

پژوهش‌های زبانی، سال ۱۱، شماره ۱، بهار و تابستان ۱۳۹۹



10.22059/jolr.2020.298600.666587

Print ISSN: 1026-2288-Online ISSN: 2676-3362

<https://jolr.ut.ac.ir>

Durational Parameters of Speech Rhythm, A Measure for Detection of the Persian Speakers' Use of Disguise?

Maral Asiaee

Ph. D Candidate in Linguistics, Alzahra University, Tehran, Iran,

Mandana Nourbakhsh¹

Associate Professor Linguistics, Alzahra University, Tehran, Iran

Received: February 28, 2020 & Accepted: June 9, 2020

Abstract

Using the auditory-acoustic approach, the present study examines the possibility of using durational acoustic parameters of speech rhythm for detecting the Persian speakers' use of Azeri as a form of voice disguise. To do so, continuous speech of 5 speakers of Standard Persian and 5 speakers of Azeri, Tabrizi variety, were chosen for the acoustic and statistical analysis after completing the validity procedures. All Persian speakers were monolingual and all Azeri speakers were bilingual speakers of Azeri and Persian, who spoke Azeri as their mother tongue and Persian as their second language. Each Persian speaker was asked to narrate a lifetime experience once in Persian (Persian- Persian data), and once as an imitation of Azeri (Persian- Azeri data). Azeri speakers were also asked to narrate a lifetime experience once in Azeri (Azeri- Azeri data) and once in Persian (Azeri-Persian Data). Persian-Azeri data is the type that is considered as the disguised data in this survey. The recorded data were then annotated in five tiers: segment, CV-segment, CV-segment interval, CV-interval and syllable. In order to control the effect of any unwanted variable, one minute (± 5 seconds SD) of each sound file was extracted for further acoustic and statistical analyses. A Praat script, DurationAnalyzer, was used to automatically calculate the acoustic correlates of durational parameters of speech rhythm. These parameters are: %V (the proportion which speech is vocalic), ΔC (ln) (standard deviation of the natural-log normalized duration of consonantal intervals), ΔV (ln) (standard deviation of the natural-log normalized duration of vocalic intervals), nPVI- V (rate-normalized averaged durational differences between consecutive vocalic intervals) and syllable rate. Results revealed there was a significant difference between the proposed types of data and that %V and syllable rate best discriminated between them; however, none of the above-mentioned parameters were significantly different between Persian-Azeri and Azeri-Persian data.

Keywords: Voice Disguise, Speech Rhythm, Durational Parameters of Speech Rhythm, Forensic Phonetics, Persian, Azeri, Tabrizi Variety.

1. Email of the corresponding author: nourbakhsh@alzahra.ac.ir

پارامترهای دیرش - بنیاد ریتم گفتار، سنجهای برای تشخیص تقلب

فارسی‌زبانان در گفتار؟

مارال آسیائی

دانشجوی دکتری زبان‌شناسی، دانشگاه الزهرا (س)، تهران، ایران

ماندانا نوربخش^۱

دانشیار زبان‌شناسی، دانشگاه الزهرا (س)، تهران، ایران

تاریخ دریافت مقاله: ۹۸/۱۲/۹؛ تاریخ پذیرش مقاله: ۹۹/۳/۲۰

(از ص ۱ تا ص ۲۳)

علمی - پژوهشی

چکیده

پژوهش حاضر با استفاده از رویکردی شنیداری-صوت‌شناختی به بررسی امکان استفاده از پارامترهای صوت‌شناختی و دیرش-بنیاد ریتم گفتار در تشخیص تقلب فارسی‌زبانان در تولید گونه ترکی آذری می‌پردازد. از این‌رو، گفتار پیوسته و محاوره‌ای پنج گویشور زبان فارسی معیار و پنج گویشور ترکی آذری گونه تبریزی پس از انجام مراحل اعتبارسنجی مورد تجزیه و تحلیل صوت‌شناختی و آماری قرار گرفت. تمامی گویشوران فارسی تک‌زبان و گویشوران ترکی آذری دوزبان (ترکی آذری و فارسی) بودند. از هر یک از گویشوران فارسی‌زبان درخواست شد تا ابتدا خاطره‌ای را به زبان فارسی (فارسی-فارسی) و سپس همان خاطره را به صورت تقلیدی از ترکی آذری (فارسی-ترکی) بیان نمایند. از گویشوران ترکی آذری نیز درخواست شد تا خاطره‌ای را یک‌بار به ترکی آذری (ترکی-ترکی) و بار دیگر به فارسی (ترکی-فارسی) نقل کنند. داده‌های فارسی-ترکی، داده‌هایی هستند که در این پژوهش به عنوان داده‌های تقلب فارسی‌زبانان از ترکی آذری شناخته می‌شوند. پس از ضبط داده‌ها، نمونه‌های صوتی در پنج لایه برچسب‌گذاری شدند. به منظور یکسان‌سازی و حذف هرگونه متغیر اضافی از نمونه‌های آوایی، حدود یک دقیقه (با انحراف معیار ± 5 ثانیه) از گفتار گویشوران، به منظور انجام بررسی‌های صوت‌شناختی و آماری انتخاب گردید. پارامترهای صوت‌شناختی و دیرش-بنیاد ریتم گفتار با استفاده از برنامه خودکار تجزیه و تحلیل دیرش استخراج و مورد واکاوی قرار گرفتند. این پارامترها عبارت‌اند از: میزان واکه‌ای بودن گفتار، انحراف معیار دیرش مناطق همخوانی و واکه‌ای نرمال‌سازی شده، شاخص تغییر دو به دوی واکه‌ای نرمال‌سازی شده و سرعت تولید. نتایج بررسی آماری که با استفاده از نرم‌افزار آماری اسپاس انجام شد، نشان داد که پارامترهای مناسب در ایجاد تمایز میان انواع داده‌های پژوهش حاضر، پارامترهای میزان واکه‌ای بودن گفتار و سرعت تولید هستند. اگرچه، این دو پارامتر در ایجاد تمایز کلی میان انواع داده‌های این پژوهش موفق عمل کرده‌اند، هیچ‌کدام از این پارامترها توانایی ایجاد تمایز میان داده‌های ترکی-فارسی و فارسی-ترکی را نداشتند.

واژه‌های کلیدی: تقلب در گفتار، ریتم گفتار، پارامترهای دیرش-بنیاد، آواشناسی قضایی، زبان فارسی، ترکی آذری گونه تبریزی.

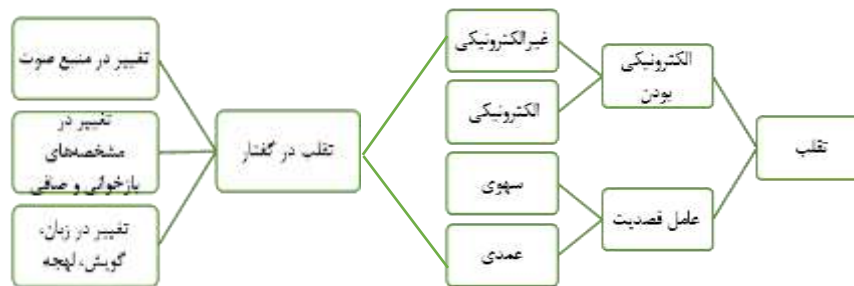
۱- مقدمه

آواشناسی قضایی^۱ «کاربرد دانش، نظریه‌ها و روش‌های آخذ شده توسط آواشناسی عمومی، در رویه‌های قضایی» تعریف شده است (یسن، ۲۰۰۸: ۶۷۱). یکی از مهم‌ترین زیر بخش‌های آواشناسی قضایی، مقایسه قضایی گفتار^۲ است. در مقایسه قضایی گفتار با تکیه بر این امر که صدای هر گویشور دارای مشخصات فرد ویژه‌ای است که امکان شناسایی فرد را فراهم می‌آورد، همواره سعی بر یافتن پارامترهایی بوده است که بتوان با استفاده از آن‌ها هویت گوینده را به درستی تشخیص داد.^(۱) در فرایند مقایسه قضایی گفتار صدای ضبط شده‌ای وجود دارد که متعلق به یک گوینده ناشناس^۳ است (به عنوان نمونه می‌توان به صدای گروگان‌گیری که تقاضای پول نموده است اشاره نمود). در عین حال صدای شخص دیگری وجود دارد که مظنون^۴ قلمداد می‌شود و آواشناس قضایی می‌کوشد با استفاده از سرنخ‌های صوتی و آوایی احتمال اینکه صدای فرد ناشناس همان صدای فرد مظنون باشد را بررسی نماید.

یکی از مواردی که همواره تشخیص هویت و رویه‌های مورد استفاده در آن را با چالش جدی مواجه کرده است، تقلب در گفتار^۵ است. «تقلب در گفتار شامل هرگونه فعالیتی است که از سوی گوینده به منظور مخدوش کردن و یا پنهان کردن هویت، جنسیت و سن خود صورت می‌پذیرد» (پروت و همکاران، ۲۰۰۷). کانزل (۲۰۰۰: ۱۵۰) تقلب در گفتار را «هر نوع تغییر ارادی و عمدی در مشخصه‌های واک^۶ گفتار و زبان به منظور پنهان کردن هویت فرد» تعریف می‌نماید. تقلب در گفتار می‌تواند با استفاده از روش‌های متعددی انجام پذیرد که از آن جمله می‌توان به هر نوع تغییر در وضعیت دستگاه گفتار (همانند گرد کردن لب‌ها، تنگ کردن حنجره و مانند آن) و یا استفاده از دستگاه‌های الکترونیکی و نرم‌افزارهای مختلف اشاره نمود. این تغییرات می‌تواند در جهت تغییر مشخصه‌های صدای یک فرد (تغییر صدا^۷) و یا تغییر صدای فرد به گونه‌ای که صدا، به

1. forensic phonetics
2. forensic voice comparison
3. unknown or disputed sample
4. known or suspect sample
5. voice disguise
6. Features of voice
7. voice transformation

صدای فردی دیگر شبیه شود (تبدیل صدا^۱) باشد. پروت و شله (۲۰۱۲) از تغییر صدا با عنوان تقلب در گفتار و از تبدیل صدا تحت عنوان جعل گفتار^۲ نام می‌برند. علاوه بر طبقه‌بندی انواع تقلب به الکترونیکی و غیر الکترونیکی عامل قصدیت^۳ نیز در طبقه‌بندی انواع تقلب توسط پژوهشگران مختلف از جمله کانزل (۲۰۰۰)، رادمن (۱۹۹۸) و ماستهف (۱۹۹۶) در نظر گرفته شده است. آنچه اغلب از آن تحت عنوان تقلب در گفتار یاد می‌شود، تغییرات عامدانه‌ای است که به صورت الکترونیکی و غیر الکترونیکی ایجاد شود. در کنار طبقه‌بندی انواع تقلب به الکترونیکی و غیر الکترونیکی، سهوی و عامدانه، می‌توان انواع تقلب عامدانه و غیر الکترونیکی را - که از این پس در سرتاسر مقاله به عنوان تقلب در گفتار شناخته می‌شود - به سه گروه کلی دسته‌بندی کرد. دسته اول شامل مواردی می‌گردد که تغییر در منبع صوت ایجاد شده است؛ از آن جمله می‌توان به زیروبمی افزوده^۴، زیروبمی کاهیده^۵، نجوا^۶ و مانند آن اشاره کرد. طبقه دوم دربرگیرنده انواع تغییر در مشخصه‌های بازخوانی^۷ و صافی صوت است که شامل قرار دادن شی خارجی در مجرای گفتار (دهان)^۸، کم/پر خیشومی‌شدگی^۹ و مانند آن است. گروه آخر نیز تغییرات در زبان، گویش و لهجه را شامل می‌شود. انگاره شماره ۱، طبقه‌بندی انواع تقلب را بازنمایی می‌کند.



شکل ۱: طبقه‌بندی انواع تقلب در گفتار

1. voice conversion
2. voice forgery
3. intentionality
4. raised pitch
5. lowered pitch
6. whisper
7. resonance features
8. foreign object in the vocal tract
9. hypo-/hypernasality

اگر در هر بزهی که صورت می‌پذیرد، یکی از انواع تقلب‌ها به کار گرفته شود، فرایند مقایسه قضایی گفتار با مشکل مواجه می‌گردد. یکی از مهم‌ترین انواع تقلب که روند تشخیص هویت گوینده را دچار چالش جدی می‌کند، تقلید گویش، لهجه و یا زبانی غیر از زبان گویش، لهجه و زبان خود فرد است (اریکسون، ۲۰۱۰). از این‌رو، شوی (۱۹۹۵) و هولین (۲۰۰۲) معتقدند که یکی از مواردی که از ابتدای بررسی‌های قضایی مورد توجه آواشناسان و کارشناسان قضایی قرار می‌گیرد بررسی گویش (گونه زبانی) افراد است. انتخاب پارامترهایی که دارای ویژگی‌های منحصربه‌فرد (فرد ویژه) بوده و با استفاده از آن‌ها بتوان میزان موفقیت فرایند مقایسه قضایی گفتار را افزایش داد همواره از چالش‌هایی بوده که آواشناس قضایی با آن روبرو است. توجه به این نکته که هیچ‌گاه، هیچ دو صدای تولیدشده توسط یک فرد دارای ویژگی‌های صوت‌شناختی صد درصد یکسان نبوده و عوامل متعددی همچون احساسات مختلف نظیر اضطراب، ترس، استرس و بیماری بر گفتار فرد تأثیرگذار است، نیز از اهمیت بسزایی در مطالعات در حوزه آواشناسی قضایی برخوردار است. از این‌رو، از مهم‌ترین ویژگی‌هایی که اکثر پژوهشگران در حوزه آواشناسی قضایی از جمله نولان (۱۹۸۳: ۱۱) و ولف (۱۹۷۲: ۲۰۴۴) به عنوان ویژگی یک پارامتر مناسب از آن یاد می‌کنند می‌توان به تغییرات بین-گوینده^۱، تغییرات درون-گوینده^۲، مقاوم بودن در برابر هرگونه تقلب و یا تقلید^۳، در دسترس بودن^۴ و قابل اندازه‌گیری بودن^۵ اشاره نمود.

در اغلب پژوهش‌هایی که در حوزه آواشناسی قضایی صورت گرفته است، مشخصه‌های طیفی گفتار^۶ نظیر فرکانس پایه (اسکارنیتزل و ونکوا، ۲۰۱۷؛ کینوشیتا، ۲۰۰۵)، فرکانس سازه‌ای (آسیائی و همکاران، ۲۰۱۹؛ اسدی و همکاران، ۲۰۱۸؛ گولد و همکاران، ۲۰۱۳؛ نولان و گریگوراس، ۲۰۰۵) و مشخصه‌های بسته طیفی^۷ (آمینو و آرای، ۲۰۰۹) مورد بررسی قرار گرفته‌اند؛ با این وجود، توجه به مشخصه‌های زمانی گفتار^۸ از جمله بررسی پارامترهای دیرش-بنیاد ریتم گفتار نیز رو به فزونی نهاده است، چراکه به عقیده دلوو و

1. between-speaker variability
2. within-speaker variability
3. mimicry
4. accessibility
5. measurability
6. spectral features of speech
7. spectral envelope features
8. temporal features of speech

همکاران (۲۰۱۵) ویژگی‌های فردی تنها بر مؤلفه‌های فرکانسی گفتار تأثیر نمی‌گذارند؛ بلکه بر سازمان‌بندی زمانی گفتار نیز اثرگذار هستند.

تاکنون پژوهش‌های متعددی بر روی زبان‌های مختلف اعم از زبان‌های هجا-زمان^۱ و تکیه-زمان^۲ صورت گرفته و توانایی مشخصه‌های صوت‌شناختی ریتمیک گفتار در ایجاد تمایز میان زبان‌های مختلف و همچنین قابلیت این مشخصه‌ها برای تمایزگذاری میان گویشوران مختلف را به محک آزمایش گذاشته‌اند. نخستین پژوهش‌هایی که با استفاده از پارامترهای کمی و دیرش-بنیاد در حوزه طبقه‌بندی ریتمیک گفتار انجام شده است، پژوهش‌های راموس و همکاران (۱۹۹۹) و گرابه و لو (۲۰۰۲) است. راموس و همکاران (همان) با معرفی پارامترهایی نظیر میزان واکه‌ای بودن گفتار (%V)، انحراف مناطق همخوانی^۳ (ΔC) و انحراف معیار مناطق واکه‌ای^۴ (ΔV) و گرابه و لو (همان) با معرفی شاخص نرمال‌شده تغییر دو به دو^۵ در مناطق واکه‌ای (nPVI-V) زبان‌های مختلفی اعم از هجا-تکیه- و مورا-زمان^۶ را طبقه‌بندی نمودند. بر اساس نتایج به دست آمده از این پژوهش‌ها، زبان‌های تکیه-زمان، دارای (ΔC) بالاتر (نمایانگر گوناگونی بالا در مناطق همخوانی) و (%V) پایین‌تر (نمایانگر گوناگونی بالا در مناطق واکه‌ای) و زبان‌های هجا-زمان دارای (ΔC) پایین‌تر و (%V) بالاتر هستند. میزان (ΔC) بالاتر در زبان‌های تکیه-زمان به دلیل دارا بودن ساختار هجایی پیچیده‌تر (که خود به دلیل مجاز بودن خوشه‌های همخوانی با تعداد همخوان زیادتر در یک زبان است) و میزان (%V) پایین‌تر به دلیل بیشتر بودن نسبت همخوان به واکه و وقوع فرایند کاهش واکه‌ای در این زبان‌ها است (دوئر، ۱۹۸۳).

با وجود پژوهش‌های متعددی که توسط پژوهشگران مختلف در طبقه‌بندی ریتمیک گفتار صورت گرفته است (باری و همکاران، ۲۰۰۳؛ وایت و ماتیس، ۲۰۰۷؛ دلوو، ۲۰۱۰؛ رتکه و اسمیت، ۲۰۱۵؛ تقوی و ابوالحسنی زاده، ۲۰۱۶) با بررسی پیشینه مطالعات در این حوزه، اندک پژوهش‌هایی را می‌توان یافت که تأثیر استفاده از تقلب بر طبقه‌بندی ریتمیک گفتار (لیمان و کولی، ۲۰۱۵) را مورد بررسی قرار داده‌اند. از اینرو، در این

1. syllable-timed
2. stress-timed
3. c- interval
4. v-interval
5. pairwise variability index
6. mora-timed

پژوهش برآنیم تا ابتدا ساختار ریتمیک زبان فارسی معیار و زبان ترکی آذری^۱ گونه تیریزی را بر اساس پارامترهای ریتمیک صوت‌شناختی مطرح‌شده راموس و همکاران (۱۹۹۹)، گرابه و لو (۲۰۰۲) و دلوو (۲۰۱۰) مورد مطالعه قرار دهیم و سپس به بررسی امکان تشخیص تقلب فارسی‌زبانان در استفاده از گونه ترکی آذری با استفاده از این پارامترها می‌پردازیم.^(۲)

۲- روش پژوهش

در این بخش اطلاعات مربوط به اعتبارسنجی داده‌های پژوهش و شرکت‌کنندگان در پژوهش، داده‌های آوایی، شیوه تقطیع داده‌ها و پارامترهای صوت‌شناختی ارائه خواهد شد.

۱-۲ داده‌های پژوهش

داده‌های این پژوهش در قالب گفتار پیوسته جمع‌آوری شده است. ضبط تمامی داده‌ها با فرکانس نمونه‌برداری ۴۴۱۰۰ هرتز و با استفاده از دستگاه ضبط حرفه‌ای ZOOM H6 در محیطی کاملاً آرام صورت گرفته است. دستگاه ضبط به صورت مورب و با فاصله ۲۰ سانتی‌متری از دهان گویشوران قرار داده شد.

داده‌های این پژوهش به دو دسته کلی طبقه‌بندی می‌شوند. دسته اول شامل گفتار گویشوران دوزبانه‌ای است که ترکی آذری گونه تیریزی زبان مادری آن افراد بوده و همگی آن‌ها به زبان فارسی به عنوان زبان دوم مسلط بوده‌اند. داده‌های گروه دوم، دربرگیرنده گفتار فارسی‌زبانانی است که سعی در تقلید گونه زبانی ترکی آذری داشته‌اند. از گویشوران ترک‌زبان درخواست شد تا تلخ‌ترین خاطره زندگی خود را یک‌بار به ترکی (داده نوع ۱) و بار دیگر به فارسی (داده نوع ۲) بازگو نمایند. از گویشوران فارسی‌زبان نیز درخواست شد تا تلخ‌ترین خاطره زندگی خود را یک‌بار به فارسی (داده نوع ۳) و بار دیگر به صورت تقلیدی از زبان ترکی نقل کنند (داده نوع ۴).^(۳)

۲-۲ اعتبارسنجی

اعتبارسنجی این پژوهش در سه مرحله و بر روی داده‌های نوع ۲ و ۴ صورت پذیرفت. نخستین مرحله از اعتبارسنجی در مرحله تقطیع و برچسب‌گذاری داده‌ها انجام گرفت. در ابتدا تمامی داده‌ها توسط اعضای گروه هسته پژوهشی آواشناسی برچسب‌گذاری شدند. برچسب‌گذاری داده‌های گونه زبانی ترکی آذری توسط گویشور این گونه صورت گرفته است. پس از برچسب‌گذاری، صحت این روند توسط نویسنده نخست مقاله بررسی شد. در

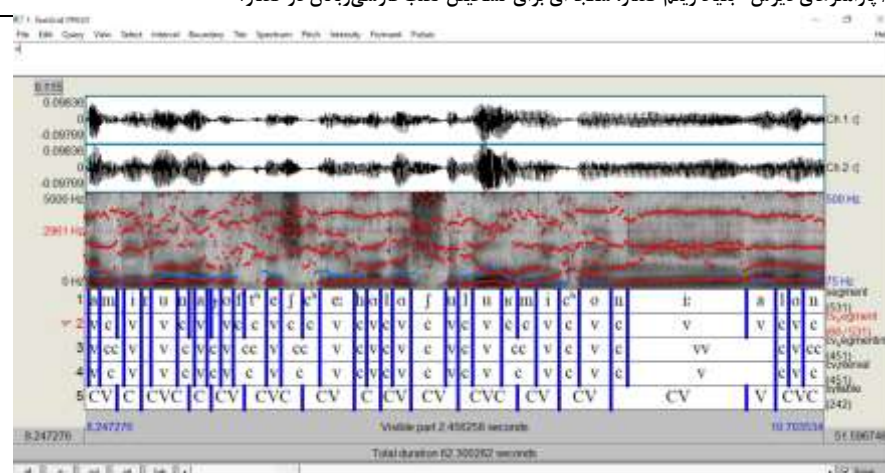
مرحله دوم، جهت کسب اطمینان از داده‌های انتخاب‌شده برای گونه زبانی ترکی، در مواردی که گویشوران این گونه زبانی به زبان فارسی صحبت کرده‌اند - تمامی داده‌ها که شامل گفتار هشت گویشور گونه ترکی آذری می‌شد برای پنج فرد دانش‌آموخته در رشته زبان‌شناسی به صورت تصادفی پخش شد و از آن‌ها درخواست شد که از میان گزینه‌های ترکی، کردی، عربی و هیچ‌کدام یک گزینه را انتخاب نمایند. پس از جمع‌آوری نتایج از تمامی پنج دانش‌آموخته زبان‌شناسی، داده‌هایی که بیش از هشتاد درصد افراد به عنوان داده مربوط به گونه ترکی آذری انتخاب کرده بودند، به عنوان داده آن‌گونه انتخاب و مرحله نهایی اعتبارسنجی روی آن‌ها صورت پذیرفت. همین مراحل برای داده‌های دسته دوم که در آن فارسی‌زبانان گونه زبانی ترکی آذری را تقلید کرده‌اند نیز صورت پذیرفت. به این ترتیب که از دانش‌آموختگان زبان‌شناسی درخواست شد تا از میان گونه‌های ترکی، کردی، عربی و هیچ‌کدام یک گزینه را به عنوان گزینه‌ای که فرد فارسی‌زبان سعی در تقلید آن داشته است انتخاب کنند. در مواردی که گونه تقلیدشده توسط شرکت‌کننده در پژوهش، بیش از هشتاد درصد آرا دانش‌آموختگان را کسب می‌کرد، آن‌گونه به عنوان گونه تقلیدشده توسط گویشوران انتخاب می‌شد. در آخرین مرحله از اعتبارسنجی، داده‌هایی که بیش از هشتاد درصد دانش‌آموختگان در مرحله دوم به عنوان داده منتخب هرگونه انتخاب کرده بودند، مجدداً مورد بررسی قرار گرفتند. در این مرحله از اعتبارسنجی، ابتدا داده‌های گونه ترکی آذری که گویشوران در آن به زبان دوم خود یعنی فارسی صحبت می‌کردند و سپس داده‌هایی که در آن گویشوران فارسی‌زبان سعی در تقلید گونه ترکی آذری داشته‌اند، مورد بررسی قرار گرفتند. در مورد داده‌های اول، گفتار گویشوران گونه ترکی آذری به زبان فارسی، برای دانش‌آموختگان پخش شد. سپس از آن‌ها سؤال شد که به چه میزان گفتار این گویشوران به گفتاری که آن‌ها به عنوان نمونه‌ای از گفتار یک فرد ترک‌زبان در نظر دارند، نزدیک است؟ این افراد می‌بایست بین گزینه‌های صفر تا بیست درصد، بیست تا چهل درصد، چهل تا شصت درصد، شصت تا هشتاد درصد و هشتاد تا صد درصد یک گزینه را انتخاب می‌کردند. در خصوص داده‌های دوم، داده‌های منتخب مرحله دوم به صورت تصادفی برای شرکت‌کنندگان در مرحله اعتبارسنجی پخش شد و از آن‌ها سؤال شد که فرد گویشور چقدر در تقلید گونه ترکی آذری موفق عمل کرده است. از دانش‌آموختگان خواسته شد به ازای هر صوت یک گزینه را در همان بازه درصدی قبلی انتخاب می‌کردند. در نهایت، از میان تمامی داده‌های نوع اول و نوع دوم، آن داده صوتی که بیش از ۶۰٪ افراد گزینه ۸۰٪ تا ۱۰۰٪ را برای آن انتخاب کرده بودند، به عنوان نمونه

برتر صوتی برای تحلیل صوت‌شناختی و آماری انتخاب شد. بدین ترتیب پس از انجام تمامی مراحل اعتبارسنجی، گفتار ۵ گویشور از گونه ترکی آذری و پنج گویشور زبان فارسی انتخاب شد. میانگین سنی گویشوران فارسی‌زبان، ۳۵/۷۵ سال با انحراف معیار $\pm ۷/۲۵$ و میانگین سنی گویشوران ترک‌زبان ۳۳/۵ سال با انحراف معیار $\pm ۶/۳۴$ است. لازم به ذکر است به منظور دستیابی به جامعه آماری همگن تمامی گویشوران به لحاظ عدم وجود اختلالات گفتاری و میزان تحصیلات کنترل شده‌اند. پس از اتمام مرحله اعتبارسنجی، به منظور یکسان‌سازی و حذف هرگونه متغیر اضافی از نمونه‌های آوایی، حدود یک دقیقه (با انحراف معیار ± ۵ ثانیه) از گفتار گویشوران منتخب برای هر نوع داده، به منظور انجام بررسی‌های صوت‌شناختی و آماری انتخاب گردید.

۲-۳ شیوه برچسب‌گذاری و محاسبه پارامترهای ریتم گفتار

تقطیع و برچسب‌گذاری داده‌های آوایی مورد استفاده در این پژوهش با استفاده از نرم‌افزار پرات^۱ و ویرایش ۰۶،۱،۰۹ صورت پذیرفت. به ازای هر نمونه صوتی یک پوشه متنی با پنج لایه ایجاد شد و داده‌های آوایی با علائم آواشناختی آی‌پی‌ای^۲ برچسب‌گذاری شدند. لایه اول مربوط به بخش آوایی است که در آن همخوان‌ها و واکه‌ها با علائم آوانگاری خود مشخص شده است. در لایه دوم بخش‌های همخوانی با علامت C و بخش‌های واکه‌ای با علامت V مشخص شده‌اند. لایه سوم بخش‌های مناطق همخوانی و مناطق واکه‌ای را به همراه تعداد عناصر آن‌ها به نمایش می‌گذارد. لایه چهارم دربرگیرنده مناطق واکه‌ای و همخوانی است و لایه پنجم سطح هجا را نشان می‌دهد. لایه‌های اول و پنجم به صورت دستی توسط اعضای هسته پژوهشی آواشناسی قضایی دانشگاه الزهرا (س) ایجاد شده است. لایه‌های دوم، سوم و چهارم نیز به صورت دستی توسط نویسنده اول مقاله ایجاد شده‌اند.

1. Praat
2. IPA



شکل ۲: نمونه‌ای از تقطیع و برچسب‌گذاری داده‌های آوایی

استخراج پارامترهای دیرش - بنیاد با استفاده از برنامه خودکار تجزیه و تحلیل دیرش^۱ ویرایش ۰,۰۳ صورت پذیرفت. این برنامه به صورت یک افزونه به نرم‌افزار پرات اضافه شده و با اجرای آن مقادیر مربوط به مشخصه‌های دیرشی در سطح بخش آوایی، هجاها، مناطق همخوانی و واکه‌ای و سایر بخش‌ها از لایه‌های مورد نظر در شبکه‌های متنی ایجاد شده قابل استخراج است. پس از استخراج پارامترهای دیرش - بنیاد، نتایج بدست آمده با استفاده از نرم‌افزار اسپاس^۲ ویرایش ۲۲ مورد تجزیه و تحلیل آماری قرار گرفتند. در این پژوهش طبقه‌بندی ریتمیک گفتار با استفاده از پارامترهای دیرش - بنیاد مرتبط با مناطق واکه‌ای و همخوانی و هجا صورت می‌پذیرد. منطقه همخوانی، منطقه‌ای در گفتار است که شامل یک یا چند همخوان بوده و یک واکه (و یا مکث) قبل و بعد از همخوان (ها) قرار می‌گیرد. مناطق همخوانی می‌توانند به لحاظ ساختاری ساده و یا پیچیده باشند. مناطق همخوانی ساده تنها شامل یک همخوان و مناطق همخوانی پیچیده دارای حداقل دو همخوان هستند. منطقه واکه‌ای، منطقه‌ای در گفتار است که شامل یک یا چند واکه (و یا بخش‌های واکه‌ای نظیر واکه دوگانه^۳ و یا واکه سه‌گانه^۴) بوده و یک همخوان (و یا مکث) قبل یا بعد از واکه‌ها قرار گیرد. بخش‌های واکه‌ای غالباً در هسته هجا قرار گرفته و قبل و

1. Duration Analyzer: Available at <http://www.pholab.uzh.ch/static/volker/software.html>

2. SPSS

3. diphthong

4. triphthong

یا بعد از آن‌ها یک همخوان قرار می‌گیرد. به این ترتیب، بخش‌های واکه‌ای معمولاً فقط شامل یک واکه و یا یک واکه دوگانه هستند. اگرچه، امکان تلاقی دو واکه در مرز واژه‌ها نیز وجود دارد.

۴-۲ پارامترهای صوت‌شناختی

پارامترهای مورد سنجش در این بررسی عبارت‌اند از: میزان واکه‌ای بودن گفتار (%V)، انحراف معیار دیرش مناطق همخوانی ($\Delta C(\ln)$) و واکه‌ای ($\Delta V(\ln)$) نرمال‌سازی شده، شاخص تغییر دو به دوی واکه‌ای (nPVI-V) که بر اساس سرعت نرمال‌سازی شده است. تمامی این پارامترها در سطح مناطق واکه‌ای و همخوانی اندازه‌گیری می‌شوند. آخرین پارامتر که در سطح هجا اندازه‌گیری می‌شود عبارت است از سرعت هجا (سرعت تولید^۱). راموس و همکاران (۱۹۹۹) با تکیه بر اینکه زبان‌های تکیه-زمان دارای ساختار هجایی پیچیده‌تری نسبت به زبان‌های هجا-زمان هستند، معتقدند که این زبان‌ها (تکیه-زمان) دارای مناطق همخوانی پیچیده‌تری نسبت به زبان‌های هجا-زمان هستند. از اینرو، ضروری است در سطح صوت‌شناختی، این تمایز، از طریق گوناگونی در دیرش مناطق همخوانی منعکس گردد. بدین ترتیب، انتظار می‌رود در زبان‌هایی که دارای پیچیدگی‌های همخوانی بیش‌تری هستند، گوناگونی در سطح دیرش مناطق همخوانی بالاتر باشد. در سطح واکه‌ای، بسته به رخداد و یا عدم رخداد کاهش واکه‌ای در یک زبان، دیرش مناطق واکه‌ای با یکدیگر متفاوت خواهد بود. از آنجاکه کاهش واکه‌ای، از طریق زمان رمزگذاری می‌شود، انتظار می‌رود در زبان‌هایی که تمایز میان واکه‌های کاهش‌یافته و واکه‌های کاهش‌نیافته در آن‌ها وجود دارد، گوناگونی در دیرش مناطق واکه‌ای بالاتر باشد. بر همین اساس، راموس و همکاران (۱۹۹۹)، با طراحی آزمایشی، انحراف معیار مناطق واکه‌ای و همخوانی و همچنین میزان پارامتر (%V) را به عنوان مقیاسی برای بررسی گوناگونی‌های دیرشی این مناطق بررسی کردند. استفاده از این پارامترها به دو علت صورت پذیرفته است: الف) زبان‌هایی که در آن‌ها کاهش واکه‌ای رخ می‌دهد (زبان‌های تکیه-زمان) شامل تعداد بیشتری از مناطق واکه‌ای، با دیرش‌های کوتاه‌تر هستند؛ بنابراین، در این زبان‌ها مدت زمان واکه‌ای بودن گفتار کمتر خواهد بود. ب) چون در زبان‌های تکیه-زمان، مناطق همخوانی به لحاظ ساختاری، دارای پیچیدگی بیشتری نسبت به زبان‌های هجا-زمان هستند، مدت زمان این مناطق در زبان‌های تکیه-زمان بیش‌تر خواهد بود. پارامتر (%V)

میزان واکه‌ای بودن جمله را نشان می‌دهد و بر اساس فرمول (۱) محاسبه می‌گردد.

$$(1) \%V = \frac{(\sum_{i=1}^{N_v} V_i)}{\sum_{i=1}^{N_c} C_i + \sum_{i=1}^{N_v} V_i} \times 100$$

در این فرمول N_v تعداد کل نمونه‌های مناطق واکه‌ای، N_c تعداد کل نمونه‌های مناطق همخوانی را نشان می‌دهد؛ به عبارت دیگر در این فرمول عبارت صورت، جمع کل دیرش مناطق واکه‌ای و عبارت مخرج جمع کل دیرش مناطق واکه‌ای و همخوانی (دیرش کل جمله در صورتی که مکث از جمله حذف شده باشد) است. انحراف معیار مناطق واکه‌ای و مناطق همخوانی از دیگر پارامترهای قابل اطمینان در تعیین ریتم زبان‌ها هستند. انحراف معیار منطقه واکه‌ای عبارت است از انحراف معیار دیرش مناطق واکه‌ای در هر جمله که با (ΔV) نشان داده می‌شود. انحراف معیار منطقه همخوانی که با (ΔC) بازنمایی می‌شود، عبارت است از انحراف معیار دیرش مناطق همخوانی در هر جمله. میزان (ΔV) و (ΔC) از طریق فرمول (۲) قابل محاسبه است.

$$(2) \Delta Int = \sqrt{\frac{N_{Int} \cdot \sum_{i=1}^{N_{Int}} Int_i^2 - (\sum_{i=1}^{N_{Int}} Int_i)^2}{N_{Int} \cdot (N_{Int} - 1)}}$$

در این فرمول N_{Int} تعداد مناطق واکه‌ای و یا همخوانی در پاره‌گفتار و Int_i دیرش i آمین منطقه واکه‌ای و یا همخوانی است.

دلوو (۲۰۱۰: ۸۷-۹۰) با اشاره به این موضوع که غالباً تحلیل آماری داده‌ها بر مبنای توزیع نرمال داده‌ها استوار است، معتقد است که دیرش مناطق گفتار لزوماً دارای توزیع نرمال نبوده و شکل توزیع گوسی^۱ مناطق همخوانی، واکه‌ای و هجا دارای چولگی^۲ و کشیدگی^۳ مثبت به سمت راست است. از اینرو، به منظور کاهش میزان چولگی و کشیدگی مثبت از نحوه محاسبه لگاریتمی برای نرمال‌سازی داده‌ها استفاده می‌شود. دلوو (۲۰۱۰) فرمول $\Delta Int(\ln)$ را به عنوان جایگزینی برای پارامترهای (ΔC) و (ΔV) معرفی می‌کند.

محاسبه پارامترهای $(\Delta C(\ln))$ و $(\Delta V(\ln))$ با استفاده از فرمول (۳) صورت می‌گیرد. در این فرمول می‌توان Int را با C و یا V جایگزین نمود.

1. Gaussian distribution
2. skewness
3. Kurtosis

$$(3) \Delta Intln = \sqrt{\frac{N_{Int} \sum_{i=1}^{N_{Int}} (\ln Int_i)^2 - (\sum_{i=1}^{N_{Int}} (\ln Int_i))^2}{N_{Int} (N_{Int} - 1)}}$$

در این فرمول، N_{Int} تعداد مناطق واکه‌ای و یا همخوانی در پاره‌گفتار و Int_i دیرش i امین منطقه واکه‌ای و یا همخوانی است. شاخص تغییر دو به دوی واکه‌ای نرمال‌سازی شده (nPVI-V) تغییرات دیرش در مناطق واکه‌ای را از طریق مقایسه دو به دوی دیرش در سطح دو منطقه واکه‌ای متوالی محاسبه می‌نماید. شاخص تغییر دو به دوی نرمال‌سازی شده از طریق فرمول (۴) محاسبه می‌گردد:

$$(4) nPVI - V = 100 \times \left[\sum_{k=1}^{m-1} \left| \frac{d_k - d_{k+1}}{(d_k + d_{k+1})/2} \right| / (m - 1) \right]$$

در این فرمول m تعداد مناطق واکه‌ای و d_k دیرش k امین منطقه واکه‌ای را نشان می‌دهد. پارامتر دیگری که همواره در بررسی ریتمیک گفتار مورد بررسی قرار می‌گیرد سرعت گفتار است. در این پژوهش، سرعت گفتار بر اساس سرعت هجا که یکی از واحدهای مناسب برای درک گفتار است، سنجیده می‌شود. سرعت هجا، تعداد هجا در ثانیه را نشان می‌دهد و از طریق فرمول (۵) محاسبه می‌گردد.

$$(5) \text{syllable-rate} = \frac{N_{syll}}{d}$$

در این فرمول N_{syll} تعداد کل هجاهای پاره‌گفتار و d دیرش کل پاره‌گفتار بدون در نظر گرفتن مکث است.

۳. تحلیل داده‌ها

اکثر پژوهش‌های پیشین زبان فارسی را به عنوان زبانی هجا-زمان طبقه‌بندی کرده‌اند (لازار، ۱۹۹۲؛ ویندفور، ۱۹۷۹)، با این وجود جیمز (۱۹۴۰: ۲۵) زبان فارسی را همانند زبان‌های انگلیسی و عربی به عنوان زبانی تکیه-زمان در نظر گرفته است. دوئر (۱۹۸۳) تمایز میان زبان‌های تکیه-زمان و هجا-زمان را به دلیل وجود ویژگی‌های واجی در نظر می‌گیرد که از مهم‌ترین آن‌ها می‌توان به ساخت هجا^۱ و کاهش واکه‌ای^۲ اشاره نمود. زبان‌های تکیه-زمان دارای ساخت هجایی پیچیده‌تری نسبت به زبان‌های هجا-زمان بوده و امکان رخداد فرایند کاهش واکه‌ای در این زبان‌ها وجود دارد (دوئر، همان). این تمایز،

1. syllable structure
2. vowel reduction

در پارامترهای صوت‌شناختی مطرح‌شده به منظور طبقه‌بندی ریتمیک گفتار، نیز قابل مشاهده است. در اغلب پژوهش‌های پیشین، پارامترهای (ΔC) (راموس و همکاران، ۱۹۹۹؛ مایرانو و رومانو، ۲۰۱۱)، $(\%V)$ (راموس و همکاران، ۱۹۹۹؛ باری و همکاران، ۲۰۰۳؛ وایت و ماتیس، ۲۰۰۷؛ دلوو، ۲۰۱۰) و $(nPVI-V)$ (گراه و لو، ۲۰۰۲؛ وایت و ماتیس، ۲۰۰۷) به عنوان پارامترهای مناسب در ایجاد تمایز ریتمیک میان زبان‌های مختلف معرفی شده‌اند. زبان‌های تکیه-زمان دارای میزان (CA) بالاتر و میزان $(\%V)$ و $(nPVI-V)$ کمتر نسبت به زبان‌های هجا-زمان هستند. از اینرو در این بخش ابتدا پارامترهای معرفی‌شده در بخش ۲-۴ در ایجاد تمایز ریتمیک میان گونه‌های زبانی فارسی و ترکی آذری گونه تبریزی مورد واکاوی قرار خواهد گرفت. سپس، تأثیر تقلب بر هر یک از متغیرهای وابسته بررسی و میزان اثرگذاری هر یک از این پارامترها در تشخیص هویت مطالعه خواهد شد.

۱-۳ نتایج و بحث

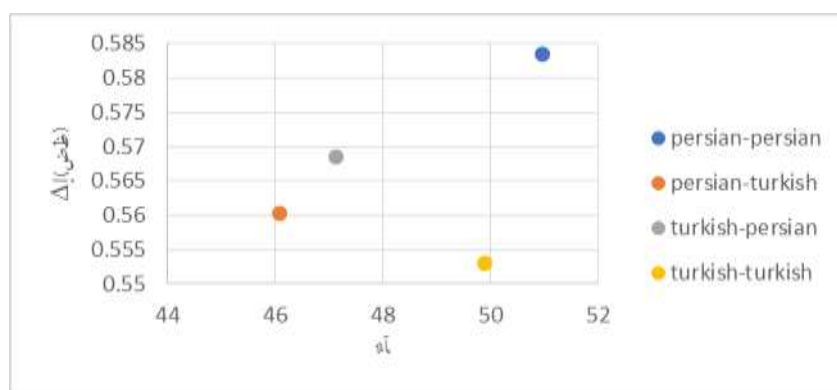
بررسی ساخت هجایی فارسی محاوره‌ای و ترکی آذری گونه تبریزی نشان می‌دهد که نمی‌توان به لحاظ پیچیدگی ساخت هجایی، تفاوتی را میان این دو زبان قائل شد. الگوی هجایی زبان فارسی به صورت $CV(C(C))$ و الگوی هجایی ترکی آذری گونه تبریزی به شکل $(C)V(C)$ (غفاروند مکاری، ۲۰۱۷؛ راه‌انداز و همکاران، ۱۳۹۳) است. به همین دلیل، انتظار می‌رود میزان پارامتر $(\Delta C(\ln))$ در این دو گونه زبانی پایین بوده و اختلاف معنی‌داری میان این دو گونه وجود نداشته‌باشد. همچنین در زبان فارسی رخداد فرایند کاهش واکه‌ای گزارش نشده است (صادقی، ۱۳۹۴)؛ از اینرو، انتظار می‌رود میزان پارامتر $(\%V)$ در این زبان بالا باشد. طبقه‌بندی ریتمیک انواع داده‌های پژوهش با استفاده از پارامترهای زبان‌های فارسی محاوره‌ای و ترکی آذری گونه تبریزی با اندازه‌گیری پارامترهای $(\%V)$ ، $(\Delta C(\ln))$ ، $(\Delta V(\ln))$ و $(nPVI-V)$ و سرعت تولید صورت پذیرفت. میانگین نتایج به‌دست‌آمده از بررسی این پارامترها برای هر یک از گونه‌ها در جدول ۱ آمده است:

جدول ۱: میانگین مقادیر به‌دست‌آمده از پارامترهای صوت شناختی

rate	nPVI-V	%V	$\Delta V(\ln)$	$\Delta C(\ln)$	نوع داده
۶/۱۵	۵۰/۹۲	۵۰/۹۸	۰/۵۴	۰/۵۹	فارسی-فارسی
۵/۰۴	۵۲/۲۵	۴۶/۱۰	۰/۵۵	۰/۵۶	فارسی-ترکی
۴/۹۶	۴۸/۸۳	۴۷/۱۳	۰/۴۸	۰/۵۷	ترکی-فارسی
۵/۲۵	۵۵/۸۹	۴۹/۹۱	۰/۵۴	۰/۵۵	ترکی-ترکی

* در تمامی جدول‌ها، نمودارها و شکل‌ها، در قسمت نوع داده در ترکیب (گونه ۱- گونه ۲) گونه نخست زبان مادری گویشوران و گونه دوم، گونه‌ای است که فرد هنگام ضبط داده به آن گونه صحبت کرده است.

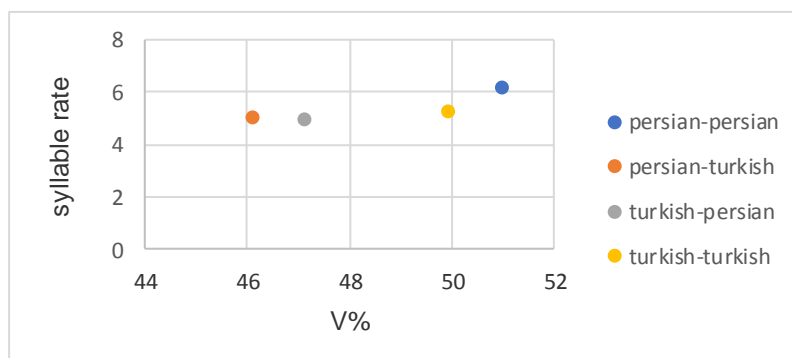
بر اساس اطلاعات موجود در این جدول، زبان فارسی دارای بالاترین میزان پارامتر $\Delta C(\ln)$ و $\%V$ است. کمترین میزان پارامتر $\Delta C(\ln)$ نیز متعلق به ترکی آذری گونه تبریزی است.



شکل ۳: میانگین پارامترهای $\Delta C(\ln)$ و $\%V$ در گونه‌های مورد بررسی در پژوهش

مقایسه مقادیر به‌دست‌آمده برای پارامترهای $\Delta C(\ln)$ و $\%V$ در گونه فارسی و ترکی آذری گونه تبریزی با مقادیر به‌دست‌آمده از همین پارامترها در سایر زبان‌ها نشان می‌دهد که نمی‌توان گونه زبانی فارسی را متعلق به طبقه هجا-زمان و یا تکیه-زمان در نظر گرفت. گونه زبانی فارسی به لحاظ ساخت هجا که از نظر صوت‌شناختی در پارامتر $\Delta C(\ln)$ متبلور می‌شود رفتاری شبیه به زبان‌های تکیه-زمان از خود نشان داده و به لحاظ پارامتر $\%V$ رفتاری مشابه زبان‌های هجا-زمان دارد. از اینرو، همسو با نظر نسپر (۱۹۹۰) مبنی بر قائل

شدن به طبقه‌ای ریتمیک با عنوان زبان‌های ترکیبی^۱، شاید بتوان فارسی محاوره‌ای را نیز به عنوان زبانی ترکیبی به لحاظ ریتمیک در نظر گرفت. بر اساس مقادیر به‌دست‌آمده از همین پارامترها ترکی آذری گونه تبری‌زی را می‌توان در دسته زبان‌های هجا-زمان قرارداد. میانگین مقادیر پارامترهای $\Delta C(\ln)$ و $V\%$ داده‌های نوع ۲ (ترکی - فارسی) و ۴ (فارسی - ترکی) را در گروه زبان‌های تکیه-زمان قرار می‌دهد. با این وجود، تفاوت میانگین پارامتر $\Delta C(\ln)$ ، در انواع داده‌ها، دارای تمایز معنی‌دار نیست. در مرحله دوم از بررسی داده‌های پژوهش، گفتار افرادی که سعی در تقلب گونه زبانی ترکی آذری داشته‌اند با گفتار فارسی‌زبانان و گویشوران ترکی که به زبان فارسی به عنوان زبان دوم خود مسلط بوده‌اند مورد مقایسه قرار گرفت. به منظور بررسی تأثیر استفاده از تقلب بر پارامترهای مختلف، از آزمون آماری تحلیل واریانس چندمتغیره^۲ استفاده شد. در این آزمون، نوع داده (گونه‌های زبانی مختلف) به عنوان عامل ثابت و پارامترهای $(V\%)$ ، $(\Delta C(\ln))$ ، $(\Delta V(\ln))$ ، $(nPVI-V)$ و سرعت تولید به عنوان متغیرهای وابسته در نظر گرفته شدند. نتایج این بررسی‌ها نشان داد که تمایز معنی‌داری میان پارامترهای صوت‌شناختی طبقه‌بندی ریتمیک گفتار بر اساس نوع داده وجود دارد ($F(15, 33/528) = 4/375$). بررسی مقادیر به‌دست‌آمده نشان داد که نوع داده تأثیر معناداری بر پارامترهای $(V\%)$ ($F(3, 16) = 16/88$)، $p < 0/05$ ، $\text{partial } \eta^2 = 0/76$ و سرعت تولید ($F(3, 16) = 8/144$)، $p < 0/05$ ، $\text{partial } \eta^2 = 0/604$ دارد. شکل ۴، تعامل میان پارامترهای $(V\%)$ و سرعت تولید را در انواع داده‌ها بازنمایی می‌کند.



1. rhythmically mixed
2. Manova

شکل ۴: میانگین پارامترهای سرعت تولید و $V\%$ در گونه‌های مورد بررسی در پژوهش

داده‌های موجود در این نمودار نمایانگر پایین‌تر بودن میزان پارامتر سرعت تولید در داده‌های نوع دوم (ترکی-فارسی) نسبت به داده‌های نوع اول (ترکی-ترکی) است. علت این امر را می‌توان تأثیر زبان مادری و زبان دوم در نظر گرفت. بر اساس نظر دیونگ و همکاران (۲۰۱۳) سرعت گفتار در زبان مادری بیشتر از زبان دوم است. داده‌های نوع ۴ (فارسی-ترکی) دارای سرعت تولید پایین‌تری نسبت به زبان مادری این گویشوران که فارسی است بوده، در عین حال، مقدار این پارامتر بالاتر از میانگین سرعت گویشوران ترک‌زبانی است که زبان فارسی، زبان دوم این افراد بوده است. به منظور بررسی تفاوت میان پارامترهای مورد سنجش در داده‌های مختلف، آزمون تعقیبی توکی^۱ اجرا شد. بر اساس نتایج موجود در جدول ۲، میانگین پارامتر $V\%$ در داده‌های فارسی-فارسی و فارسی-ترکی به لحاظ آماری دارای تفاوت معنادار است. تفاوت معنادار این پارامتر در این دو نوع داده به این معناست که بر اساس پارامتر $V\%$ گویشوران فارس‌زبانی که گونه ترکی آذری را تقلید کرده‌اند، در ایجاد تمایز میان زبان مادری خود و زبانی که سعی در تقلید آن داشته‌اند، موفق عمل کرده و گفتار خود را به گونه ترکی آذری نزدیک کرده‌اند. نتایج به‌دست‌آمده از این بررسی در تعارض با پژوهش‌هایی نظیر دلوو و همکاران (۲۰۰۹) است که پیش‌تر عنوان کرده بودند میزان واکه‌ای بودن گفتار که از طریق پارامتر $V\%$ بازنمایی می‌شود، فرد ویژه بوده و تقلید یک گویشور از گونه‌ای دیگر باعث ایجاد تغییر در این پارامتر نمی‌گردد. داده‌های موجود در جدول ۲ همچنین نشان می‌دهد که میزان عملکرد موفق گویشوران فارسی به حدی بوده است که نمی‌توان بر اساس پارامتر $V\%$ تمایز معناداری میان داده‌های فارسی-ترکی و ترکی-فارسی قائل شد. از اینرو، بر اساس نتایج به‌دست‌آمده از این پژوهش، پارامتر $V\%$ ، پارامتر مناسبی برای ایجاد تمایز میان داده‌های فارسی-ترکی و ترکی-فارسی نیست. داده‌های این پژوهش، همچنین مؤید تفاوت معنی‌دار میان داده‌های فارسی-فارسی و ترکی-فارسی هستند. به همین دلیل، می‌توان پارامتر $V\%$ را به عنوان پارامتری مناسب در ایجاد تمایز میان گویشورانی که فارسی زبان مادری آن افراد بوده و گویشورانی که فارسی زبان دوم آن‌ها بوده، در نظر گرفت. نتیجه مشابه در مورد گویشورانی که ترکی آذری زبان مادری آن افراد بوده و افرادی که سعی در تقلید ترکی داشته‌اند نیز صادق است؛ بنابراین، به نظر می‌رسد که

بتوان از پارامتر (%V) به عنوان پارامتری موفق در ایجاد تمایز میان زبان مادری و زبان دوم/ غیر مادری افراد نام برد. واکاوی نتایج آزمون تعقیبی توکی در مورد پارامتر سرعت تولید نشان می‌دهد که تفاوت میان میانگین‌های این پارامتر در داده‌های فارسی-فارسی و فارسی-ترکی، ترکی-فارسی و ترکی-ترکی معنادار است. به این ترتیب، این پارامتر در ایجاد تمایز میان گویشوران فارسی که به فارسی صحبت می‌کرده‌اند با گویشوران ترک‌زبان که به ترکی و یا فارسی صحبت می‌کرده‌اند، موفق عمل می‌کند.

جدول ۲: آزمون تعقیبی توکی برای پارامترهای (%V) و سرعت تولید

متغیر وابسته	نوع داده	نوع داده	متغیر وابسته	نوع داده	نوع داده	متغیر وابسته
P-value			P-value			P-value
۰/۰۰۴	فارسی-ترکی*	فارسی-فارسی	۰/۰۰۰	فارسی-ترکی*	فارسی-فارسی	% V
۰/۰۰۲	ترکی-فارسی*		۰/۰۰۱	ترکی-فارسی*		
۰/۰۲۰	ترکی-ترکی*		۰/۵۴۸	ترکی-ترکی		
۰/۰۰۴	فارسی-فارسی*	فارسی-ترکی	۰/۰۰۰	فارسی-فارسی*	فارسی-ترکی	
۰/۹۸۹	ترکی-فارسی		۰/۵۶۶	ترکی-فارسی		
۰/۸۷۱	ترکی-ترکی		۰/۰۰۱	ترکی-ترکی*		
۰/۰۰۲	فارسی-فارسی*	ترکی-فارسی	۰/۰۰۱	فارسی-فارسی*	ترکی-فارسی	
۰/۹۸۹	فارسی-ترکی		۰/۵۶۶	فارسی-ترکی		
۰/۷۱۰	ترکی-ترکی		۰/۰۱۴	ترکی-ترکی*		
۰/۰۲۰	فارسی-فارسی*	ترکی-ترکی	۰/۵۴۸	فارسی-فارسی	ترکی-ترکی	
۰/۸۷۱	فارسی-ترکی		۰/۰۰۱	فارسی-ترکی*		
۰/۷۱۰	ترکی-فارسی		۰/۰۱۴	ترکی-فارسی*		

* علامت ستاره در جدول نشانه معنادار بودن تمایز میان میانگین پارامترها در انواع داده است

همانند پارامتر (%V)، پارامتر سرعت تولید در گویشوران فارسی‌زبانی که گونه ترکی را تقلید کرده‌اند، دارای تمایز معنادار با گفتار همین افراد در هنگامی است که به گونه فارسی صحبت می‌کنند؛ به همین دلیل، نمی‌توان این پارامتر را نیز، پارامتری فرد ویژه در نظر گرفت که تحت تأثیر تقلب دستخوش تغییر نمی‌شود.

همچنین، پارامتر سرعت تولید، دارای تمایز معناداری میان داده‌های فارسی-ترکی و ترکی-فارسی نیست. اگرچه این پارامتر، نظیر پارامتر (%V) قابلیت ایجاد تمایز میان داده‌های نوع ۲ (ترکی-فارسی) و نوع ۴ (فارسی-ترکی) را ندارد، اما در ایجاد تمایز معنی‌دار میان زبان مادری و زبان دوم/ غیر مادری موفق عمل کرده است.

۴- نتیجه

در پژوهش حاضر، به بررسی تأثیر تقلب فارسی‌زبانان در استفاده از گونه ترکی آذری بر پارامترهای صوت‌شناختی دیرش-بنیاد طبقه‌بندی ریتمیک گفتار، در چارچوب آواشناسی قضایی پرداختیم. ابتدا و با استفاده از پارامترهای میزان واکه‌ای بودن گفتار (٪V) و انحراف معیار دیرش مناطق همخوانی نرمال‌سازی شده ($\Delta C(\ln)$) که در پیشینه پژوهش به عنوان پارامترهای مناسب در طبقه‌بندی ریتمیک زبان‌های مختلف معرفی شده‌اند، طبقه‌بندی ریتمیک فارسی محاوره‌ای و ترکی آذری گونه تبریزی و داده‌های نوع ۲ (ترکی-فارسی) و ۴ (فارسی-ترکی) انجام شد. نتایج این پژوهش نشان داد که نمی‌توان پارامتر انحراف معیار دیرش مناطق همخوانی نرمال‌سازی شده ($\Delta C(\ln)$) را به عنوان پارامتری مناسب در ایجاد تمایز میان داده‌های مختلف این پژوهش در نظر گرفت. بر اساس نتایج کسب‌شده از سنجش متغیرهای میزان واکه‌ای بودن گفتار (٪V)، انحراف معیار دیرش مناطق همخوانی نرمال‌سازی شده ($\Delta C(\ln)$)، انحراف معیار دیرش مناطق واکه‌ای نرمال‌سازی شده ($\Delta V(\ln)$)، شاخص تغییر دو به دوی واکه‌ای نرمال‌سازی شده (nPVI-V) و سرعت تولید، تنها میانگین دو پارامتر میزان واکه‌ای بودن گفتار (٪V) و سرعت تولید، دارای تمایز معنی‌دار، در انواع داده‌ها هستند. با این وجود، هیچ‌کدام از پارامترهای صوت‌شناختی ریتم گفتار توانایی ایجاد تمایز میان داده‌های ترکی-فارسی و فارسی-ترکی را نداشتند. اگرچه پارامترهای میزان واکه‌ای بودن گفتار (٪V) و سرعت تولید، در ایجاد تمایز میان داده‌های فارسی-ترکی و ترکی-فارسی موفق عمل نکردند، اما می‌توان از این دو پارامتر به عنوان پارامترهای مناسب در ایجاد تمایز میان زبان مادری و زبان دوم/ غیر مادری یاد کرد. به طور کلی، می‌توان اذعان نمود که پارامترهای ریتم گفتار، پارامترهای مناسبی در تشخیص تقلب در گفتار نیستند.

پی‌نوشت

۱. یکی از نکاتی که باید به آن اشاره کرد، جایگزینی واژه مقایسه قضایی گفتار به جای واژه تشخیص هویت گوینده در بسیاری از منابع است که پس از انتشار مقاله سکس (Michael J. Saks) و کولر (Johnathan J. Koehler) (۲۰۰۵) در خصوص لزوم بازنگری و ایجاد تغییرات در علوم قضایی مرتبط با تشخیص هویت، به وجود آمده است. جهت کسب اطلاعات بیشتر در این خصوص رجوع کنید به مقاله (The coming paradigm shift in the forensic identification science).
۲. میرهاف (Miriam Meyerhoff) (۲۰۱۱: ۷۴) «گونه زبانی را اصطلاحی خنثی برای ارجاع به زبان، گویش و لهجه بدون در نظر گرفتن تمایز میان زبان، گویش و لهجه» تعریف کرده است. در این پژوهش نیز، گونه زبانی، با در نظر گرفتن تعریف میرهاف (۲۰۱۱) بکار برده خواهد شد.

۳. به عقیده لباو (۲۰۰۶: ۷۱-۶۵) در مواردی که از افراد درخواست می‌گردد که خاطره تلخی را بیان نمایند، به دلیل اینکه احساسات افراد در بیان آن خاطره و یا اتفاق دخیل می‌گردد، امکان دستیابی به گفتار در طبیعی‌ترین حالت فراهم می‌شود.

منابع

- راه‌انداز، سعید، آسیائی، مارال و نقش‌بندی، شهرام (۱۳۹۳). خوشه همخوانی و ساخت هجا در زبان ترکی آذری، در دبیرمقدم، محمد، مجموعه مقالات نهمین همایش زبان شناسی ایران (۵۳-۵۶۴). تهران، دانشگاه علامه طباطبائی
- صادقی، وحید (۱۳۹۴). بررسی آوایی کاهش واکه‌ای در زبان فارسی، جستارهای زبانی، دوره شش، شماره ۳، ۱۶۵-۱۷۸.
- Amino, K. and T. Arai. 2009. Speaker-Dependent Characteristics of the Nasals, *Forensic Science International* 185(1): 21-28. <http://www.sciencedirect.com/science/article/pii/S0379073808004672>.
- Asadi, H., M. Nourbakhsh, F. Sasani, and V. Dellwo. 2018. Examining Long-Term Formant Frequency as a Forensic Cue for Speaker Identification: An Experiment on Persian. In M. Nourbakhsh, H. Asadi, and M. Asiaee (eds), *Proceedings of the First International Conference on Laboratory Phonetics and Phonology (21-28)*,. Tehran: Neveesh Parsi Publications.
- Asiaee, M., M. Nourbakhsh, and R. Skarnitzl. 2019. Can LTF Discriminate Bilingual Speakers? In *Proceedings of the 28th Annual Conference of the International Association for Forensic Phonetics and Acoustics (IAFPA)*, Istanbul, 41-42.
- Barry, W. J., B. Andreeva, M. Russo, S. Dimitrova and T. Kostadinova. 2003. Do Rhythm Measures Tell Us Anything about Language Type? In D. Recasens, M. J. Solé, and J. Romero (eds), *Proceedings of the 15th International Congress of Phonetic Sciences (15th ICPHS) (2693-96)*, Barcelona.
- Boersma, P. and Weenink, D. 2020 Praat: doing phonetics by computer. <http://www.praat.org>, Accessed 02 March 2020.
- Dauer, R. M. 1983. Stress-Timing and Syllable-Timing Reanalyzed, *Journal of Phonetics* 11(1): 51-62. [https://doi.org/10.1016/S0095-4470\(19\)30776-4](https://doi.org/10.1016/S0095-4470(19)30776-4).
- De Jong, N. H., R. Groenhout, R. Schoonen, and J. H. Hulstijn. 2015. Second Language Fluency: Speaking Style or Proficiency? Correcting Measures of Second Language Fluency for First Language Behavior, *Applied Psycholinguistics* 36(2): 223-43.
- Dellwo, V., S. Ramyeard and J. Dankovicova. 2009. The influence of voice disguise on temporal characteristics of speech. *Abstract presented at the IAFPA conference*, Cambridge: UK
- Dellwo, V., A. Leemann, and M. J. Kolly. 2015. Rhythmic Variability between Speakers: Articulatory, Prosodic, and Linguistic Factors, *The Journal of the Acoustical Society of America* 137(3): 1513-28.

- Dellwo, V. 2010. Influences of Speech Rate on the Acoustic Correlates of Speech Rhythm: An Experimental Phonetic Study Based on Acoustic and Perceptual Evidence. Doctoral Dissertation, University of Bonn (Rheinische Friedrich-Wilhelms-Universität Bonn).
- Eriksson, A. 2010. The Disguised Voice: Imitating Accents and Speech Styles and Impersonating Individuals, *Language and identities* (January 2010): 86–96.
- Gold, E., P. French, and P. Harrison. 2013. Examining Long-Term Formant Distributions as a Discriminant in Forensic Speaker Comparisons under a Likelihood Ratio Framework, *Proceedings of Meetings on Acoustics* 19(May).
- Grabe, E., and E. L. Low. 2002. Durational Variability in Speech and the Rhythm Class Hypothesis. In C Gussenhoven and A Warner (eds), *Laboratory Phonology 7* (515–46), Berlin: Mouton de Gruyter.
- Hollien, H. 2002. *Forensic Voice Identification*. San Diego: Academic Press.
- IBM Corp. 2012. IBM SPSS Statistics for Windows (version 22.0). Armonk, NY: International Business Machines Corporation.
- Jessen, M. 2008. Forensic Phonetics, *Language and linguistics compass*, 2(4): 671–711.
- Kinoshita, Y. 2005. Does Lindley's LR Estimation Formula Work for Speech Data? Investigation Using Long-Term F0, *Forensic Linguistics*, 12(2): 235–54.
- Künzel, H. J. 2000. Effects of Voice Disguise on Speaking Fundamental Frequency, *Forensic Linguistics*, 7 (2): 149–179. <https://www2.scopus.com/inward/record.uri?eid=2-s2.0-54249140687&partnerID=40&md5=91a9ecd533c278f5e6fc8f1d80299550>.
- Labov, W. 2006. *The Social Stratification of English in New York City*, 2nd ed, Cambridge: Cambridge University Press.
- Lazard, G. 1992. *Grammar of Contemporary Persian*. Costa Mesa, CA: Mazda Publishers.
- Leemann, A. and M. J. Kolly. 2015. Speaker-Invariant Suprasegmental Temporal Features in Normal and Disguised Speech. *Speech Communication*, 75:97–122. <https://www2.scopus.com/inward/record.uri?eid=2-s2.0-84946575988&doi=10.1016%2Fj.specom.2015.10.002&partnerID=40&md5=dfbba8701a51254a6dc359b266ce994e>.
- Low, E. L., E. Grabe, and F. Nolan. 2000. Quantitative Characterizations of Speech Rhythm: Syllable-Timing in Singapore English, *Language and Speech*, 43(4): 377–401.
- James, A. 1940, *Speech Signals in Telephony*, London: Pitman
- Mairano, P., and A. Romano. 2011. Rhythm Metrics for 21 Languages, *ICPhS XVII* (August): 1318–21.
- Masthoff, H. 1996. A Report on a Voice Disguise Experiment, *Forensic Linguistics* 3(1): 160–67. <https://journals.equinoxpub.com/index.php/IJSL/article/view/17245>.

- Meyerhoff, M. 2011. *Introducing Sociolinguistics*, New York, NY: Routledge.
- Ghaffarvand Mokari, P. and S. Werner. 2017. Illustrations of the IPA: Azerbaijani, *Journal of the International Phonetic Association*, 47(2): 207-12.
- Nespor, M. 1990. On the rhythm parameter in phonology. In I. Roca (eds.), *the Logical Problem of Language Acquisition (157-175)*, Foris. Dordrecht.
- Nolan, F., and C. Grigoras. 2005. A Case for Formant Analysis in Forensic Speaker Identification, *International Journal of Speech, Language and the Law*, 12(2): 143-73.
- Nolan, F. 1983. *The Phonetic Bases of Speaker Recognition*, New York: Cambridge University Press.
- Perrot, P. and Chollet, G. 2012. Helping the Forensic Research Institute of the French *Gendarmerie* to Identify a Suspect in the Presence of Voice Disguise or Voice Forgery, In A. Neustein and H. A. Patil (eds.), *Forensic Speaker Recognition: law enforcement and counterterrorism (pp.469-503)*. New York: Springer. https://www2.scopus.com/inward/record.uri?eid=2-s2.0-84943230355&doi=10.1007%2F9781461402633_16&partnerID=40&md5=026de43ee05d64da537d50753b60e535.
- Perrot, P., G. Aversano, and G. Chollet. 2007. Voice Disguise and Automatic Detection: Review and Perspectives. In Y. Stylianou, M. Faundez-Zanuy and A. Esposito (eds), *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)* 4391 LNCS(101-17), https://doi.org/10.1007/978-3-540-71505-4_7.
- Rahandaz, S., M. Asiaee & Sh. Naghshbandi. 2015. Consonant Cluster and Syllable Structure in Azerbaijani." In M. Dabirmoghaddam (ed.), *9th Iranian Conference on Linguistics (553-564)*, Tehran: Allameh Tabataba'ei University [In Persian].
- Ramus, F., M. Nespor, and J. Mehler. 1999. Correlates of Linguistic Rhythm in the Speech Signal, *Cognition*, 75(1): 265-92.
- Rathcke, T. V. and R. H. Smith. 2015. Speech Timing and Linguistic Rhythm: On the Acoustic Bases of Rhythm Typologies, *The Journal of the Acoustical Society of America*, 137(5): 2834-45.
- Rodman, R. D. 1998. Speaker Recognition of Disguised Voices: A Program for Research. In A. Demirekler, M. Saranlı, A. Altıncay and H. Paoloni (eds.), *Proceedings of the Consortium on Speech Technology Conference on Speaker Recognition by Man and Machine: Directions for Forensic Applications (9-22)*. <http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.116.7877>.
- Sadeghi, V. (2015). A phonetic study of vowel reduction in Persian, *Language Related Research*, 30: 165-187. [In Persian]

- Saks, M. J. and J. J. Koehler. 2005. The Coming Paradigm Shift in Forensic Identification Science, *Science* 309(5736): 892–95.
- Shuy, R. W. 1995. Dialect as Evidence in Law Cases, *Journal of English Linguistics*, 23(1/2): 195–208.
- Skarnitzl, R., and J. Vaňková. 2017. Fundamental Frequency Statistics for Male Speakers of Common Czech, *Acta Universitatis Carolinae – Philologica 3, Phonetica Pragensia XIV*: 7–17.
- Taghva, N. and V. Abolhasani Zadeh. 2016. Comparison of English Language Rhythm and Kalhori Kurdish Language Rhythm, *Advances in Language and Literary Studies*, 7(2): 226–30.
- White, L. and S. L. Mattys. 2007. Calibrating Rhythm: First Language and Second Language Studies, *Journal of Phonetics*, 35(4): 501–22.
- Windfuhr, G. L. 1979. *Persian Grammar: History and State of its Study*, New York: Mouton de Gruyter. <https://doi.org/10.1515/9783110800425>
- Wolf, Jared J. 1972. Efficient Acoustic Parameters for Speaker Recognition, *The Journal of the Acoustical Society of America* 51(6B): 2044–56. <https://doi.org/10.1121/1.1913065>.

