages 3–19.

Turing, A. (1950). Computing machinery and intelligence. *Mind*, 59:433–460, New Series.

Van Rijsbergen, C. (1979). *Information Retrieval*. 2nd ed. London Butterworths. 208 pp.

Xanthos, A. (2003). An incremental implementation of the utterance-boundary approach to speech segmentation, In *Proceedings of the 14th Meeting of Computational Linguistics in the Netherlands*, pages 171–180.