



اطلس گویشی چرداول: بررسی تنوعات آوایی و اندازه‌گیری فاصله گویشی

محسن حیدری زادی^{۱*}

گلناز مدرسی قوامی^۲

چکیده

گویش سنجی رویکردی آمارگرا در مطالعات زبانی است که در آن تنوعات زبانی در یک پهنه‌ی جغرافیایی بر پایه‌ی روش‌ها و اصولی مشخص سنجیده و سپس در قالب نمودارها و شکل‌های تفسیری ارائه می‌شوند و قابلیت تعیین نسبی مرکز گویش‌ها و زبان‌ها را فراهم می‌آورد. این پژوهش با بررسی حوضه آبریز رود چرداول در استان ایلام که در آن گونه‌هایی از زبان‌های کردی، لکی و لری رایج هستند با استفاده از شیوه‌های گویش‌سنجی و ارائه روش‌شناسی انجام کار کوشیده است تصویری مناسب از چگونگی توزیع گونه‌های زبانی در منطقه ارائه نماید و علاوه بر ارائه‌ی تصویری کلی از همبستگی زبانی - جغرافیایی منطقه، برای هر مکان خاص فواصل گویشی از سایر مکان‌ها را بسنجد. برای این منظور، داده‌های زبانی ۳۵ آبادی بر اساس فهرست واژگان پایه مورس سوادش در این قلمرو گردآوری و آوانگاری گردیده است و پس از لحاظ کردن داده‌های جغرافیایی منطقه با استفاده از نرم‌افزار Gabmap تحلیل‌های مربوطه به‌دست آمده است. ارائه دندوگرام، نمودار مقیاس‌گذاری چند بُعدی و نیز نقشه پرتوی از منطقه مورد بررسی نمونه‌هایی از این تحلیل‌ها هستند. که در آنها تناسب میان تنوعات زبانی - جغرافیایی به‌گونه‌ای منعکس شده است که مرزهای نسبی زبانی در کردی، لکی و لری قابل تشخیص است.

کلیدواژه‌ها: گویش‌سنجی، فاصله‌ی گویشی، تنوعات آوایی - واژگانی، چرداول.

۱- دانشجوی دکتری زبان‌شناسی همگانی؛ دانشگاه علامه طباطبائی، دانشکده ادبیات | zanjire@gmail.com (نویسنده مسئول).

۲- دانشیار رشته زبان‌شناسی همگانی، دانشگاه علامه طباطبائی، دانشکده ادبیات فارسی | gmodarresi@yahoo.com و زبان‌های خارجی.

۱- مقدمه

اهمیت ثبت و سپس دسته‌بندی داده‌های زبانی در یک گستره‌ی جغرافیایی مشخص، به‌طوری‌که بتوان بر اساس آن به تبیین‌های زبان‌شناختی دست یافت، بر کمتر کسی پوشیده است. بررسی تفاوت‌های ساختاری، معنایی یا تلفظ کلمات از ناحیه‌ای به ناحیه دیگر بر روی نقشه‌های جغرافیایی به ترسیم اطلس زبانی^۱ منتهی می‌شود. اطلس زبانی مجموعه‌ای از نقشه‌های جغرافیایی است که توزیع مشخصه‌های زبانی را در یک محدوده جغرافیایی مشخص به نمایش می‌گذارد (هارتمن و جیمز^۲، ۲۰۰۲: ۸۸). مطالعه‌ی نظام‌مند گونه‌های زبانی پراکنده در پهنه‌های جغرافیایی که با عنوان جغرافیای زبانی^۳ یا جغرافیای گویشی^۴ شناخته می‌شود، مورد علاقه و دغدغه بسیاری از زبان‌شناسان بوده است؛ زیرا پاسخ به بسیاری از پرسش‌های مطرح در علم زبان‌شناسی را می‌توان از لابه‌لای این مطالعات دریافت. توصیف درست گونه‌های پراکنده در یک محدوده جغرافیایی زمانی میسر است که در طول مطالعات گسترده انجام شده در آن منطقه، همانندی‌ها از ناهمانندی‌ها تمیز داده شوند. گستره‌ی این گونه مطالعات می‌تواند دربرگیرنده‌ی واج‌شناسی، صرف، نحو، واژگان و مانند آن باشد. اطلس‌های زبانی گوناگونی تاکنون در جای‌جای جهان تدوین شده‌اند، که هم در شناخت اصول زبانی حاکم بر یک زبان خاص و هم در تدوین اطلس جهانی زبان‌ها^۵ نقش‌آفرینی می‌کنند. تدوین اطلس‌های زبانی به‌ویژه برای زبان‌های در معرض خطر^۶، می‌تواند نقش مهمی در حفظ میراث زبانی در جوامع گوناگون ایفا کند؛ بر همین اساس تلاش‌های زیادی در جهت ثبت و ضبط داده‌های زبان‌های گوناگون صورت پذیرفته است.

امروزه به کمک فناوری‌های ثبت و تحلیل انبوه داده‌ها تهیه‌ی اطلس‌ها، به شیوه رایانه‌ای صورت می‌پذیرد و به سبب امکان دسته‌بندی و خوشه‌بندی داده‌ها با سهولت بیشتر، قابلیت سنجش کمی این داده‌ها نیز به‌صورت انبوه فراهم است. اندازه‌گیری آن دسته از تفاوت‌های گویشی که عموماً توسط جغرافیا رقم می‌خورد، ذیل عنوان گویش‌سنجی^۷ قرار می‌گیرد؛ اگرچه برخی ممکن است گویش‌سنجی را در گستره‌ای وسیع‌تر بگنجانند، آنجا که مطالعات تغییرات و تنوعات زبانی علاوه بر جغرافیا، متغیرهای دیگری چون عوامل اجتماعی، جنسیت، سن و مانند آن را در برمی‌گیرد. انتظار می‌رود روش‌های گویش‌سنجی، در آینده برای سنجش تفاوت‌های زبانی بر اساس این عوامل به‌کار گرفته شوند. آنچه در گویش‌سنجی نوین در کانون توجه قرار دارد ارائه نتایج آماری با استفاده از تحلیل انواع داده‌هاست؛ این شیوه‌ی کار از اعمال سلیقه در مرزبندی

-
1. linguistic atlas
 2. Hartmann, R.R.K. and James, G.
 3. linguistic geography
 4. dialect geography
 5. World Atlas of Languages
 6. endangered languages
 7. dialectometry

گوشش‌ها که یکی از معایب اصلی گویش‌شناسی سنتی است، جلوگیری می‌کند. دقت و سرعت و نیز توانایی انجام کار بر روی داده‌های بیشتر از مزایای دیگر گویش‌سنجی نوین محسوب می‌شود. هدف این پژوهش بررسی تنوعات آوایی در منطقه چرداول (حوزه آبریز رود چرداول) استان ایلام با استفاده از شیوه‌های گویش‌سنجی نوین و بر اساس فهرست واژگان مفاهیم پایه مورس سوادش^۱ است. قرار گرفتن در حد فاصل میان استان‌های، کرمانشاه، لرستان، و ایلام که در آنها به گونه‌هایی از کردی، لکی و لری تکلم می‌شود، باعث شده است چرداول، منطقه‌ای با تنوعات زبانی قابل ملاحظه باشد که در آن اثرگذاری متقابل گونه‌ها بر همدیگر مشهود است.

یکی از نخستین تلاش‌ها در جهت تدوین اطلس‌های زبانی مربوط به کوراث^۲ (۱۹۴۹) است؛ وی نخستین اطلس زبانی انگلیسی ایالات متحده آمریکا را ترسیم نموده است. چمبرز و ترادگیل^۳ (۲۰۰۴) چارچوبی را برای بررسی‌های جامعه‌شناختی زبان و مسائل پیرامون آن ارائه نموده‌اند و به بررسی متغیرهای مؤثر در گوناگونی‌های زبانی پرداخته‌اند. از جمله‌ی این متغیرها، تأثیر جغرافیا بر گویش است، که از آن با عنوان جغرافیای گویشی نام برده می‌شود. تهیه‌ی اطلس زبانی اسپانیولی در نیومکزیکو و کلرادو جنوبی توسط بیلز و ویجیل^۴ (۲۰۰۸) صورت پذیرفته است. در این اطلس زبانی که بر پایه‌ی تنوعات آوایی و واژگانی بنا شده است، انواع نقشه‌های زبانی نشان‌دهنده‌ی مرزهای همگویی^۵ ارائه شده است. همسو با تلاش‌های فراوان در پهنه‌ی جهان، در داخل کشور نیز تلاش‌هایی در جهت تدوین اطلس‌های زبانی گونه‌های رایج، صورت گرفته است که از جمله‌ی آنها می‌توان به جهانگیری (۱۳۵۲)، قلی فامیان (۱۳۸۹)، الیاسی (۱۳۸۹)، مولایی (۱۳۹۰)، بصیرت (۱۳۹۱)، ملایی پاشایی (۱۳۹۳)، صادقی (۱۳۹۴) و سنایی (۱۳۹۵) اشاره کرد. علاوه بر این، مجموعه طرح ملی گویش‌شناسی مشتمل بر چند طرح پژوهشی و پایان‌نامه دانشجویی است که در چند استان از جمله فارس، هرمزگان، کرمانشاه، ایلام، لرستان، کردستان، یزد، همدان، گیلان، مازندران، اصفهان، کرمان و نیشابور تدوین گردیده است. در این طرح عموماً از هر شهرستان چند روستای محدود به‌عنوان نمونه انتخاب شده‌اند. اطلس‌های زبانی متعددی نیز در میان پایان‌نامه‌های دانشجویی دانشگاه‌های شرق کشور به‌ویژه دانشگاه سیستان و بلوچستان و دانشگاه فردوسی مشهد به چشم می‌خورد که می‌توان از آن میان به اطلس زبانی بیرجند، سریشه، نهبندان، طیس، سیرجان، زابل، پشت‌آب و... اشاره نمود.

اگرچه به‌طور کلی بر روی گونه‌های زبانی استان ایلام پژوهش‌هایی صورت گرفته است، اما هدف پژوهش حاضر ارائه جزئیات فنی‌تر و دقیق‌تر از قلمروی محدودتر با استفاده از ابزارهای گویش‌سنجی و معرفی کارکرد آنها است تا جزئیات زبانی بیشتری را بتوان نشان داد. آنچه در پژوهش‌های پیشین منطقه‌ای،

1. Swadesh, M.

2. Kurath, H.

3. Chambers, J. K. & P. Trudgill

4. Bills, Garland D. and Neddy A. Wigil

5. Isoglosses

یکدست نمایش داده شده است خود دارای تنوعات زبانی قابل توجه است؛ این امر به درک بهتر وضعیت گونه‌های زبانی موجود در منطقه کمک می‌کند.

۱-۱- گویش سنجی

گویش سنجی به معنای اندازه‌گیری میزان تفاوت‌ها و شباهت‌های میان گویش‌ها (گونه‌ها) است و به تبع آن می‌توان الگوهای گویشی منطقه مورد مطالعه را آشکار کرد. از گویش سنجی به منظور ثبت سابقه گویشی یک منطقه، آشکار ساختن اثرات زبان‌شناختی ناشی از مهاجرت، برخوردهای زبانی و یا تنوعات درون‌زبانی می‌توان بهره گرفت. برنامه‌ریزان زبان، زبان‌آموزان، خبرنگاران و طیف وسیعی از افراد مخاطب این شاخه از زبان‌شناسی هستند.

سگوی^۱ را می‌توان آغازگر مطالعات زبانی مبتنی بر گویش سنجی، دانست. او شیوه‌ای را در دهه ۱۹۷۰ میلادی بنا نهاد که به کمک آن اندازه‌گیری فواصل زبانی میان گویش‌هایی که در مکان‌های مختلف بدان‌ها تکلم می‌شود فراهم می‌شد. در شیوه‌ی سگوی، با مقایسه‌ی صورت‌های ثبت‌شده در دو مکان مجاور و شمارش تفاوت‌های آنها جدولی تهیه می‌شد که در آن مقادیر عددی تفاوت زبانی هر مکان با مکان مجاورش ثبت می‌گردید. از این جدول برای تعیین مرزهای گویشی^۲ استفاده می‌شد، به‌گونه‌ای که بیشترین مقادیر تفاوت زبانی محاسبه شده به‌عنوان مرز گویشی تلقی و بر روی نقشه ترسیم و نواحی گویشی^۳ تعیین می‌شدند. این مرزهای گویشی جایگزین دسته‌های همگویی^۴ در گویش‌شناسی سنتی شدند. هرچند بعدها سگوی نیز تعیین مرزهای گویشی را غیردقیق دانست و رسیدن به مرز دقیق گویش‌ها را ناممکن (فرانسیس^۵، ۱۹۸۳: ۱۵۸).

هانس گوبل^۶ را باید نخستین کسی دانست که از رایانه در پردازش داده‌های گویشی بهره برد و به تبع آن این امکان برای وی وجود داشت که مقادیر بیشتری از داده‌ها را پردازش نماید. او نگاه‌ها را از تفاوت‌های زبانی به شباهت‌های زبانی معطوف کرد و محاسبات گویش‌سنجی را که سگوی فقط در مورد مکان‌های مجاور به کار می‌برد به‌گونه‌ای بسط داد که تشابهات زبانی هر مکان با تمامی مکان‌های دیگر، چه مجاور و چه غیرمجاور سنجیده می‌شد و حاصل آن جدولی متقارن بود که به اندازه تعداد مکان‌ها، طول و عرض داشت. معرفی تحلیل خوشه‌ای^۷ به‌عنوان مهم‌ترین موفقیت گوبل در گویش‌سنجی شناخته می‌شود که ابزاری

-
1. Seguy, Jean
 2. dialect boundaries
 3. dialect areas
 4. bundles of isogloss
 5. Francis, W. N.
 6. Goebel, Hans
 7. Cluster Analysis

مهم در طبقه‌بندی عددی^۱ محسوب می‌شود. از تحلیل خوشه‌ای برای دسته‌بندی مکان‌های دارای ویژگی‌های زبانی مشابه استفاده می‌شود.

نربون و هرینگا^۲، با معرفی معیارهای مشخص امکان ورود آواشناسی را به عرصه گویش‌سنجی فراهم کردند و از این رو امکان سنجش فواصل مدرج زبانی را فراهم آوردند. روش‌های قبلی همگی مبتنی بر سنجش صفر و یک بودند و تنها وقتی خاصیت مدرج پیدا می‌کردند که به صورت انبوه تحلیل می‌شدند. آنها همچنین توانستند با معرفی نمودار مقیاس‌گذاری چند بُعدی^۳ به‌عنوان ابزاری در گویش‌سنجی مسئله نمایش چند بُعدی متغیرهای زبانی را در فضای دو بُعدی محور مختصات X و Y فراهم آورند که در صفحات نمایش رایانه و برای نوشتارهای کاغذی ملموس‌ترند (نربون و هرینگا، ۲۰۰۱: ۳۹۹).

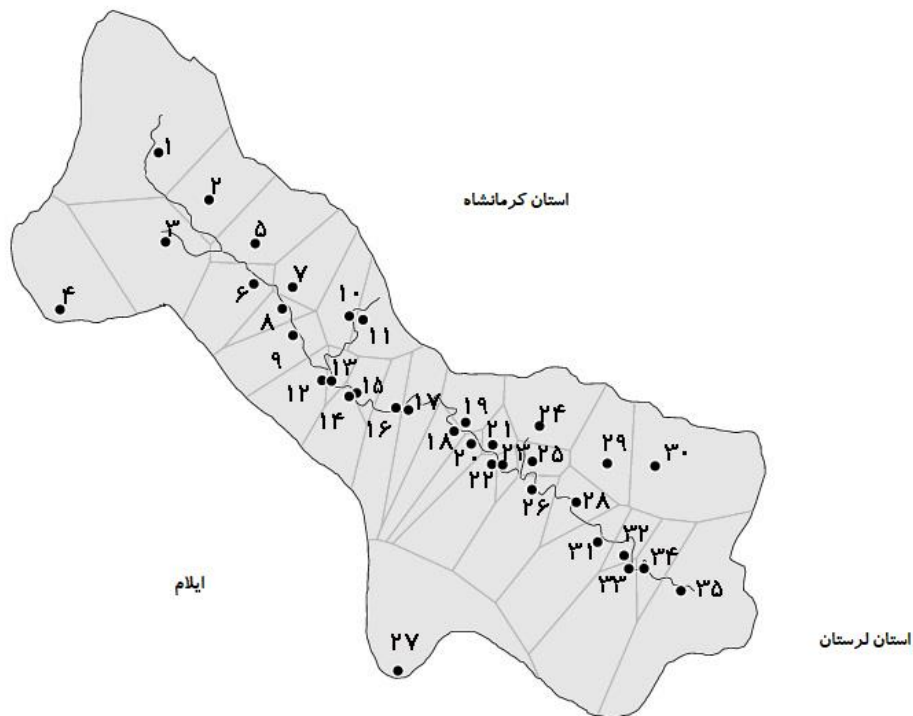
۲- روش تحقیق

قلمرو پژوهش حاضر منطقه جغرافیایی چرداول (حوضه آبریز رود چرداول) در استان ایلام است و در مجموع تعداد ۳۵ آبادی در این قلمرو مورد بررسی قرار گرفته‌اند. شیوه انتخاب روستاها به‌صورت نقطه به نقطه بوده است، به‌طوری‌که تقریباً تمامی شهرها و روستاها را در این حوضه شامل می‌شود. از منظر زبان‌شناختی، این منطقه، دربرگیرنده طیفی از گونه‌های زبانی است که کردی، لکی و لری را شامل می‌شود. این مطالعه از روستای زنجیره علیا که یکی از سرچشمه‌های رود چرداول است آغاز و در امتداد مسیر رودخانه به روستای چم‌چرود رسیده است. علاوه بر این آبادی‌ها، داده‌های روستاهای حسن گاوداری (شماره ۲۷) و چشمه خزانه (شماره ۴) که از ویژگی‌های متمایز زبانی در منطقه برخوردارند، نیز لحاظ گردیده‌اند. ناگفته پیداست که از ویژگی‌های خوب اطلس‌های رایانه‌ای خاصیت بسط‌پذیری آنها است که به راحتی می‌توان بر تعداد سایت‌های آنها در هر زمان افزود.

جدول ۱: سایت‌های بررسی شده برای تهیه اطلس رایانه‌ای منطقه چرداول استان ایلام

۱. زنجیره علیا	۲. زنجیره سفلی	۳. سرابله	۴. چشمه خزانه	۵. باخله	۶. موشکان	۷. شباب
۸. چوب شیرعلی	۹. چوب سرخ	۱۰. لتاب	۱۱. سنگ سفید	۱۲. دراشکفت	۱۳. چم کبود	۱۴. چم قوله
۱۵. قریب	۱۶. بردبل	۱۷. هلسم	۱۸. قاسم آباد	۱۹. طاق گاورین	۲۰. سعدآباد	۲۱. کله قطار علیا
۲۲. علیمرادخانی علیا	۲۳. کله قطار سفلی	۲۴. شورابه	۲۵. چغاسبز	۲۶. علیمرادخانی سفلی	۲۷. حسن گاوداری	۲۸. چوب بور
۲۹. بلاوه تره	۳۰. بلاوه خشکه	۳۱. نورآباد	۳۲. چم پوش	۳۳. کلنگ بر	۳۴. چوب شله	۳۵. چم چرود

1. Numerical Taxonomy
2. Nerbonne, John. and W. Heeringa
3. Multidimensional Scaling Plot



شکل ۱: نقشه منطقه جغرافیایی چرداول در استان ایلام (پلیگون^۱ مستخرج از گوگل ارث^۲)

۲-۱- شیوه گردآوری داده‌ها

در این پژوهش، از شیوه مصاحبه برای گردآوری داده‌های آوایی استفاده شده است. برای این منظور فهرست واژه‌های سوادش که شامل صد واژه است مبنای کار قرار گرفته است؛ این فهرست طیفی از واژه‌های پایه را که در تمامی زبان‌های بشری برای آنها معادل یافت می‌شود در بر می‌گیرد که شامل مفاهیمی چون اسامی حیوانات، اشیاء، ضمائر، اندام‌های انسان و حیوانات، برخی فعل‌های پایه، رنگ‌ها و مانند آن است. گردآوری داده‌ها حتی‌المقدور با استفاده از تصاویر تهیه شده از آن مفاهیم بوده است و در فاصله زمانی مه‌رمه ۹۵ تا آذر ۹۶ صورت گرفته است.

1. Polygon
2. Google Earth

۲-۲- مصاحبه‌شوندگان

مصاحبه‌شوندگان در این پژوهش، به صورت تصادفی از آبادی‌ها انتخاب شده‌اند. به منظور گردآوری داده‌ها، در هر مکان سعی شده است حداقل با سه نفر مصاحبه حضوری صورت پذیرد. سن مصاحبه‌شوندگان عموماً بین ۳۵ تا ۴۵ سال بوده است و به دلیل محدودیت‌های پژوهش غالباً مرد بوده‌اند.

۳-۲- تحلیل نرم افزاری

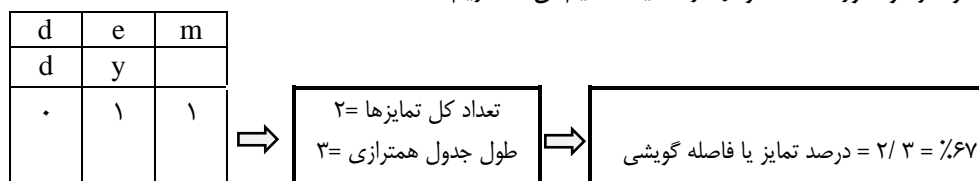
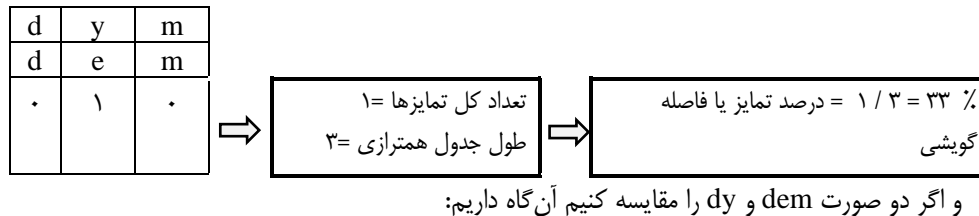
برای تحلیل جغرافیایی-زبانی داده‌ها از نرم‌افزار آنالین Gabmap بهره گرفته شده است. این نرم‌افزار، داده‌های زبانی و جغرافیایی را در قالب‌های مناسب و مورد قبول خود دریافت می‌کند و قادر است دسته‌بندی‌ها و نقشه‌های زبانی-جغرافیایی را تولید نماید.^۱ تعیین فواصل جغرافیایی-زبانی و ارائه انواع نقشه‌ها، نمودارها و میزان همبستگی داده‌ها، به صورت خروجی‌های مورد نیاز از مزیت‌های این نرم‌افزار است.

۳- تحلیل داده‌ها

همان‌گونه که پیش‌تر گفته شد، برای گویش‌سنجی از روش‌های مختلفی بهره گرفته شده است که از آن جمله می‌توان گویش‌سنجی با شیوه ارزش تفاوت نسبی^۲، ارزش تفاوت وزنی^۳ و فاصله لونشتین^۴ را نام برد. فاصله لونشتین را گاه فاصله ویرایشی^۵ نیز می‌نامند؛ این مفهوم را ولادیمیر لونشتین^۶ در سال ۱۹۵۶ برای نخستین بار مطرح کرد و در نظریه اطلاعات^۷، علوم رایانه و بعدها در گویش‌سنجی به کار گرفته شد و عبارت است از کمترین هزینه^۸ (= عملیات^۹) برای انطباق^{۱۰} دو زنجیره‌ی زبانی^{۱۱}. به بیان دیگر، این فاصله حداقل تعداد جایگزینی، درج و یا حذف برای تبدیل یک رشته زبانی به رشته دیگر است. در گویش‌سنجی برای محاسبه فاصله لونشتین، مقدار تفاوت‌ها را می‌توان با کمک جدول همترازی^{۱۲} به دست آورد. برای مثال، در سطح تلفظی اگر دو صورت زبانی [dym] و [dem] را که معادل برای واژه «دُم» در دو مکان مختلف هستند در نظر بگیریم، آن‌گاه با استفاده از جدول همترازی داریم:

۱. برای اطلاعات بیشتر با استفاده از مرورگر موزیلا فایر فاکس ر. ک. <https://www.gabmap.nl>

2. Relative Difference Value
3. Weighted difference value
4. Levenshtein Distance
5. edit distance
6. Vladimir Levenshtein
7. Information Theory
8. cost
9. operation
10. mapping
11. Linguistic string
12. alignment



در همترازی سازی زبانی، واژه‌ها با واژه‌ها و نیز همخوان‌ها با همخوان‌ها انطباق می‌یابند. تنوعات میان گویش‌ها را معمولاً بر اساس تعداد مشخصی از واژه‌ها می‌سنجند؛ ناگفته پیداست که برای به‌دست آوردن فاصله گویشی، هرچه تعداد واژه‌هایی که بررسی می‌شوند بیشتر باشند، دقت اندازه‌گیری بیشتر می‌شود. حال باید دید که این تنوعات چگونه در قالب اعداد و ارقام مشخص می‌توانند معرف فواصل گویشی باشند. اطلاعات جدول ۲ حاوی معادل‌های شش واژه (از میان صد واژه) و پنج مکان (از میان سی و پنج مکان) بررسی شده در این پژوهش می‌باشد. فرض کنیم بخواهیم فواصل گویشی روستاهای جدول زیر را بر اساس واژه‌های شش‌گانه آن جدول محاسبه کنیم.

جدول ۲: داده‌های زبانی پنج آبادی بر اساس شش واژه

واژه / آبادی	خوب	استخوان	خشک	دندان	مورچه	ناخن
زنجیره علیا	xas	seqan	hejk	degan	muriɜ	nyxan
شیاب	xas	seqan	hejk	denan	meruɜ	nexan
چم چرود	xu	sexan	hwejk	denan	meryɜ	naxwen
حسن گاوداری	xu	sexo	hejk	deno	meriɜ	naxo
سرابله	xas	seqan	hejk	denan	meryɜ	naxon

حال دو مکان زنجیره علیا و شیاب را از جدول فوق در نظر می‌گیریم؛ می‌توان با ترسیم جدول‌های همترازی مربوط به هر واژه در دو مکان انتخابی، فاصله زبانی دو مکان را در آن واژه به‌دست آورد. با محاسبه فاصله n واژه معادل در دو مکان انتخاب شده، می‌توان فاصله گویشی دو نقطه مکانی را بر اساس n واژه به‌دست آورد. جدول ۳ با نشان دادن فاصله هر یک از معادل‌ها در دو مکان زنجیره علیا و شیاب، بیانگر وجود ۴ تمایز زبانی از مجموع ۲۷ واحد طول جدول‌های همترازی رسم شده است که معادل ۱۴/۸ درصد می‌باشد.

جدول ۳: خلاصه جدول‌های همترازی دو مکان زنجیره علیا و شباب

واژه	زنجیره علیا	شباب	تعداد تمایزات (تغییرات لازم جهت انطباق)	طول جدول همترازی
خوب	xas	xas	۰	۳
استخوان	seqan	seqan	۰	۵
خشک	heʃk	heʃk	۰	۴
دندان	degan	denan	۱	۵
مورچه	muriʒ	meruʒ	۲	۵
ناخن	nyxan	nexan	۱	۵
۱۴٪ = تمایز دو روستا → ۲۷ / ۴		جمع تعداد تمایزها = ۴		
		جمع طول جدول‌های همترازی = ۲۷		

برای تعیین فاصله گویشی هر روستا از روستای دیگر، جدولی مانند جدول ۳ لازم است. با تکرار این جدول‌ها و به‌دست آوردن فاصله گویشی هر روستا از سایر روستاها، می‌توان به جدول ۴ رسید (اعداد بر حسب درصد هستند).

جدول ۴: فاصله گویشی پنج مکان بر اساس شش واژه

زنجیره علیا	شباب	چم چرود	حسن گاوداری	سرابله
۰	۱۴/۸	۳۴/۴۸	۴۴/۴۴	۱۸/۵۱
۱۴/۸	۰	۲۵/۸	۳۹/۲۸	۱۱/۱۱
۳۴/۴۸	۲۵/۸	۰	۳۲/۱۴	۲۱/۴۲
۴۴/۴۴	۳۹/۲۸	۳۲/۱۴	۰	۳۳/۳۳
۱۸/۵۱	۱۱/۱۱	۲۱/۴۲	۳۳/۳۳	۰

داده‌های جدول ۴، مبین فواصل گویشی هر یک از مکان‌های مورد بررسی (بر اساس شش واژه پیش گفته) با سایر مکان‌هاست. به منظور نمایش بهتر و داشتن درک مناسب‌تر در گویش‌سنجی، از روش‌های نمایش فواصل گویشی استفاده می‌شود که در ذیل به چند شیوه اشاره می‌گردد. از تحلیل خوشه‌ای برای گروه‌بندی گویشی استفاده می‌شود. برای گروه‌بندی پنج مکان جدول ۴، کوچک‌ترین عدد را می‌یابیم که مربوط به فاصله گویشی میان شباب و سرابله است و برابر ۱۱/۱۱ درصد است؛ بنابراین شباب و سرابله را یک خوشه در نظر می‌گیریم و برای به‌دست آوردن فاصله سایر مکان‌ها از این خوشه، جدول زیر را ترسیم می‌کنیم؛ براساس فاصله هر یک از مکان‌های دیگر از شباب و نیز سرابله، فاصله آنها از خوشه را به‌دست می‌آوریم که برابر با میانگین دو عدد قبلی است. به‌عنوان مثال فاصله زنجیره

علیا از خوشه (شباب و سرابله) برابر است با میانگین فاصله زنجیره علیا از شباب (۱۴/۸) و زنجیره علیا از سرابله (۱۸/۵۱) یعنی: $(۱۴/۸۰ + ۱۸/۵۱) / ۲ = ۱۶/۶۵$

جدول ۵: فاصله هر مکان تا خوشه (شباب و سرابله)

(شباب و سرابله)	سرابله	شباب	
۱۶/۶۵	۱۸/۵۱	۱۴/۸۰	زنجیره علیا
۲۳/۶۱	۲۱/۴۲	۲۵/۸۰	چم چرود
۳۶/۳۰	۳۳/۳۳	۳۹/۲۸	حسن گاوداری

با تشکیل خوشه (شباب و سرابله) جدول ۴ به صورت زیر بازنویسی می شود.

جدول ۶: فاصله گویشی پنج مکان بر اساس شش واژه

حسن گاوداری	چم چرود	(شباب و سرابله)	زنجیره علیا	
۴۴/۴۴	۳۴/۴۸	۱۶/۶۵	۰	زنجیره علیا
۳۶/۳۰	۲۳/۶۱	۰	۱۶/۶۵	(شباب و سرابله)
۳۲/۱۴	۰	۲۳/۶۱	۳۴/۴۸	چم چرود
۰	۳۲/۱۴	۳۶/۳۰	۴۴/۴۴	حسن گاوداری

برای تشکیل دادن سایر خوشه‌ها، عملیات بالا را با اعداد جدول جدید تکرار می‌کنیم؛ بدین صورت که کوچکترین عدد را می‌یابیم که مربوط به فاصله میان زنجیره علیا و خوشه (شباب و سرابله) است و خوشه جدید را ترسیم و فاصله گویشی سایر مکان‌ها را از خوشه جدید به دست می‌آوریم.

جدول ۷: فاصله هر مکان تا خوشه (زنجیره علیا و (شباب و سرابله))

(زنجیره علیا و (شباب و سرابله))	زنجیره علیا	(شباب و سرابله)	
۲۹/۰۵	۳۴/۴۸	۲۳/۶۱	چم چرود
۴۰/۳۷	۴۴/۴۴	۳۶/۳۰	حسن گاوداری

و بدین ترتیب می‌توانیم جدول ۶ را نیز به صورت جدول زیر بازنویسی کنیم.

جدول ۸: فاصله گویشی پنج مکان بر اساس شش واژه

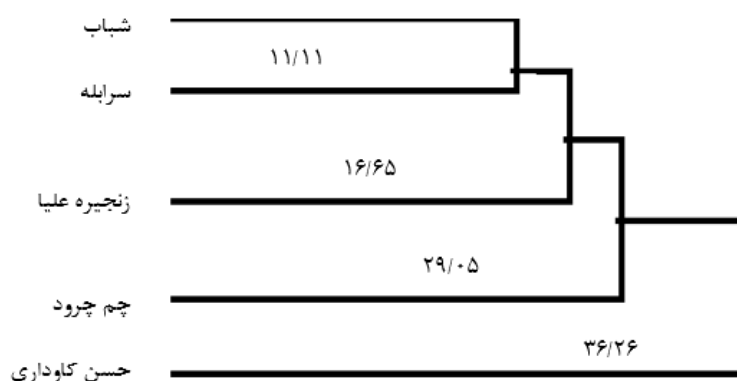
حسن گاوداری	چم چرود	(زنجیره علیا و (شباب و سرابله))	
۴۰/۳۷	۲۹/۰۵	۰	(زنجیره علیا و (شباب و سرابله))
۳۴/۱۴	۰	۲۹/۰۵	چم چرود
۰	۳۴/۱۴	۴۰/۳۷	حسن گاوداری

از میان سه فاصله باقیمانده، کوچکترین عدد مربوط به فاصله چم چرود و خوشه (زنجیره علیا و (شباب و سرابله)) است. پس داریم:

جدول ۹: فاصله حسن گاوداری تا خوشه (چم چرود و (زنجیره علیا و (شباب و سرابله)))

چم چرود	(زنجیره علیا و (شباب و سرابله))	(چم چرود و (زنجیره علیا و (شباب و سرابله)))	حسن گاوداری
۴۰/۳۷	۳۲/۱۴	۳۶/۲۶	

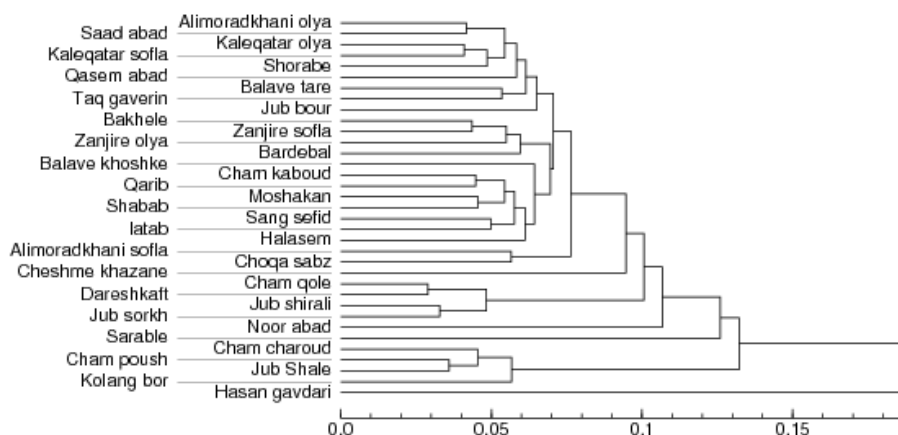
اطلاعات به دست آمده از تحلیل خوشه بالا را می‌توان در قالب دندوگرام^۱ زیر ترسیم نمود.



نمودار ۱: دندوگرام پنج مکان بر اساس شش واژه

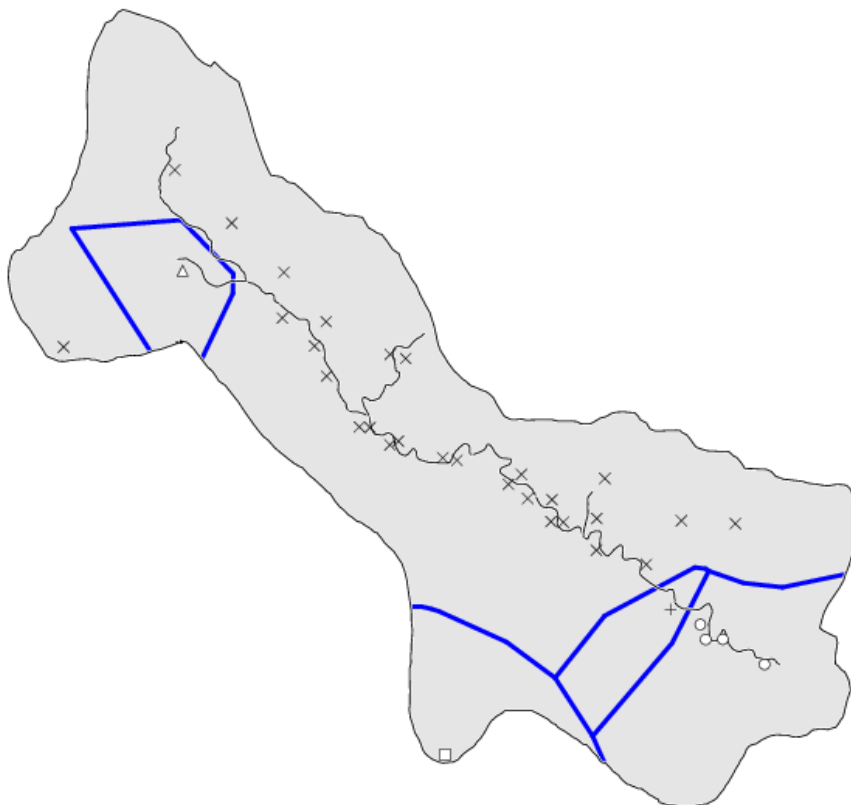
اگر تعداد مکان‌های مورد بررسی را به همه ۳۵ مکان مورد بررسی در این پژوهش بسط دهیم و نیز تعداد واژه‌ها را از شش واژه فوق به صد واژه فهرست واژگان سوادش برسانیم، آن‌گاه با ملاحظاتی که در ادامه خواهد آمد، نمودار دندوگرام زیر به دست می‌آید. ناگفته پیداست که با لحاظ کردن تعداد بیشتر واژه‌ها و مکان‌ها، دندوگرام جدید لزوماً بر دندوگرام قبل منطبق نیست.

1. dendogram



نمودار ۲: دندوگرام منطقه مورد مطالعه بر اساس صد واژه لیست سوادش

در جداول همترازی که پیشتر درباره آنها سخن گفتیم، از ارزش‌های صفر و یک بهره گرفته شده است. به منظور مدرج کردن ارزش‌گذاری در فواصل گویشی، می‌توان با اعمال تعدیل‌هایی، سنجش‌های واقعی‌تری را رقم زد. در ارزش‌گذاری مقادیر جدول‌های همترازی که مبنای دندوگرام فوق قرار گرفته است، برای هر جایگاه تبدیل C به C (همخوان به همخوان) ارزش عددی ۱ لحاظ شده است. در مواردی که تبدیل V به V (واکه به واکه) رخ داده است، ارزش عددی ۰/۵ در نظر گرفته شده است. اگر تبدیل V به C (واکه به همخوان) و بالعکس رخ داده باشد، ارزش عددی ۲ در نظر گرفته شده است. اگر بخواهیم نتایج تحلیل خوشه‌ای داده‌ها را بر روی نقشه نمایش دهیم، با توجه به تعداد دسته‌هایی که در نظر می‌گیریم، ممکن است حالت‌های مختلف نمایش داده شود. نقشه زیر منطقه مورد مطالعه را به پنج قسمت زبانی دسته‌بندی کرده است.

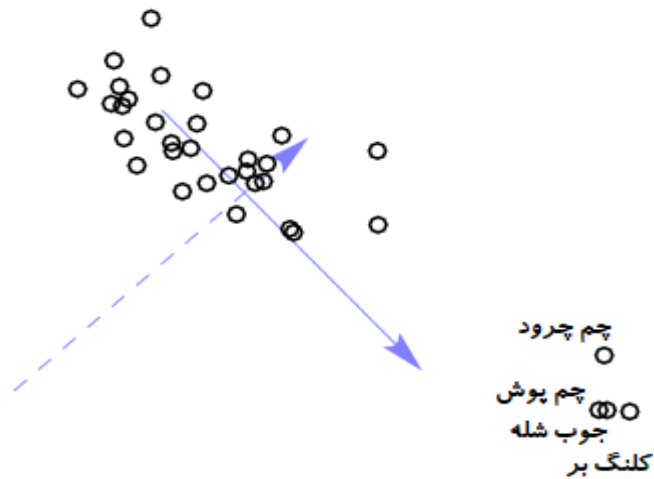


شکل ۲: نقشه تفکیکی منطقه مورد مطالعه بر اساس تقسیم مناطق گویشی به پنج منطقه

یکی دیگر از شیوه‌های نمایش فاصله گویشی میان مکان‌ها، استفاده از نمودار مقیاس‌گذاری چندوجهی است. در آمار و علوم وابسته به آن، مقیاس‌گذاری چندوجهی شیوه‌ای برای ارائه و نمایش فاصله‌ها و تمایزات میان دسته‌های داده‌هاست (کروسکال و ویش، ۱۹۷۸). هر مکان حاوی مجموعه‌ای از داده‌های زبانی منحصر به فرد است. مکان‌هایی که دارای شباهت بیشتر در داده‌های زبانی خود باشند (به دیگر سخن، فاصله زبانی کمتری داشته باشند)، بر روی نمودار مقیاس‌گذاری چندبعدی به همدیگر نزدیک‌تر خواهند بود. همان‌گونه که پیش‌تر گفته شد، استفاده از نمودار مقیاس‌گذاری چندبعدی، شیوه‌ای برای کاهش در تعداد ابعاد اطلاعات چندبعدی است به‌گونه‌ای که بتوان میزان همبستگی موجود در داده‌ها را در تعداد ابعاد کمتر نمایش داد (بویا و همکاران، ۲۰۰۷). همان‌گونه که در نمودار مقیاس‌گذاری چندبعدی زیر مشاهده می‌شود، روستای

1. Kruskal & Wish
2. Buja et al

لر زبان حسن گاوداری (شماره ۲۷) دارای بیشترین فاصله از سایر مکان‌هاست. چهار روستای لک زبان (شماره‌های ۳۲ تا ۳۵) نیز در جایگاهی نزدیک به هم و دور از سایر مکان‌ها قرار دارند.

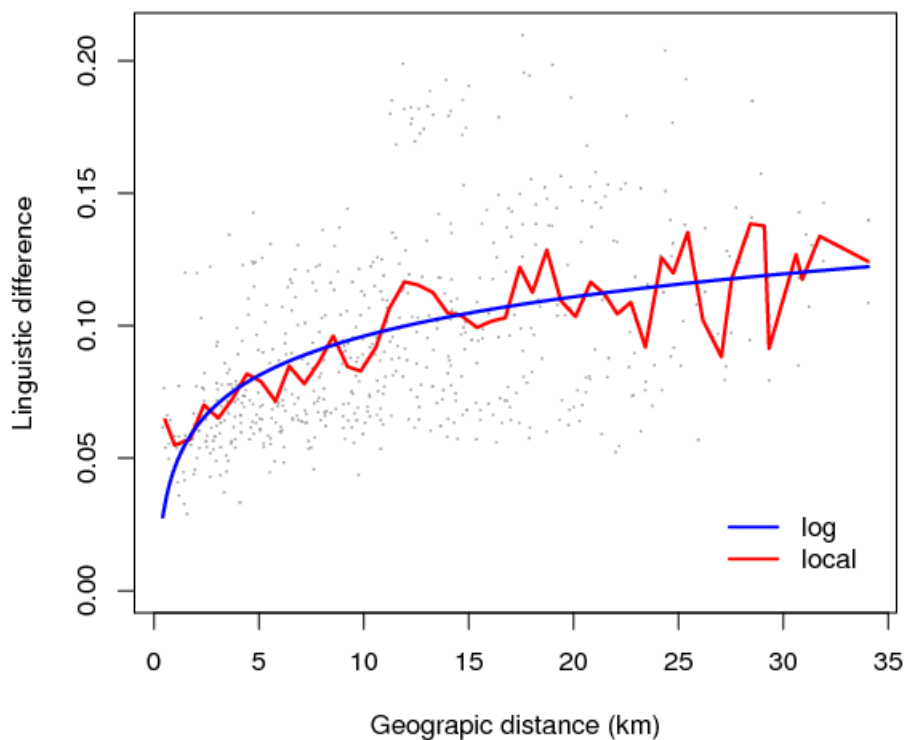


حسن گاوداری

نمودار ۳: نمودار مقیاس گذاری چند بعدی منطقه مورد مطالعه

ارزش تفاوت‌ها میان فواصل کوچک‌تر، به صورت نسبی بزرگ‌تر است از ارزش تفاوت‌ها میان فواصل بزرگ‌تر؛ این بدان معناست که ارزش تفاوت میان ۲ و ۳ به صورت نسبی بزرگ‌تر از ارزش تفاوت میان ۲۰ و ۲۱ است؛ زیرا در حالت اول مقدار تفاوت دو عدد برابر است با ۵۰ درصد ارزش عدد کوچکتر (یعنی $1/2$) و در حالت دوم مقدار تفاوت دو عدد برابر است با ۵ درصد ارزش عدد کوچکتر (یعنی $1/20$). برای نشان دادن هرچه بهتر این موضوع، در آمار برای سنجش فاصله‌ها از فواصل لگاریتمی^۱ استفاده می‌شود. نمودار زیر فاصله لگاریتمی و فاصله خطی میان متغیرهای جغرافیایی - زبانی را در پژوهش حاضر نمایش می‌دهد.

1. Logarithmic distances

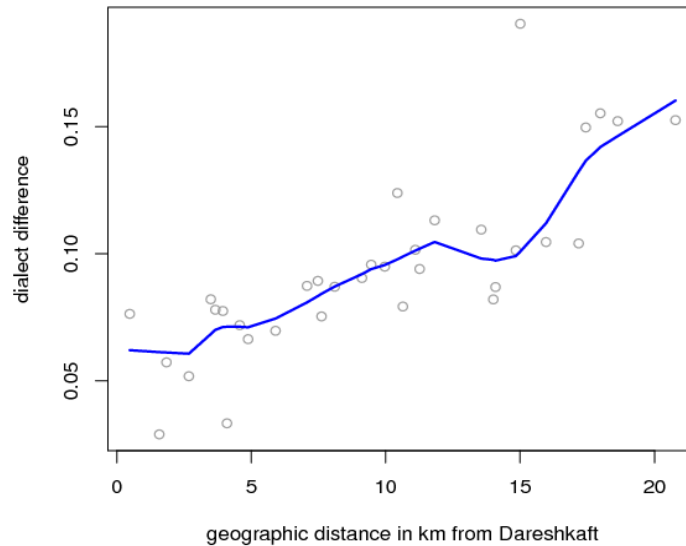


نمودار ۴: نمایش فاصله خطی و فاصله لگاریتمی میان متغیرهای جغرافیایی-زبانی در منطقه مورد پژوهش

اگر بخواهیم داده‌های زبانی یک مکان مشخص و یا هر گونه زبانی معیار دیگری را به‌عنوان نقطه مرجع^۱ تعیین کنیم، و فواصل جغرافیایی-زبانی دیگر را نسبت به آن سنجش کنیم، از نموداری مانند نمودار زیر استفاده می‌کنیم که در آن همبستگی متغیرهای فاصله جغرافیایی تفاوت‌های گویشی، نشان داده شده است. این نمودار نشان می‌دهد که چگونه با افزایش فاصله جغرافیایی، تفاوت‌های زبانی افزایش یافته‌اند و آهنگ این تغییر، از نقطه مرجع (در این مثال روستای دراشکفت مکان شماره ۱۲) به چه میزان است.^۲

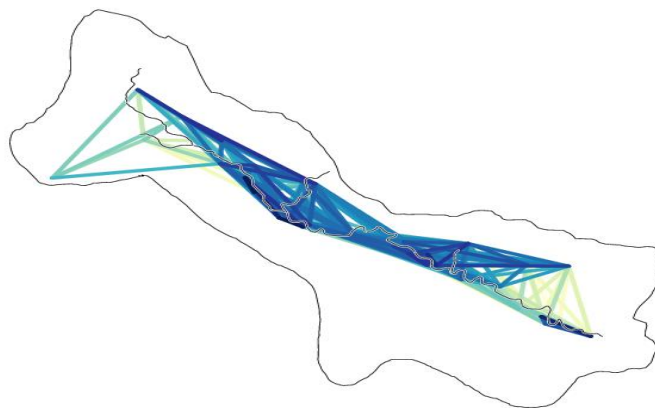
1. Reference point

۲. در نرم‌افزار Gabmap نقطه مرجع به‌گونه‌ای قابل تعیین است که برای همه مکان‌ها همبستگی جغرافیایی-زبانی را بتوان به‌دست آورد.



نمودار ۵: نمایش همبستگی جغرافیایی-زبانی منطقه مورد مطالعه از نقطه مرجع روستای دراشکفت

یکی دیگر از شیوه‌های نمایش فواصل جغرافیایی-زبانی، استفاده از نقشه پرتوی^۱ است. در این شیوه هر مکان، به وسیله یک خط مستقیم بر هر یک از سایر مکان‌ها متصل می‌شود. هر چه شباهت زبانی بیشتر باشد، این خط تیره‌تر و هر چه فاصله زبانی بیشتر باشد، این خط کم‌رنگ‌تر خواهد بود. شکل ۳ نقشه پرتوی منطقه مورد مطالعه را نمایش می‌دهد.



شکل ۳: نقشه پرتوی منطقه مورد مطالعه

۴- نتیجه‌گیری

گوش‌سنجی، به‌عنوان رویکردی آمارگرا، با ارائه معیارهای مشخص در تعیین فاصله‌ی زبانی میان گونه‌های زبانی و ارتباط آن با عوامل غیرزبانی مانند جغرافیا، سعی دارد تا عوامل مؤثر در گوناگونی‌های زبانی را تبیین نماید. از ویژگی‌های بارز مطالعات مبتنی بر گوش‌سنجی به‌دست آوردن نتایج تحلیلی به دور از پیش‌فرض‌ها و اتکای صرف به داده‌های زبانی، شفافیت و عدم سوگیری است، به‌گونه‌ای که به کمک نمودارها و اعداد و ارقام می‌توان فواصلی گویشی را تعیین و میزان تشابه و افتراق گونه‌های زبانی را در یک پهنه‌ی جغرافیایی مشخص از همدیگر باز شناخت؛ خواه این مکان‌ها از هم دور باشند خواه نزدیک. هر یک از نمودارهای معرفی‌شده در این بررسی اطلاعاتی را درباره‌ی ارتباط مؤلفه‌های زبانی و مؤلفه‌های جغرافیایی ارائه می‌کند که برخی از آنها مانند نمودارهای ۲ و ۳ تبیین‌کننده‌ی اطلاعات کلی منطقه مورد مطالعه هستند و برخی دیگر مانند نمودار ۵ جزئیات بیشتری را از یک مکان خاص و رابطه‌ی زبانی-جغرافیایی آن با سایر مکان‌ها نمایان می‌سازد. همان‌گونه که پیش‌تر گفته شد، نرم‌افزار Gabmap این امکان را فراهم آورده است که درباره‌ی یک مکان خاص، هر یک از مؤلفه‌های زبانی-مکانی به‌صورت معین با اطلاعات بیشتری نمایش داده شوند که نمودارهای ارائه شده تنها به‌عنوان نمونه ارائه شده‌اند.

انطباق یافته‌های گوش‌سنجی درباره‌ی مرزهای گویشی با عوامل مؤثر در ایجاد گوناگونی‌های زبانی از مباحث درخور توجه است. تنوعات زبانی ناشی از عواملی چون فرهنگ، قومیت، مذهب، مرزهای سیاسی (چون استان شهرستان دهستان و مانند آن) و عوامل طبیعی (چون کوه‌ها، دره‌ها، رودها و مانند آن) را می‌توان با ابزارهای گوش‌سنجی اندازه‌گیری نمود و میزان تأثیرگذاری هر یک از این عوامل را بر مشخصه‌های زبانی سنجید.

موقعیت جغرافیایی منطقه چرداول به‌گونه‌ای است که در همسایگی استان‌های کرمانشاه و لرستان و در شمال استان ایلام قرار دارد. در این منطقه، به‌گونه‌های مختلف کردی، لکی و لری تکلم می‌شود و می‌توان تأثیرپذیری متقابل گونه‌های زبانی را از همدیگر در میان داده‌های زبانی مشاهده کرد. این پژوهش کوشیده است تا با معیارهای گوش‌سنجی، مرز نسبی میان گونه‌های زبانی را مشخص کند. به‌عنوان نمونه، نمودار ۳ که مربوط به نمودار مقیاس‌گذاری چند بعدی منطقه مورد مطالعه است، تناسب میان تنوعات زبانی و جغرافیا را منعکس می‌کند که در آن سه دسته زبانی شکل گرفته‌اند: روستاهای کردنشین که بیشترین فراوانی را به خود اختصاص داده‌اند، روستاهای لک‌نشین و نیز تک روستای لرنشین منطقه مورد مطالعه در این نمودار قابل مشاهده‌اند.

همسویی و یکپارچگی در ارائه‌ی تحلیل‌های آماری از ویژگی‌های نتایج حاصل از رویکردهای گوش‌سنجی است. مثلاً در نمودارهای ۲ و ۳ و نیز شکل‌های ۲ و ۳ می‌توان به وضوح تفاوت زبانی مکان شماره ۲۷ (حسن گاوداری) را در قیاس با مکان‌های مجاور مشاهده نمود؛ این تفاوت فاصله زبانی در همه نمودارهای ارائه شده منعکس شده است. از این منظر می‌توان روستای لرنشین حسن گاوداری را یک جزیره‌ی زبانی دانست که دارای بیشترین فاصله زبانی با روستاهای مجاور است که عموماً کردنشین هستند.

انطباق یافته‌های پژوهش حاضر با باورهای سنتی حاکم در منطقه مورد مطالعه درباره دسته‌بندی گونه‌های زبانی نیز، خود می‌تواند موضوع تحقیقی مستقل باشد.

منابع

- اسدپور، هیوا (۱۳۹۰)، اطلس زبانی رایانه‌ای استان آذربایجان غربی: رویکردی رده‌شناختی ادراکی در زبان‌شناسی جغرافیایی، پایان‌نامه کارشناسی‌ارشد رشته زبان‌شناسی، دانشگاه آزاد اسلامی واحد علوم و تحقیقات.
- اسدپور، هیوا (۱۳۹۱)، «بررسی گونه‌های زبانی در استان آذربایجان غربی با روش فاصله‌سنجی گویشی»، ویژه‌نامه فرهنگستان (زبان‌ها و گویش‌های ایرانی)، دوره جدید، ش ۱، ۱۷۳-۲۰۲.
- آفاگل‌زاده، فردوس (۱۳۹۰)، «الگوی برای گردآوری داده‌های گویشی تنوعات لهجه‌ای و تهیه نقشه زبان‌شناسی (اطلس زبانی) ایران»، مجموعه مقالات نخستین همایش بین‌المللی گویش‌های مناطق کویری ایران (جلد اول). سمنان: دانشگاه سمنان.
- بصیرت، علی و فامیان، علی‌رضاقلی (۱۳۹۰)، «به سوی طراحی اطلس زبانی گویش گیلکی»، ادب پژوهی، ش ۱۷، ۱۸۳-۱۶۵.
- پرمون، یدالله (۱۳۸۶)، طرح ملی اطلس زبانی ایران. کتابچه جامع تدوین اطلس (مستندسازی، پایگاه داده‌ها، نقشه)، تهران: سازمان میراث فرهنگی، صنایع دستی و گردشگری، پژوهشکده زبان و گویش.
- پرمون، یدالله و زارعی، زهره (۱۳۸۵)، اطلس زبانی ایران: مبانی و روش‌شناسی: گزارش اطلس زبانی اصفهان، اصفهان: سازمان میراث فرهنگی، صنایع دستی و گردشگری استان اصفهان.
- جهانگیری، نادر (۱۳۵۲)، بررسی تعداد، توزیع و گسترش گویش‌های رایج در استان مازندران، هدف تهیه‌ی اطلسی از گویش‌های رایج، پایان‌نامه کارشناسی‌ارشد زبان‌شناسی دانشگاه تهران.
- خادمی، شیرین (۱۳۸۲)، طرح ملی گویش‌شناسی (کردستان و کرمانشاه)، پایان‌نامه کارشناسی‌ارشد رشته زبان‌شناسی، دانشگاه آزاد اسلامی واحد تهران مرکزی.
- دورقمری، حمیرا (۱۳۸۲)، طرح ملی گویش‌شناسی (استان‌های تهران و خراسان)، پایان‌نامه کارشناسی‌ارشد رشته زبان‌شناسی، دانشگاه آزاد اسلامی واحد تهران مرکزی.
- دیهیم، گیتی (۱۳۸۸)، اطلس گویش‌شناختی قصران داخل، تهران: انتشارات سروش با همکاری دفتر پژوهش‌های فرهنگی.
- رستم‌بیک تفرشی، آتوسا (۱۳۸۲)، طرح ملی گویش‌شناسی (استان‌های یزد و همدان)، پایان‌نامه کارشناسی‌ارشد رشته زبان‌شناسی، دانشگاه آزاد اسلامی واحد تهران مرکزی.
- زارعی، زهرا (۱۳۹۱)، طرح ملی گویش‌شناسی (شهرستان نیشابور)، پایان‌نامه کارشناسی‌ارشد رشته زبان‌شناسی، دانشگاه آزاد اسلامی واحد تهران مرکزی.
- سنایی، یاسر (۱۳۹۵)، بررسی تنوع گویشی نواحی شمال استان ایلام: تدوین اطلس گویشی، پایان‌نامه کارشناسی - ارشد رشته زبان‌شناسی، دانشگاه پیام نور.

- صادقی، مصطفی (۱۳۹۴)، اطلس زبانی برخی صورت‌های واژگانی زبان ترکی بخش سرولایت نیشابور، پایان‌نامه کارشناسی‌ارشد رشته زبان‌شناسی، دانشگاه فردوسی مشهد.
- مقدم، صفیه (۱۳۸۶)، طرح ملی گویش‌شناسی (گویش‌های استان‌های کرمان و اصفهان)، پایان‌نامه کارشناسی‌ارشد رشته زبان‌شناسی همگانی، پژوهشگاه علوم انسانی و مطالعات فرهنگی.
- ملایی پاشایی، سیف‌اله (۱۳۹۳)، گویش‌سنجی رایانشی دامنه شمالی البرز مرکزی بر پایه الگوریتم لونستین: تدوین یک اطلس زبان‌شناختی، رساله دکتری زبان‌شناسی همگانی، دانشگاه پیام نور.
- ملایی قاسم، مسعود (۱۳۷۴)، بررسی گسترش جغرافیایی و ترسیم اطلس زبانی بخش پشت‌آب سیستان، پایان‌نامه کارشناسی‌ارشد زبان‌شناسی. دانشگاه فردوسی مشهد.
- مولایی، حامد (۱۳۹۰)، بررسی توزیع ۱۱۴ متغیر زبان فارسی و رسم اطلس زبانی بخش‌های کوهبنان و خرم‌دشت در استان کرمان، پایان‌نامه کارشناسی‌ارشد زبان‌شناسی همگانی، دانشگاه سیستان و بلوچستان.
- الیاسی، محمود (۱۳۸۹)، بررسی گسترش جغرافیایی و ترسیم اطلس زبانی ۱۰۰ صورت زبانی در حدفاصل شهرهای تربت‌جام و مشهد، پایان‌نامه کارشناسی‌ارشد زبان‌شناسی دانشگاه فردوسی مشهد.
- Bills, Garland D. and Neddy A. Wigil (2008), *The Spanish language of New Mexico and Southern Colorado, A linguistic Atlas*. University of New Mexico Press. New Mexico.
- Buja, A. Deborah F Swayne, Michael L Littman, Nathaniel Dean, Heike Hofmann, Lisha Chen (2008), *Data Visualization with Multidimensional Scaling*. Journal of Computational and Graphical Statistics. Volume 17, 444 – 472.
- Chambers, J. K. & P. Trudgill, (2004), *Dialectology* (2nd edition). Cambridge: Cambridge University Press, 2nd edition.
- Dryer, Matthew S.; D. Gil; B. Comrie; H. Jung & C. Schmidt (2005), *The World Atlas of Language Structures*. Oxford: Oxford University Press.
- Francis, W. Nelson (1983), *Dialectology: An Introduction*. London. Longman
- Goebel, H. & G. Schiltz (1997), "A dialectometrical compilation of CLAE 1 and CLAE 2: Isoglosses and dialect integration". In: Viereck, W.; Ramisch, H. (Ed.). *Computer Developed Linguistic Atlas of England (CLAE)*. Tübingen: Max Niemeyer Verlag, v. 2.
- Hartmann, R. R. K and James, Gregory (2002), *Dictionary of lexicography*. London and New York. Routledge.
- Kessler, B. (1995), "Computational dialectology in Irish Gaelic". In *Proceedings of 7th Conference of the European Charter of the Association for Computational Linguistics (EACL 7)*, Dublin, 60-66.
- Kruskal, Joseph B. & Myron Wish (1978), "Multidimensional Scaling". Sage University Publication. London.
- Labov, W., Sh. Ash & Ch. Boberg (2005), *The Atlas of North American English: Phonetics, Phonology and Sound Change*. Berlin: Walter de Gruyter.
- Nerbonne, John and William kretzschmar (2003), *Introducing Computational Techniques in Dialectometry*. Kluwer Academic Publishers, Netherlands.
- Nerbonne, John. and W. Heeringa & P. Kleiweg (2003), "Lexical distance in LAMSAS". *Computers and the Humanities*. 37 (4). 339-357.

- Nerbonne, John. and W. Heeringa (2001), "Computational Comparison and Classification of Dialects". *Dialectologia et Geolinguistica. Journal of the International Society for Dialectology and Geolinguistics*, 2001(9). 69-83.
- Séguy, Jean, (1971), "La relation entre la distance spatiale et la distance lexicale," *Revue de Linguistique Romane* 35, 335-57.
- Séguy, Jean, (1973), *La Dialectométrie Dans l'Atlas Linguistique de la Gascogne*. Société de linguistique romane.
- Wieling M, Nerbonne J, Baayen RH (2011), *Quantitative Social Dialectology: Explaining Linguistic Variation*.