

بنام خدا

وزارت علوم، تحقیقات و فناوری

مرکز منطقه ای اطلاع رسانی علوم و فناوری

گزارش نهایی طرح تحقیقاتی

تهیه پیکره (فرهنگ) املائی واژه‌های هم‌آوا و هم‌نویسه در زبان فارسی

مجری :

شاپوررضا برنجیان

مربی گروه پژوهشی زبانشناسی رایانه‌ای

آذر ماه ۱۳۹۸

## فهرست مندرجات

صفحه	عنوان
۴.....	چکیده
۵.....	فصل اول: کلیات تحقیق
۵ .....	۱-مقدمه
۶ .....	۱-۱- بیان مسئله
۱۲.....	فصل دوم:
۱۲.....	۲-پیشینه تحقیق
۱۴.....	۲-۱- میانی نظری
۱۷.....	۲-۱-۱- انواع پیکره
۲۰ .....	۲-۲- ضرورت انجام طرح
۲۱ .....	۲-۳- اهداف پژوهش
۲۳ .....	۲-۴- جامعه پژوهش
۲۵.....	فصل سوم: روش و شیوهی تحقیق
۲۵ .....	۳-روش پژوهش
۲۷.....	فصل چهارم: تجزیه و تحلیل داده‌ها
۲۷ .....	۴-روش و ابزار گردآوری داده‌ها
۲۷.....	۴-۱- خروجی‌های این طرح

۲۸	۴-۲- این طرح می‌تواند در موارد زیر بکار رود .....
۲۸	۴-۳- یافته‌ها .....
۲۹	فصل پنجم: بحث و نتیجه‌گیری .....
۲۹	۵- روابط واژگانی .....
۳۰	۵-۱- هم‌ها .....
۳۱	۵-۲- هم معنایی .....
۳۲	۵-۳- هم معنا، هم شکل، غیر هم آوا .....
۳۲	۵-۴- چند معنایی .....
۳۳	۵-۵- هم آوا هم معنا، غیر هم شکل .....
۳۳	۵-۶- نتیجه‌گیری .....
۳۴	منابع و ماخذ .....
۳۷	پیوست ۱ .....
۴۸	پیوست ۲ .....
۷۲	پیوست ۳ .....
۸۷	پیوست ۴ .....

## سپاسگزاری

تهیه، تنظیم و تدوین این طرح بدون همکاری و یاری افراد گوناگون امکان پذیر نبود، لذا در اینجا بر خود لازم می‌دانم تا از تک تک آن بزرگواران تشکر و سپاسگزاری نمایم.

۱) جناب آقای دکتر محمد جواد دهقانی، ریاست محترم مرکز منطقه‌ای که در تمام مراحل کار با دقت و حوصله نظارت مستمر داشته و همواره راهنما و مشوق بوده‌اند.

۲) جناب آقای دکتر محمدرضا صالحی، معاونت محترم پژوهشی و فناوری مرکز منطقه‌ای که در تمام مراحل کار اعم از تهیه پروپوزال و ... با دقت و حوصله همواره راهنما و مشوق بوده و نظرات و پیشنهادات سازنده‌ای ارائه فرموده و امکان اجرایی شدن این طرح را فراهم نموده‌اند.

۳) اعضای محترم شورای علمی که در هنگام تصویب طرح نظرات سازنده‌ای ارائه نمودند.

شاپوررضا برنجیان

مربی گروه پژوهشی زبانشناسی رایانه‌ای

## تهیه پیکره (فرهنگ) املائی واژه‌های هم‌آوا و هم‌نویسه در زبان فارسی

شاپور رضا برنجیان

مربی گروه پژوهشی زبانشناسی رایانه‌ای

berenjia@ricest.ac.ir

### چکیده

یکی از مهم‌ترین گام‌ها در طراحی و ساخت سیستم‌های بازشناسی گفتار و تصحیح خطاهای آوایی و غلط‌یاب‌ها و استاندارد سازی متون فارسی نحوه‌ی تهیه‌ی یک پیکره و واژه‌نامه‌ی بی‌عیب و نقص در یک زبان خاص است. برای چنین عملی، داشتن مقادیر بسیار زیادی از داده‌های تصفیه شده‌ی استاندارد (که قبلاً بارها بررسی شده باشند) از آن زبان خاص ضروری است. هدف این پژوهش تهیه‌ی گنجینه‌ای از واژه‌های هم‌آوا و واژه‌های هم‌نویسه جهت جوابگویی به نیازهایی از این دست می‌باشد. جامعه‌ی پژوهش این طرح را تمامی واژه‌های بخش اول فرهنگ معین (۱۳۷۱) تشکیل می‌دهد. این بخش شامل پنج هزار و دویست و هفتاد و هفت (۵۲۷۷) صفحه و حدوداً بیش از سی و هفت هزار (۳۷۰۰۰) مدخل می‌باشد. روش پژوهش ما ترکیبی از روشهای زبانشناسی و رایانشی است. نتایج به دست آمده نشان داد که کل واژه‌های استخراج شده ۱۳۷۷۰ مدخل همراه با معانی واژه‌ها می‌باشند که از این تعداد ۲۸٪ را «هم‌آواها» و ۳۸٪ را «هم‌نویسه‌ها» و ۱۵٪ را «امکان خطا در املاء» و ۱۹٪ را نیز «هم‌نام‌ها» تشکیل می‌دهند. در این تحقیق برای اولین بار واژه‌نامه‌ی «امکان خطا در املاء» که شامل هم‌آواها و واژه‌های دارای واج‌های مشابه بودند، استخراج گردید تا بتواند در اختیار آموزگاران و مدرسین دوره‌های ابتدایی برای تدریس واژه‌های دشوار قرارگیرد. این پیکره‌ها برای اشتراک گذاری و دسترسی آسان پژوهشگران می‌توانند به صورت برخط در سامانه مرکز منطقه‌ای اطلاع رسانی علوم و فناوری برای استفاده کاربران نصب گردند.

کلید واژه‌ها: پیکره، واژه‌های هم‌آوا، واژه‌های هم‌نویسه، زبان فارسی.

## فصل اول: کلیات تحقیق

### ۱- مقدمه

پیکره در لغت به معنی "زمینه، شالوده، اساس" و در اصطلاح امروز "شکل، نقش، تصویر، صورت، تمثیل، چهره، عکس و صورت نگاشت" می‌باشد (دهخدا، ۱۳۷۱).

بنابر تعریف، پیکره زبانی عبارت است از مجموعه‌ای از متن‌های نوشتاری یا گفتاری آوانویسی شده که می‌توان در توصیف و تحلیل زبان از آن بهره گرفت (Kennedy: 1998). به عبارت دیگر پیکره واژگانی (corpus) به مجموعه‌ای از داده‌های خام واژگانی زبانی گفته می‌شود که می‌توان از آن‌ها برای تحلیل زبان بهره گرفت. گرچه در گذشته، بسیاری از زبان‌شناسان بر اهمیت پیکره زبانی در بیشتر بررسی‌هایشان تأکید کرده‌اند، اما در دوران جدید است که تکیه بر داده‌های واقعی زبانی به صورت گسترده‌ای رواج یافته و شرط اساسی بسیاری از پژوهش‌های نظری و کاربردی مانند: گویش‌شناسی، دستورنویسی، سازماندهی، تنظیم و تفکیک، جستجو و دستیابی سریع داده‌های زبانی شده است و افق‌های تازه‌ای در برابر پژوهندگان گشوده و باعث پیدایش شاخه‌ای تخصصی در حوزه زبان‌شناسی رایانه‌ای شده است. این شاخه، با نام زبان‌شناسی پیکره‌ای تنها در آخرین دهه‌های قرن بیستم ایجاد شد و در همین زمان کوتاه تبدیل به یکی از فعال‌ترین و پرکاربرد-ترین زمینه‌ها شده است. تالیف کتاب‌های متعدد و ایجاد مجامع علمی، مراکز پژوهشی، نشریه‌های تخصصی و بین‌المللی و سایت‌های اینترنت برای این شاخه، گویای این واقعیت است. پیکره هم‌آوا و هم‌نویسه حاضر، از فرهنگ معین (۱۳۷۱) استخراج گردید، در فرایند تهیه این پیکره‌ها توانستیم دو پیکره دیگر نیز تحت عنوان پیکره «هم‌نام‌ها (چند معنایی)» و «امکان خطا در املاء» را نیز استخراج نماییم.

انسان‌ها غالباً به دلیل بی‌توجهی و یا عدم دقت یا عدم دانش کافی خطای تایپی انجام می‌دهند. خطاهای تایپی دسته بزرگی از خطاهای معمول در متون الکترونیکی و تایپی هستند که توسط انسان و یا نرم‌افزارها ایجاد می‌شوند (دستغیب، ۱۳۹۵). در هنگام نوشتن نامه یا املاء یا هر سند دیگر ممکن است در اثر بی‌توجهی یا بی‌دقتی یا کم‌بود دانش نویسنده، واژه‌های مختلف به دلیل مشکلات موجود در خط فارسی به شکل‌های گوناگون نوشته شوند که فقط یکی از آن‌ها در جمله صحیح باشد و مابقی نادرست محسوب گردد.

به جهت کمک به حل مشکلات نوشتاری زبان فارسی، ما به انجام این پژوهش اقدام کردیم و سه واژه نامه که بترتیب شامل هم‌آواها و هم‌نویسه‌ها و هم‌آوا - هم‌نویسه که مستخرج از فرهنگ شش جلدی معین می‌باشند را تهیه نمودیم، تا ضمن پاسخ‌گویی به مسئله‌ی فوق، در پردازش رایانه و تهیه‌ی غلط یاب مورد استفاده قرارگیرند و با کاربردی کردن این محصول در سایر زمینه‌ها بتوانیم خدمتی هر چند اندک در راستای آموزش صحیح خط و زبان فارسی به نو آموزان انجام داده باشیم. در ادامه به پژوهش‌های انجام‌گرفته در مورد اشکالات خط فارسی و چگونگی تهیه‌ی پیکره‌ها خواهیم پرداخت.

### ۱-۱- بیان مسئله

صفوی (۱۳۹۲) هم‌آوا و هم‌نویسه را در یکجا آورده و می‌نویسد:

معمولاً در کتاب‌های معنی‌شناسی به هنگام بحث در باره‌ی روابط مفهومی، بخشی نیز به معرفی هم‌آوا - هم‌نویسه یا تشابه اختصاص می‌یابد که الفاظی متشابه‌اند و اتفاقاً هیچ رابطه‌ی مفهومی‌ای با یکدیگر ندارند. در سنت مطالعات ادبی، هم‌آوا - هم‌نویسه تحت عنوان جناس تام معرفی می‌شود که آمیزه‌ای از جناس خط و جناس لفظ است، جناس خط یا آنچه هم‌نویسه نامیده می‌شود، شرایطی است که دو یا چند واژه، به یک شکل نوشته می‌شود، ولی تلفظ‌های مختلف داشته باشند، مانند «کرم» / karam /، «کرم» / kerem /، «کرم» / korom /، «کرم» / kerm / در این مورد می‌توان «کرم» / karam / به معنی «کر هستم» را نیز مطرح ساخت که با «کرم» به معنای «بخشش» هم‌نویسه به حساب می‌آیند و به دلیل اختلاف در جایگاه تکیه، به یک شکل تلفظ نمی‌شوند (صفوی، ۱۳۹۲: ۱۱۱). هم‌آوایی یا آنچه جناس لفظ نامیده می‌شود، شرایط چند واژه مختلف است که به یک شکل تلفظ می‌شوند و به صورت‌های مختلف نوشته می‌شوند، برای نمونه «خوار» و «خار»، «قضا» و «غذا»، «خویش» و «خیش» و جز آن، هم‌نویسی ناشی از نظام نوشتاری است و هم‌آوایی از قرض‌گیری و تحول آوایی ناشی می‌شود.

در ادامه صفوی (۱۳۹۲) برای توضیح بیشتر در مورد هم‌آوا - هم‌نویسه یا چند معنایی (که ما آن‌ها را در

اینجا هم‌نام می‌نامیم) چنین می‌نویسد:

**هم آوا - هم نویسه (هم نام‌ها)** یا جناس تام، شرایطی است که چند واژه‌ی متفاوت، بدون هیچ ارتباطی میان معانی‌شان، به یک شکل نیز نوشته شوند، به گونه‌ای که انگار یک واژه چند معنای متفاوت دارد. در این باره می‌توان سه واژه‌ی مختلف «شانه» {وسیله‌ی آرایش مو}، «شانه» {کتف} و «شانه» {جای تخم‌مرغ} یا «دوش» {کتف}، «دوش» {وسیله‌ی حمام} و «دوش» {دیشب} را نمونه آورد (صفوی، ۱۳۹۲). اهل زبان هیچ ارتباطی میان مفهوم «دوش» که واژه‌ای قرضی از فرانسه است و ابزاری برای حمام کردن به حساب می‌آید و «دوش» به معنی دیشب نمی‌یابند؛ به همین دلیل، اینها را واژه‌هایی متمایز از یکدیگر می‌دانند. در فرهنگ‌های مختلفی آورده می‌شوند تا مشخص شود که واژه‌های متفاوتی به شمار می‌روند، به عبارت دیگر چند معنایی که (نام دیگر هم‌نام‌ها و هم‌آوا- هم‌نویسه‌ها است) از جمله آشناترین روابط مفهومی‌ای است که در سنت مطالعه‌ی معنی به آن پرداخته‌اند و آن را شرایطی معرفی می‌کنند که یک واحد زبانی از چند معنی برخوردار شود. چند معنایی را می‌توان در سطوح مختلفی چون تکواژ، واژه، گروه و جمله مورد بررسی قرار داد. هر چند در سنت، چند معنایی تنها در سطح واژه اهمیت یافته و تحلیل آن نیز به همین سطح محدود گشته است. به هنگام طرح چند معنایی در سطح واژه، به دونوع چند معنایی هم‌زمانی و در زمانی برمی‌خوریم که یک صورت زبانی می‌تواند یا در یک مقطع زمانی از چند معنی برخوردار باشد و یا در طول زمان معنای متفاوتی بیابد، نمونه‌ای از چند معنایی در زمانی را می‌توان در واژه‌هایی نظیر «گران» {سنگین / پرهزینه} «شوخ» {چرک / بزه‌گو} دانست (صفوی، ۱۳۹۲).

**امکان خطا در املاء:** لازم به ذکر است که با توجه به حروف زبان فارسی تعدادی واژه نیز در این زبان قابل توجه است و علت اصلی آن‌ها وجود حروف هم‌آوا (و نه هم‌نویسه) در نظام نوشتاری فارسی است که گاهی می‌توانند برای نوآموزان در نوشتار و املا خط فارسی تردید و مشکل ایجاد نمایند مانند: دو واژه‌ی «آبد و عبد» و «آتر و عطر» و ... که ما علاوه بر تهیه فهرست واژه‌های هم‌آوا و هم‌نویسه برای اولین بار به تهیه فهرست اینگونه واژه‌ها نیز پرداختیم همچنین برای اولین بار تعدادی واژه را نیز که نه هم‌آوا هستند و نه هم‌نویسه و نه هم‌نام اما در هنگام نوشتار، نوآموزان را دچار تردید می‌کنند (مانند: طَبّ و تَبّ) نیز استخراج کردیم و همه‌ی آن‌ها را «امکان خطا در املاء» نام نهادیم.



در زبان فارسی ۵ همخوان و یک واکه وجود دارد که در برابر آن‌ها، الفبای فارسی ۱۵ حرف به کار می‌برد که عبارتند از:

S/ س - ث - ص / q/ غ - ق  
t/ ت - ط / h/ ه - ح  
Z/ ز - ذ - ض - ظ / ā/ و / a/ و / e/ آ - ع

واج S/ که به آوای حروف «س»، «ث» و «ص» تعلق دارد، همخوان ششی، برونسو، سخت، بیواک، سایشی، دهانی و لثوی است.

واج t/ که به آوای حروف «ت» و «ط» تعلق دارد، همخوان ششی، برونسو، سخت، بیواک، دمیده، انفجاری، دهانی و دندانی می‌باشد.

واج Z/ که به آوای حروف «ز»، «ذ»، «ظ» و «ض» تعلق دارد، همخوان ششی، برونسو، نرم، واکدار، سایشی، دهانی و لثوی است.

واج q/ که به آوای حروف «ق» و «غ» تعلق دارد، همخوان ششی، برونسو، نرم، واکدار، انفجاری، دهانی و ملازی است.

واج h/ که به آوای حروف «ه» و «ح» تعلق دارد همخوان ششی، برونسو، سخت، بیواک، سایشی، دهانی و چاکنایی است.

واج ā/ که به آوای حروف «آ» و «ع» تعلق دارد. واکه‌ی پیشین، باز، گسترده‌ی باز، کوتاه است.

واج e/ که به آوای حرف «ا» و «ع» تعلق دارد. واکه‌ی پیشین، نیم‌بسته، گسترده، کوتاه است.

واج ā/ کشیده، پسین، باز، گرد.

بر همین اساس در خط فارسی اگرچه واج‌ها به یک حالت تلفظ می‌شوند اما به صورت‌های گوناگون تحریر می‌گردند و این عامل باعث تغییر معنایی (مانند حیاط و حیات) می‌شوند. همچنین بدلیل اینکه مصوت‌های کوتاه ( َ ) در خط و نوشتار فارسی نمایش داده نمی‌شوند، در هنگام خوانش نوآموز را دچار سردرگمی می‌-

کنند، لذا اینگونه واژه‌ها باعث مشکلاتی در املاء و آموزش نوشتاری فارسی بویژه برای نوآموزان می‌شود (مانند، طب و تب).

با کمی گشت و گذار در وب، نظرات تارنماهای گوناگون، فیس بوک و دیگر شبکه‌های اجتماعی به این نتیجه می‌رسیم که غلط‌های املائی فاحشی در نوشتار افرادی حتی با مدارک کارشناسی و کارشناسی‌ارشد به چشم می‌خورد و در نتیجه باید پذیرفت که امروزه نوشتن املاء مشکل بزرگ نسل جدید شده است. (اگرچه این نسل از نوشتن در فضای مجازی ترسی ندارند زیرا در این فضا گویا نوشتن املائی غلط جزو خصوصیات منفی افراد محسوب نمی‌شود و در صورتی که شخص احساس کند ممکن است ارزش علمی وی پایین بیاید بلافاصله با تغییر صفحه کلید مطالبش را به صورت فارسی انگلیسی "فنگلیش" می‌نویسد.) اما با این وجود هیچ کس دوست ندارد نوشته‌هایش غلط باشد. بی‌انصافی است اگر تمامی تقصیرهای این مسئله را فقط به گردن نویسندگان بیندازیم، زیرا خط فارسی سرشار از مشکلات است و کمتر کسی را می‌توان پیدا کرد که هنگام نوشتن دچار تردید نشود. در زیر به تعدادی از اشکالات این خط اشاره می‌کنیم.

### اشکالات خط فارسی

اولین مطلب در اغلب کتاب‌های دستور زبان مربوط به حروف الفبا است. شمار حروف الفبای فارسی را ۳۲، ۳۳، ۳۴، ۳۵ و گاهی تا ۳۷ حرف تخمین زده‌اند.

در لغت نامه دهخدا (۱۳۷۳) و فرهنگ فارسی معین (۱۳۷۱) در زیر کلمه الفبا نوشته شده که «الفبا، یا حروف فارسی ۳۳ حرف است» و در صفحه ۱۱۷ در زیر مدخل ابث (abtas) ۲۸ حرف الفبای عربی آورده شده و در ادامه نوشته شده که «ایرانیان در این میان حروف ذیل را افزوده‌اند «پ، چ، ژ، گ»». بدین ترتیب شمار حروف الفبای فارسی ۳۲ حرف فرض شده است. حسن عمید (۱۳۸۰) در فرهنگ عمید حروف زبان فارسی را ۳۲ عدد میدانند. در صورتی که فرهنگستان زبان (۱۳۸۱) نشانه‌های خط فارسی را ۳۴ نشانه ذکر کرده و علاوه بر آن ۹ نشانه ثانوی را نیز تحت عنوان نشانه‌های خط فارسی ذکر نموده است. که کلاً ۴۳ نشانه محسوب می‌شود.

علاوه بر این، به مسائل و مشکلات اساسی خط فارسی می‌توان بطور اجمال چنین اشاره کرد.

- نبودن سه واکه کوتاه ( َ ) در داخل خط فارسی، که در زمان خواندن مطلب می‌بایست سه واکه را حدس زد.

- مشخص نبودن تشدید.

- ناتوانی خط در نشان دادن واکه مرکب (ow) و همچنین عدم توانایی آن برای نشان دادن استرس-های یک گویش خاص.

- وجود حرف‌هایی که عملاً یکی هستند و نقطه در آن‌ها ایجاد تفاوت کرده است، مانند: «ت-پ-ث».
- اشکال نقطه در خط این است که چون با اضافه و کم کردن نقطه حرف مبدل به حرف دیگری می‌شود لذا عملاً اگر بطور تصادفی در نوشتن، نقطه از قلم افتاده باشد، خواننده باید با حدس و گمان محل نقطه و بودن یا نبودن آنرا دریابد. مانند: «بر - بز - پز - یز - تز - یر - تر - پر و ...» تازه این همه حدس در مورد دو حرف است، حال اگر این تعداد به سه یا چهار یا پنج حرف برسد بی‌شک مدت‌ها طول می‌کشد تا به نتیجه برسیم.

- خط فارسی از راست به چپ نوشته می‌شود، در حالی که فرمول‌ها و اعداد در علوم مختلف مثل ریاضی، فیزیک و شیمی از چپ به راست نوشته می‌شوند و مخلوط این دو، امکان پذیر نیست و در زمان نوشتن مشکلاتی را ایجاد می‌کنند.

- اکثر حروف در اول، وسط و آخر به یک شکل نوشته نمی‌شوند مانند: (ع - ع - ع - ع) (ی - ی - ی - ی) توجه داشته باشید که اگر حروف فارسی را با قرار گرفتن آن‌ها در جاهای مختلف کلمه حساب کنیم به بیش از ۱۲۵ شکل می‌رسد.

- ناتوانی خط فارسی در جداسازی واژه‌ها.

- کمبود منابع الکترونیکی در حوزه NLP برای فارسی.

حال، وجود این همه مشکل در خط و زبان فارسی و سردرگمی در هنگام آموزش آن، ما را بر آن داشت، تا بتوانیم هر چند اندک در کمک به رفع بنیادی این مشکلات با تهیه‌ی پیکره‌ی فرهنگ املائی واژه‌های هم‌آوا و هم‌نویسه اقدام نماییم، و با این کار، راه را برای محققین و پژوهشگران جهت دستیابی به پیکره‌ها هموار نماییم.

## فصل دوم: پیشینه تحقیق و مبانی نظری

### ۲- پیشینه تحقیق

اصطلاح پیکره را به ویژه زبان‌شناسان ساختگرا به کار می‌بردند و همواره تاکید می‌کردند که توصیف یک زبان یا گویش باید مبتنی بر داده‌های گردآوری شده و تحلیل این داده‌ها باشد و با آنکه فراگیری و بزرگی پیکره عامل تعیین‌کننده‌ای در افزایش دقت و اعتبار نتایج به شمار می‌آمد، محدودیت‌های نیروی انسانی و زمان در برابر این عامل قرار داشت و در نتیجه، بیشتر به نمونه‌های برگزیده‌ای اکتفا می‌شد. تنها از دهه ۱۹۶۰ بود که توانستند با کمک رایانه، مجموعه‌های بسیار بزرگی از متن‌های معتبر زبانی فراهم نمایند و آغازگر همه این‌ها، پیکره زبانی براون بود که نمونه‌ای معیار از زبان انگلیسی امروز برای استفاده در رایانه‌های رقمی به شمار می‌آید (Svartvik: 1992: 7). این پیکره را نلسون فرانسیس و هنری کوچرا (۱۹۶۴) در دانشگاه براون با گردآوری نزدیک به یک میلیون واژه از پانصد نمونه متنی پیوسته ایجاد کردند و کوشیدند انواع مقاله‌ها و گونه‌های زبانی را پوشش دهند. کار از ۱۹۶۱ آغاز شد و در سال ۱۹۶۴ به پایان رسید. این مدت در مقایسه با کارهای گذشته و با توجه به حجم نسبتاً زیاد داده‌ها در آن زمان بسیار کم به نظر می‌رسید (Francis and Kucera: 1964).

یکی از مهم‌ترین پیکره‌هایی که در همین دهه، اما بدون بهره‌گیری از رایانه فراهم آمد و از این نظر آن را در شمار پیکره‌های پیش از دوران الکترونیک قرار می‌دهند، پیکره بررسی کاربرد زبان انگلیسی با seu است که راندلف کرک از ۱۹۵۹ در دانشگاه لندن آغاز کرد و گردآوری و آماده‌سازی آن تا سال ۱۹۸۹ طول کشید، چرا که برای یک میلیون واژه‌ی آن، یک میلیون برگه با هفده سطر از متن‌های مربوط به سال‌های ۱۹۵۳ تا ۱۹۸۷ با دست نوشته شد. البته در کنار این پیکره‌ی نوشتاری، بخش گفتاری آن با عنوان پیکره لندن - لوند (بررسی انگلیسی گفتاری ...) به وسیله "یان سوارتویک"، در دانشگاه لوند سوئد از سال ۱۹۷۵ آغاز شد. از آنجا که این بخش با کمک رایانه فراهم شد، از همتای نوشتاری خود پیشی گرفت و با حدود نیم میلیون واژه در سال ۱۹۸۰ ارائه شد (Kennedy: 1998: 32).

پیکره زبانی می‌تواند بسیار بزرگ، فراگیر و نماینده تمامی یک زبان و یا گونه‌ای از آن باشد، به شکل برگه‌های یادداشت یا پرونده‌های رایانه‌ای شامل متن‌های کامل یا گزیده‌هایی از آن‌ها، و یا بخش‌های پیوسته‌ای از

متون یا گزیده‌ای از نقل قول‌ها و نکات و حتی فهرست‌های واژگانی پیکره می‌تواند، ویژه بررسی خاصی فراهم شود و یا در برگیرنده مجموعه‌ی عظیم و بی‌ساختاری از متون گوناگون باشد که برای منظوره‌ای گوناگون به کار رود.

امروزه به دلیل نیازهای پژوهشی جامعه علمی در ایران، سازمان‌های مختلف نسبت به تهیه پیکره‌های مختلف اقدام نمودند که از آن جمله می‌توان به موارد زیر اشاره نمود:

**پیکره‌ی به‌روزشونده‌ی پارسی**، این پیکره، یک پیکره‌ی به‌روزشونده (monitor) متون خبری در زبان فارسی است که هر روز متون جدید از خبرگزاری‌های معتبر به صورت خودکار در شش دسته‌ی موضوعی «سیاسی»، «اقتصادی»، «اجتماعی»، «فرهنگی»، «علمی» و «ورزشی» به آن اضافه می‌شود (در آینده متون غیرخبری نیز به پیکره اضافه خواهند شد). پیکره حاضر از این حیث که تنوعات زبانی را با انتخاب داده از ژانرهای مختلف پوشش می‌دهد،

**پیکره‌ی بی‌جن‌خان** مجموعه‌ای است از متون فارسی شامل بیش از ۲ میلیون و ۶۰۰ هزار کلمه که با ۵۵۰ نوع برچسب POS برچسب‌گذاری شده‌اند. این پیکره که در پژوهشکده‌ی پردازش هوشمند علائم تهیه شده است همچنین شامل بیش از ۴۳۰۰ تگ موضوعی چون سیاسی، تاریخی و ... برای متون است. نسخه‌ی کامل پیکره بی‌جن‌خان با عنوان «پیکره‌ی متنی زبان فارسی».

**فرهنگ بسامدی بر اساس پیکره متنی زبان فارسی امروز که توسط محمود بی‌جن خان و مهدی محسنی تهیه شده است**، این فرهنگ بسامدی میزان کاربرد کلمات زبان را بر حسب فراوانی وقوع آن‌ها در یک پیکره زبانی که نماینده آماری زبان است نشان می‌دهد. کلمات پرسامد و کم‌بسامد می‌توانند شاخص ارزشمندی برای پی بردن به مولفه‌های فرهنگی یک جامعه زبانی باشند. میزان بسامد کلمات در یک فرهنگ بسامدی تابع بازه زمانی است که پیکره فرهنگ در آن زمان گردآوری شده است. بنابراین اطلاعات موجود در فرهنگ بسامدی رابطه مستقیم با تحولات تاریخی یک جامعه زبانی دارد. امروزه فرهنگ بسامدی در تحقیقات بین‌رشته‌ای مانند آموزش زبان دوم به خارجی‌ان، زبان‌شناسی نظری و کاربردی و فناوری‌های زبان در محیط رایانه از قبیل ترجمه ماشینی و طراحی مدل زبانی برای سامانه‌های پردازش گفتار، کاربرد موثری دارد.

پایگاه داده های زبان فارسی به منظور ایجاد مجموعه ای بزرگ از پیکره های گوناگون زبان فارسی امروز توسط پژوهشگاه علوم انسانی و مطالعات فرهنگی طراحی شده است و در آن متن های برگزیده ادبی، علمی، هنری، سیاسی و مانند این ها از گونه های نوشتاری و گفتاری فارسی وجود دارد. این سامانه امکان جستجوی واژه ها، ترکیب ها، باهمایی ها و بررسی بسامد آن ها را به همراه گزارش های آماری متنوعی از متون فراهم می کند. افزون بر اصل متن ها و واژه های به کاررفته در آن ها، معنی، مقوله دستوری، آوانگاشت و ریشه یا (بن واژه) بسیاری از واژه ها نیز در پایگاه وجود دارد.

کاربران قبل از معرفی شدن به سیستم می توانند از امکاناتی که به صورت پیش فرض برای آنان در نظر گرفته شده است استفاده کنند که شامل مشاهده و گزارش گیری از چند اثر انتخاب شده از سوی مدیر وبگاه می باشد. چنانچه کاربر بخواهد متن های بیشتری را مشاهده کند باید در قسمت معرفی کاربر، خود را به سیستم معرفی کند. در معرفی کاربر ۳ فیلد اطلاعاتی نام کاربری، گذرواژه و نام، از موارد الزامی هستند.

یکی از مهم ترین گام ها در طراحی و ساخت سیستم های بازشناسی گفتار پیوسته با دایره لغات بزرگ (LVCSR)، تعلیم مدل زبانی (Language Model) برای مدل کردن نحوه توالی کلمات در یک زبان خاص است. برای انجام چنین کاری، داشتن مقادیر بسیار زیادی از دادگان متنی (Text) از آن زبان همراه با برچسب های لازم ضروری است. از طرفی در بسیاری از پروژه های تحقیقاتی در زمینه پردازش زبان طبیعی (Natural Language Processing)، چنین دادگانی مورد نیاز است. پیکره متنی زبان فارسی، دادگانی است که توسط پژوهشگاه توسعه فناوری های پیشرفته جواب گوی این گونه نیازها در شناسایی گفتار فارسی و پردازش زبان طبیعی است.

**پایگاه انتشار و تولید داده های زبانی پیکره گان؛ این پیکره ی متنی زبان فارسی، که توسط پژوهشکده**

پردازش هوشمند علائم تهیه و منتشر شده است، مجموعه ای از متون نوشتاری و گفتاری رسمی زبان فارسی است که از منابع واقعی همچون روزنامه ها، سایت ها و مستندات از قبل تایپ شده، جمع آوری شده، تصحیح گردیده و برچسب خورده است. حجم این دادگان حدوداً ۱۰۰ میلیون کلمه است و از منابع مختلف تهیه گردیده و دارای

تنوع بسیار زیادی است. ۱۰ میلیون کلمه از این پیکره با استفاده از ۸۸۲ برجسب نحوی-معنایی به صورت دستی توسط دانشجویان رشته‌ی زبان‌شناسی برجسب‌دهی شده‌اند و هر پرونده بر حسب موضوع و منبع آن طبقه‌بندی شده است. این پیکره که توسط پژوهشکده پردازش هوشمند علائم تهیه شده است، برای استفاده در آموزش مدل زبانی و سایر پروژه‌های مربوط به پردازش زبان طبیعی مناسب است.

## ۱-۲- مبانی نظری

یکی از راه‌های توصیف و تحلیل معنا، بررسی روابط مفهومی واژه‌ها و مقایسه‌ی آن‌ها با یکدیگر است، در این میان «هم‌ها» که هم‌رده‌ی تضادمعنایی و شمول‌معنایی و چندمعنایی هستند، زیر مجموعه‌ی روابط معنایی واژه‌ها محسوب می‌گردند و دسته‌ی واژگانی مهمی را تشکیل می‌دهند.

### هم‌ها

سه اصطلاح «هم‌آوا»، «هم‌نویسه» و «هم‌معنا» که به «هم‌ها» (Homos) معروف هستند، برای توصیف روابط بین کلمه‌های زبان به کار می‌روند، توضیح کامل هر کدام در زیر آمده است.

۱- هم‌آوا (Homophony) یعنی واژه‌هایی که شبیه هم تلفظ می‌شوند اما بطور متفاوت نوشته می‌شوند و اغلب معانی متفاوت نسبت به هم دارند. مثلاً واژه‌های فارسی «خوار - خار؛ خواست - خاست؛ فترت - فطرت؛ ناضر - ناظر» دارای نوشتار و معنای متفاوتی هستند در حالیکه تلفظ یکسانی دارند. همچنین در فارسی کلمات «تیر - اسیر - عصیر»، «صور (شیپور) - سور»، «قضا - غذا - غزا» و «عرض - ارز»، واژه‌های هم‌آوا محسوب می‌شوند. در دستور سنتی به اینگونه واژه‌ها جناس لفظی یا لفظی می‌گویند.

۲- هم‌نام: هم‌نام‌ها (Homonymy) واژه‌هایی هستند که به یک صورت نوشته می‌شوند و به یک صورت نیز تلفظ می‌شوند اما معنای متفاوتی دارند (لانگمن و اشمیت، ۲۰۰۲: ۲۴۱) در سنت مطالعات ادبی، هم‌نامی تحت عنوان جناس تام معرفی می‌شود که ترکیبی از جناس خط و جناس لفظ است. هم‌نامی یا جناس تام، شرایطی است که چند واژه‌ی متفاوت، بدون هیچ ارتباطی میان معانی‌شان، به یک شکل تلفظ شده و به یک شکل نیز نوشته می‌شوند، گویا اینکه آن یک واژه، چند معنای کاملاً متفاوت دارد. در این مورد می‌توان سه واژه‌ی مختلفِ شانه (وسیله‌ی آرایش مو)، شانه (کتف) و شانه (جای تخم‌مرغ) و همچنین شیر (لبنی) و شیر

(حیوان) و شیر (آب) یا دوش (وسیله حمام) و دوش (دیشب) را ذکر نمود. اهل زبان ارتباطی میان مفهوم دوش که واژه‌ای قرضی از زبان فرانسه است و ابزاری برای حمام کردن محسوب می‌شود و دوش به معنای (دیشب) نمی‌یابند؛ به همین دلیل، آن‌ها را واژه‌هایی متمایز از یکدیگر می‌دانیم. در فرهنگ لغت‌ها، این واژه‌ها تحت مدخل‌های مجزایی آورده می‌شوند تا مشخص شود که واژه‌های متفاوتی به شمار می‌روند (صفوی، ۱۳۷۹: ۱۱۰). اغلب تشخیص بین کلمات هم‌نام و چند معنا (polysomy) واضح نیست زیرا واژه‌های چندمعنایی، واژه‌هایی هستند با یک صورت نوشتاری، چندین معنای مرتبط و نیز ریشه مشترک دارند. مثلاً واژه head به بالایی بدن (سر) یا بالاترین مقام یک شرکت یا واحد اطلاق می‌شود که از این لحاظ به دلیل فرم نوشتاری یکسان با واژه‌های «هم‌نام» اشتباه می‌شوند. با این حال یکی از راه‌های تمیز بین آن‌ها، به اعتقاد یول (۲۰۰۶)، رجوع به مدخل خاص هر کلمه در فرهنگ لغت است به این معنا که اگر کلمه‌ای دارای چند معنای مرتبط به هم باشد فقط یک مدخل به آن تعلق می‌گیرد که زیر آن مدخل فهرستی از معانی مختلف کلمه آورده می‌شود ولی چنانچه دو کلمه هم‌نام باشند مدخل جداگانه‌ای به هر یک از آن‌ها اختصاص داده می‌شود. مثلاً واژه‌های فارسی شیر، شانه، دوش نمونه‌هایی از کلمات هم‌نام هستند که هر یک از معانی مختلف آن‌ها مدخلی جداگانه دارد. به عبارت دیگر کلمات هم‌معنا هر یک با توجه به داشتن معانی مختلف و غیر مرتبط، مدخل‌های جداگانه‌ای دارند. در حالیکه در کلمات چندمعنا یک مدخل به هر یک اختصاص پیدا می‌کند و معنای مختلف مرتبط به هم در همان مدخل وارد می‌شوند. نتیجه آن که در مورد واژه‌های چندمعنا، سنت غالب آن است که واژه‌ها، زیر یک مدخل درج و معانی مختلف، با شماره متمایز می‌شوند، زیرا بین معانی مختلف نوعی پیوند معنایی مشاهده می‌شود. تامس (۱۹۹۵) معتقد است که کلمات چندمعنا و هم‌نام برای گویشوران غیربومی ایجاد اشکال می‌کند زیرا آن‌ها به دانش زمینه‌ای فرهنگی که سخنوران بومی آن زبان بر آن تسلط دارند، آشنا نیستند. تا آنجا که، نمی‌دانند کدام معنا را در بافت خاص خود بکار ببرند. حال آنکه سخنور بومی زبان حتی اگر معنای یک کلمه را هم نداند با مراجعه به فرهنگ آن زبان به راحتی می‌تواند معنای صحیح بافت مورد نظر خود را بیابد (رحیمیان و رزمجو، ۱۳۸۹: ۱۸۰). بنابراین همانگونه که قبلاً نیز ذکر شد، هم‌نام‌ها واژه‌هایی هستند که شکل آوایی و نوشتاری یکسان، اما معانی مختلف دارند، مانند واژه «شیر» (= شیر آب، شیر جنگل، شیر لبنی). در هم‌نامی،



یکسانی، هم در صورت آوایی و هم در صورت نوشتاری وجود دارد. باید توجه داشت که «هم آوایی» و «هم-نویسی» مواردی از «هم نامی ناقص» هستند

۳- هم نویسه: واژه‌های هم نویسه (Homography) واژه‌هایی هستند که به صورت یکسان نوشته شده اما بطور متفاوت تلفظ می‌شوند و معانی گوناگونی هم دارند (لانگمن و پلات ۱۹۹۲ و لانگمن و اشمیت، ۲۰۰۲: ۲۴۱). واژه‌های فارسی، مانند: شکر، شکر؛ جَنگ، جَنگ؛ مَهر، مَهر، مَهر واژه‌هایی هم نویسه هستند. صفوی (۱۳۷۹: ۱۱۱) معتقد است که هم نویسه همان جناس خط است. برای نمونه واژه‌های کَرَم، کَرَم، کَرَم، کَرَم. در این مورد می‌توان کَرَم به معنای گر هستم را نیز مطرح نمود که با کَرَم به معنای بخشش هم نویسه محسوب می‌شوند ولی به دلیل تفاوت در جایگاه تکیه، به یک شکل تلفظ نمی‌شوند. در سنت مطالعات ادبی، هم نویسه‌ای تحت عنوان جناس ناقص معرفی می‌شود که شامل واژه‌هایی است که فقط هم شکل می‌باشند اما از نظر تلفظ متفاوت هستند. مانند: گزیدن - گزیدن، خَلق - خَلق و قَمَری - قَمَری. در سنت ادبی گاهی جناس مرکب را نیز ذکر می‌کنند و آن بدین صورت است که کلمات مرکب شبیه کلمات دیگر می‌شوند، مانند: کَمند و کَمند (کم هستند) یا شهریار و شهریار (همایی، ۱۳۸۳).

۴- خطاهای املایی (هم واجی) گاهی ممکن است نوآموزان و کودکان در هنگام نوشتن املاء واج‌های هم‌آوا را بجای یکدیگر بکار ببرند. مانند: «تَب» بجای «طَب» یا «صُور» (ج صورت) بجای «سور» (جشن)، که اینگونه واژه‌ها نه جزو هم‌آواها هستند و نه جزو هم‌معناها و نه هم‌نویسه‌ها، اینگونه خطاها فقط در واژه‌های نوآموزان زبانهایی امکان دارد مشاهده شود که دارای واج‌های هم‌آوا هستند. ما اینگونه واژه‌ها را «امکان خطا در املاء» یا «هم واجی» می‌نامیم.

#### ۱-۱-۲- انواع پیکره

همان گونه که پیش‌تر اشاره شد، پیکره‌های زبانی می‌توانند عام و گسترده بوده، و یا نمونه تصادفی از یک زبان و یا تخصص و نماینده یک گونه خاص زبان باشند. پیکره‌ها را می‌توان برای منظوره‌های گوناگون به کار گرفت، از جمله برای فرهنگ‌نگاری، معناشناسی، بررسی‌های دستوری، آموزش زبان و مانند اینها، پیکره‌ها را

می‌توان از نظر اندازه و گستره به دسته‌های محدود، متوسط و عظیم تقسیم کرد. بر اساس رویکرد «فرم و کارکرد» نیز، می‌توان انواع زیر را برای پیکره‌ها در نظر گرفت:

- پیکره اشباع شده (saturated corpus) که با اشاره به عوامل و متغیرهای معین با اطلاعات اضافی پر شده است. اشباع، ویژگی آماری است که با نسبت غنای واژگانی (موارد کاربردی سرواژه یا...) ارتباط دارد. در این مورد، کل متن به پاره‌های مشابه بخش شده است.

- پیکره موقعیتی (situation corpus)، که بنا بر نیازهای خاص یک بررسی، فراهم می‌شود.

- پیکره پایشگر (monitor corpus)، که دارای ابعاد تاریخی است و برای ثبت تغییرات زبانی به کار می‌رود و از برهه‌های گوناگون زبانی، نمونه‌های مشابه را دارد.

- پیکره تطبیقی (comparative corpus)، پیکره‌ای چند زبانه با نمونه‌های مشابه از هر زبان است.

- پیکره موازی (parallel corpus)، پیکره‌ای چند زبانه با متنی اصلی از یک زبان و ترجمه آن در زبان‌های دیگر (یا زبان‌های هدف) است.

- پیکره دوسویه (two sided corpus)، پیکره‌ای چند زبانه با متن اصلی به همه زبان‌ها و ترجمه آن‌ها به همه زبان‌ها است (Teubert؛ 1999).

در آغاز دهه ۱۹۸۰ شاید می‌توانستیم پیکره‌های الکترونیکی را با انگشتان یک دست بشماریم. در آغاز دهه ۱۹۹۰ تعداد طرح‌هایی که برای ایجاد پیکره‌های زبانی فعال بودند به ده‌ها می‌رسید و چند سال بعد در کشور-های مختلف جهان، فعالیت‌های بسیاری برای زبان‌های گوناگون آغاز شده بود. اکنون پس از گذشت تنها دو دهه، با جستجوی ساده در شبکه جهانی رایانه‌ها (اینترنت) بیش از سیصد پیکره بزرگ و کوچک برای بسیاری از زبان‌های جهان فهرست میشوند.

علت این گسترش شتابنده در نخستین وهله، راه یافتن رایانه‌ها در این حوزه و تغییرات بنیادی در روش-های پردازش متن و ذخیره سازی و دستیابی به آن است. ((بدین سان، زبان شناسی پیکره‌ای اکنون ناگزیر با

رایانه پیوند یافته است و همین وسیله است که سرعت شگفت انگیز، شمار پذیری کامل، تکرار و روبرداری دقیق، صحت اماری و امکان به کارگیری حجم عظیم داده‌ها را به همراه آورده است)) (Kennedy:1998).

برای زبان شناسان قدیمی‌تر، اصطلاح «زبان شناسی پیکره‌ای» اغلب یادآور دوران رونق پیکره‌زبانی یعنی دهه ۱۹۵۰ است، دوران هریس، فریس، هیل و دیگر ساختگرایان آمریکایی. اما از آن دوران به بعد است که این اصطلاح، کاربرد بیشتری یافته و به ویژه در اوایل دهه ۱۹۹۰ است که به عنوان شاخه‌ای مشخص از زبان‌شناسی، مانند زبان‌شناسی اجتماعی، روان‌شناسی‌زبان و زبان‌شناسی‌متنی که در دهه‌های ۱۹۶۰ و ۱۹۷۰ شکل گرفته‌اند، شناخته شده است (Leech:1992). اما زبان‌شناسی پیکره‌ای با رشته‌های یاد شده تفاوت دارد، زیرا نماینده یک حوزه مشخص از مطالعات زبانی نیست بلکه بنیادی روش شناختی برای پژوهش‌های زبانی به شمار می‌آید. در اصل و عملاً زبان‌شناسی پیکره‌ای، به آسانی با شاخه‌های دیگر زبان‌شناسی می‌آمیزد. می‌توان با کمک پیکره، به بررسی‌های آوایی، نحوی و اجتماعی یا دیگر زمینه‌های زبان پرداخت و در این صورت می‌گوییم که روش‌ها و فنون زبان‌شناسی پیکره‌ای را با موضوعات آوایی، نحوی و اجتماعی زبان و مانند آن آمیخته‌ایم (Leech:1992).

از آنجا که حجم مواد زبانی (در اغلب زبان‌های امروز جهان) برای گنجاندن در پیکره بسیار زیاد است. در برخی موارد می‌توان پیکره را به صورت نمونه ایجاد کرد. تعیین حدود و گستره پیکره، وظیفه زبان‌شناسان است. برخی از متغیرهایی که در تعیین حدود دخالت دارند، عبارتند از: زبان، بُرش یا محدوده‌ای از زمان، منطقه، شرایط و موقعیت، ویژگی‌های بیرونی و درونی متنی و ....

از این‌رو ما به تهیه پیکره املائی واژه‌های «هم‌آوا» و «هم‌نویسه» در زبان فارسی پرداختیم.

صادقی و زندی مقدم (۱۳۸۵) نسخهٔ رقومی فرهنگ‌املائی خط‌فارسی براساس دستور خط‌فارسی، مصوب فرهنگستان زبان و ادب فارسی را تهیه نمودند. این پژوهشگران در این مورد می‌نویسند:

از آنجایی که به خاطر سپردن تمام قواعدی که در دستور خط فارسی مصوب فرهنگستان زبان و ادب فارسی آمده، برای همگان امکان پذیر نیست، و از سوی دیگر، عموم نویسندگان، اعم از محققان، نثر نویسندگان، شاعران، ویراستاران، روزنامه نگاران، منشیان و غیره، هنگام نوشتن کلمات مرکب و مشتق و بعضی کلمات دو

املائی یا سه املائی دیگر دچار تردید می‌شوند، فرهنگستان راه چاره را در این دید که دست به انتشار یک فرهنگ املائی مبتنی بر قواعد دستور خط فارسی بزند و املائی کلماتی را که در این دستور به عهده سلیقه و ذوق نویسندگان گذاشته بود. تا آنجا که امکان دارد، بیشتر قاعده‌مند سازد، یا شیوه مرجع خود را (در مواردی که کلمه دو املائی است) در آن بگنجانند. از این رو، مقرر شد که گروه دستور زبان فارسی مجموعه‌ای از کلمات رایج فارسی امروز را با قواعد دستور خط فارسی منطبق و منتشر کند (صادقی و زندی مقدم؛ ۱۳۸۵).

این اثر با وجود اینکه بسیار سودمند است، اما چون به صورت کتاب و PDF است، لذا امکان استفاده از آن محدود می‌باشد.

سمیعی (۱۳۶۷) معتقد است؛ برخی از دشواری‌ها نوعاً مربوط می‌شوند به عناصری از زبان که صورت لفظی آن‌ها به عینه در زبان زنده ادبی باقی مانده، لیکن معنایی که در متون از آن‌ها مراد گرفته شده، منسوخ یا مهجور گردیده یا در عرف زبان کاربرد کمتری دارد.

این تشابهاتِ صوری را می‌توان به انواع زیر تقسیم کرد:

**الف** - تشابه در کتابت و تلفظ (مانند «افسوس» که در متون کهن به معنی استهزاء نیز به کار رفته است). ممکن است دو متشابه؛ دو یا چند لغت جداگانه باشند (مانند «بید» به معنی «باشید» و به معنی «درخت بید» و به معنی «نوعی حشره که به فرش آسیب می‌رساند»).

**ب** - تشابه در کتابت و فرق در تلفظ به جهت فرق در حرکات یا وجود «واومعدوله» و «واو» و «یای-مجهول» (مانند: «شسته: شسته، شسته»؛ «پول: به «واو معلوم» و به «واو مجهول» که لغتی است در معنی «پُل»؛ «گو: فعل امر از گفتن به ضم کاف فارسی»، «گو به معنی گودال»).

**ج** - تشابه در کتابت و فرق در تلفظ به جهت فرق در تکیه (مانند، انگشتی: انگشتر + یای نکره، با تکیه در هجای اول؛ انگشتی با یای جزء کلمه و با تکیه در هجای آخر).

**د** - تشابه در کتابت و فرق در تلفظ به جهت فرق در مکث بالقوه و تقطیع تکواژی (اوزار: او/ زار؛ اوزار به معنی افزار).

**ه** - تشابه در کتابت و فرق در تلفظ به جهت فرق در دو یا چند عامل سابق الذکر. فرق معنایی این کلمات

و عبارات متشابه نیز علل گوناگون دارد، از جمله:

۱- منسوخ یا مهجور شدن معنای مراد («خانه» به معنی «اتاق»).

۲- مهجوری معنای اصطلاحی («رخصت» در مقابل «عزیمت»، در اصطلاح فقهی؛ «حفاظ» در اصطلاح اهل فتوت). - فرق در تقطیع تکواژی و ساخت واژه‌ای («دستوری» به معنی «اجاره» و «خدمتی» به معنی «تحفه» و «بربند» به معنی «سینه بند» در مقابل «دستوری» و «خدمتی» با بای نکره و «بربند» فعل امر از فعل مرکب «بربستن»).

-مهجوری قاعده دستوری به کاررفته (حذف شناسه فعلی به قرینه).

- اختلاف در ریشه («اهداء»: از «هدی» به معنی هدیه فرستادن، از «هدء» به معنی آرام دادن).

- اختلاف در تعلق زبانی (دمدمه: بر وزن فعلله در عربی به معنی خشم گرفتن، هلاک کردن؛ در فارسی به معنی مکر و فریب و افسون). در تشابه صوری و فرق معنایی عناصر گویشی و لهجه‌ای و تحول آوایی نیز، همواره دخیل است. سیاهه این تشابهات برای استفاده دانشجویان و علاقه‌مندان به متون فارسی مفید خواهد بود (سمیعی، ۱۳۶۷).

اگرچه تحقیقات زیادی در این حوزه انجام پذیرفته است، اما هیچکدام از آن‌ها به صورت عمومی منتشر نشده‌اند، و تنها آگاهی ما از این محصولات همانا مقالات این پژوهشگران ارجمند می‌باشد.

هدف این پژوهش تهیه فرهنگ املائی واژه‌های هم‌آوا و واژه‌های هم‌نویسه برای زبان فارسی است. طبق پژوهش‌های انجام شده اخیر، پژوهشگران (کاشفی و همکاران (۲۰۱۴)، پیله ور و همکاران (۲۰۱۲)، دستغیب و همکاران (۲۰۱۶)) به این نتیجه رسیده‌اند که یکی از بزرگترین دسته خطاها، آن دسته از خطاهایی هستند که کاربر، دانش تصحیح آنرا ندارد، این دسته همانا دسته «خطاهای هم‌آوا» هستند.

دستغیب (۱۳۹۵) می‌نویسد: غالب روش‌های تصحیح خطا به روش تصحیح واژه‌ها به صورت منفرد تاکید دارد و تصحیح واژه را بدون در نظر گرفتن اطلاعات زبانشناسی و همچنین متنی که واژه در آن قرار گرفته است انجام می‌دهد. این نوع تصحیح خطا نمی‌تواند کسری از خطاها را (که شامل خطاهای آوایی (Phonetic)، شناختی (cognitive)، گرامری و چاپی (typographic) هستند) شناسایی و تصحیح نماید. خطاهای آوایی

معمولاً در استفاده‌ی نادرست از حروف هم‌آوا پدید می‌آید و خطای شناختی نیز به دو دلیل پدید می‌آید، نخست در اثر عدم تسلط کاربر به زبان پدید می‌آید که معمولاً کاربر معنا و املا صحیح واژه را نمی‌داند، دوم بدلیل بیماری یا اختلالات شناختی است. خطای تایپی نیز زمانی رخ می‌دهد که کاربر املا صحیح لغت را می‌داند، ولی در زمان تایپ واژه دچار مشکل شده و حرف کنار حرف مورد نظر را فشار می‌دهد و در نتیجه واژه را غلط تایپ می‌کند. خطای شناختی به دلیل عدم آگاهی کاربر از هجی و یا املا صحیح واژه نشأت می‌گیرد (مانند خواندن که می‌شود خاندن)، یا شخص هنگام نوشتن به موضوع دیگری می‌اندیشد و در نوشته‌ی خود کلمه‌ای از آن موضوع را به متن می‌افزاید. خطاهای هم‌آوایی که حدوداً ۴۴٪ از خطاها را تشکیل می‌دهند زیر شاخه‌ای از خطاهای شناختی محسوب می‌شوند و بدین صورت است که یک حرف با حرف دیگر هم‌صدا یا هم-آوا جایگزین می‌شود (مانند غلطک که می‌شود گلتک). در بسیاری از موارد تشخیص و تصحیح خطا بدون در نظر گرفتن متن امکان پذیر نمی‌باشد. استفاده از متن واژه‌ها در شناسایی و تصحیح خطا علاوه بر امکان شناسایی و تصحیح خطاهای معنایی (Real word error) در تصحیح واژه‌های مرکب و ... نیز کاربرد دارد.

بنا براین در این پژوهش گنجینه‌ای از واژه‌های هم‌آوا و واژه‌های هم‌نویسه گردآوری می‌شوند، تا بتوانند زمینه‌ای برای سایر پژوهش‌های حوزه پردازش زبان طبیعی باشند. در ضمن بسیاری از کاربردهای حوزه‌ی NLP مانند غلطیاب‌ها به چنین منابعی نیاز دارند بنابراین، این گنجینه لغت املائی می‌تواند در راستای تولید منابع برای زبان فارسی مفید باشد.

## ۲-۲- ضرورت انجام طرح

ضرورت اجرای این طرح بیشتر و در درجه‌ی اول در بازبایی اطلاعات در مراکز اطلاع رسانی الکترونیکی و رایانه‌ها در اسناد و مدارک به زبان فارسی و قابل رویت است، زیرا اگر کاربر در هنگام جستجو نتواند املا صحیح و واژه‌ای را تایپ نماید، بی شک اسناد بازبایی شده بسیار متفاوت‌تر از نتایج مورد انتظار خواهد بود و نتایج نهایی این پژوهش می‌تواند کمک شایانی برای رفع اینگونه نواقص در بر داشته باشد. همچنین از نتایج این پژوهش می‌توان در آموزش زبان فارسی به غیر فارسی زبانان و نوآموزان زبان فارسی و ساخت و برنامه نویسی غلطیاب‌های املائی و تصحیح متون و تهیه‌ی ویراستارهای الکترونیکی سود برد. همچنین با توجه به

کمیود منابع زبانی برای زبان فارسی، این پیکره‌ها می‌توانند در مواردی مانند: غلط‌یاب‌ها، شناسایی فوری نویسه‌ها، تبدیل گفتار به متن و .... مورد استفاده قرار گیرند.

باید توجه داشته باشیم که فعالیت‌های دیگر این حوزه مانند "فرهنگ املائی خط فارسی" (صادقی و دیگران؛ ۱۳۸۵)، علیرغم ارزش والائی که دارند، به صورت کتاب ویا فایل pdf تهیه شده‌اند و مستقیماً قابل استفاده به صورت الکترونیکی نیستند.

باید دقت داشت که خطا یا غلط نویسی در هم‌آواها معمولاً زمانی اتفاق می‌افتد که کاربر یا نویسنده معنای واژه‌ای را می‌داند ولی شکل لغوی واژه و املای درست آنرا نمی‌داند. بنابراین، فرهنگ حاضر می‌تواند در این راستا به کاربران کمک کند تا نرم افزارهای کاربردی، از این پیکره به عنوان یک منبع برای تهیه کاربرد غلط‌یابی استفاده نمایند.

اکثر سامانه‌هایی که محتوای متنی تولید می‌کنند می‌بایست دارای اجزای ویراستاری و بررسی لغوی متنی باشند، به همین علت نتیجه این پژوهش می‌تواند برای سامانه‌ها مورد استفاده قرار گیرد.

خط و زبان فارسی علاوه بر مشکلات تایپی و نگارشی دستی، مشکلات متعددی از نظر نگارش در رایانه نیز دارند. برای تهیه سامانه‌های محتوایی و واژه‌نامه‌های مورد نیاز اینگونه سامانه‌ها در کنار سایر منابع، نیاز است ابزارهای پردازش نیز برای زبان فارسی آماده گردد تا بتوان از آن به نحو مناسب در پژوهش‌های گوناگون استفاده نمود.

بدیهی است بکارگیری ابزارهای استاندارد موجب افزایش کیفیت و بازدهی پژوهش‌های مختلف خواهد بود. در حال حاضر عدم دسترسی به منابع استاندارد زبانی برای پژوهش‌های مختلف در حوزه زبان فارسی، موجب شده تا پژوهش‌های این حوزه هم با مشکلات متعدد روبرو شود و هم از کیفیت مناسب برخوردار نباشد. لذا به نظر می‌رسد ساخت و تولید منابع و ابزارهای مناسب و با کیفیت، برای زبان فارسی بسیار مهم و از جمله پیش-نیازها و فعالیت‌های بنیادی این حوزه باشد. از مهمترین دلایل استخراج این پیکره‌ها می‌توان به نیاز مبرم نرم-افزارهای فارسی از جمله غلط‌یاب‌ها به یک پیکره کامل از یک منبع موثق و قابل استناد اشاره نمود.

### ۲-۳- اهداف پژوهش

هدف از این پژوهش تولید و تهیه واژه‌نامه‌هایی است که متشکل از فرهنگ‌های صرفاً «هم‌آوا» (غیرهم‌معنا و غیرهم‌نویسه) (مانند سَفَر/ صِفِر) و صرفاً «هم‌نویسه» (غیرهم‌معنا) می‌باشند (مانند: آبکار/ اِبکار)، اما ما در فرایند این پژوهش سعی نمودیم تا بتوانیم واژه‌نامه‌های دیگری نیز که جزو اهداف اولیه‌ی این طرح نبودند را نیز از این فرایند استخراج کنیم، و بر این اساس توانستیم دو واژه‌نامه‌ی دیگر که عبارتند از: فرهنگ «هم‌نام‌ها» (در صفحه‌ی ۲۴ تفاوت بین هم‌نام‌ها و چندمعنایی‌ها شرح داده شده) و واژه‌نامه‌ی (پیکره‌ی) «امکان خطا در املاء» براساس فرهنگ معین استخراج کنیم. واژه‌نامه‌ی اخیر شامل واژه‌هایی است که نه هم‌آوا هستند، نه هم‌نویسه و نه هم‌معنا (مانند مزارع و مضارع / ادات و عُدات / تَب و طِب)، اینگونه پیکره برای اولین بار است که استخراج می‌شود، دلیل استخراج این واژه‌ها این بود که ممکن است نوآموزان در هنگام نوشتن املاء دچار شک و تردید یا اشتباه گردند.

همچنین در فرایند پژوهش به واژه‌هایی پی‌بردیم که عبارتند از: هم‌معنا- هم‌نویسه (غیرهم‌آوا) (مانند: سَفَر/ سِفِر تفاوت‌ها در گویش‌ها)، هم‌معنا- هم‌آوا (غیرهم‌نویسه) (تهران و طهران - طشت و تشت) {که شرح همه‌ی اینها در جدول شماره ۱ ذکر گردیده است.}. ولی به دلیل ضیق وقت در انجام پژوهش، تهیه‌ی این واژه‌نامه‌ها امکان پذیر نبود، امید است در فرصتی دیگر نگارنده بتواند نسبت به تهیه‌ی این واژه‌نامه‌ها نیز اقدام نماید. بی شک وجود چنین پژوهش‌های زیربنایی، موجب می‌شود تهیه و عمل بازیافت خطا در نوشتار فارسی با درصد خطای کمتر و توفیق بیشتر صورت پذیرد.

## ۲-۴- جامعه پژوهش

جامعه‌ی پژوهش طرح حاضر را تمامی واژه‌های هم‌آوا و هم‌نویسه‌ی موجود در فرهنگ ۶ جلدی معین تشکیل می‌دهند که این واژه‌ها به دو دسته‌ی مجزا تقسیم می‌گردند:

دسته نخست عبارتند از واژه‌هایی که دارای یکی از ۱۵ حروف زیر هستند:

t: ت - ط

s: ث - س - ص

h: ح - ه -



Z: ذ - ز - ض - ظ -

q: غ - ق

/ā/ و /a/ و /e/: آ - ع

این واژه‌ها به دو دلیل به وجود آمده‌اند نخست به دلیل استفاده از رسم الخط عربی دوم به دلیل نفوذ آن زبان در زبان فارسی (مانند: ثواب و صواب، حیاط و حیات، زهر و ظهر، غریب و قریب).

دسته دوم واژه‌هایی که فارسی می‌باشند و به دلایل تاریخی بوجود آمده‌اند (مانند: خوار و خار).

ما از فرهنگ معین (۱۳۷۱)، «واژه‌های هم‌آوا» را که می‌توانند شامل «هم‌نام‌ها و تعدادی از هم‌نویسه‌ها» باشند را استخراج کردیم و همچنین «هم‌نویسه‌ها» را که می‌توانند «هم‌نام‌ها و واژه‌های غیر هم‌آواها» را نیز شامل شوند، استخراج نمودیم. همچنین واژه‌های «هم نام (چند معنایی)» که شرح آن گذشت و واژه‌های «امکان خطا در املاء» که شامل «هم‌آواها مانند: هما و حمی» و «واژه‌هایی که هنگام نوشتن ایجاد تردید می‌کنند» نیز می‌باشند را استخراج نمودیم.

فرهنگ معین شامل یک مقدمه و سه بخش است:

بخش ۱- لغات.

بخش ۲- ترکیبات خارجی.

بخش ۳- اعلام.

واژه‌های مورد نظر ما مستخرج از بخش اول فرهنگ معین (۱۳۷۱) می‌باشند. این بخش شامل پنج هزار و دویست و هفتاد و هفت (۵۲۷۷) صفحه و حدوداً بیش از سی و هفت هزار (۳۷۰۰۰) مدخل می‌باشد، از این تعداد، چهار هزار و صد و نود (۴۱۹۰) مدخل هم‌آوا و پنج هزار و شش صد و پنجاه و چهار (۵۶۵۴) مدخل هم‌نویسه، که شامل هم‌نام و غیر هم‌آوا و.... نیز می‌باشد. همچنین دو هزار و نه صد و دوازده (۲۹۱۲) مدخل واژه‌های هم‌نام و دو هزار و صد و هشتاد و هشت (۲۱۸۸) مدخل نیز شامل امکان خطا در املاء می‌باشد.

## فصل سوم: روش و شیوهی تحقیق

### ۳- روش پژوهش

روش مورد استفاده در تحقیق حاضر ترکیبی است از روش‌های زبانشناسی و رایانه، بدین معنی که ابتدا برای تشخیص و چگونگی تفاوت بین واژه‌های هم‌آوا و هم‌نویسه و ... از دانش زبانشناسی و برای سهولت و دسته‌بندی و الفبایی کردن و ... از رایانه، بویژه از نرم افزار اکسل استفاده کردیم؛ در این تحقیق از روش توصیفی - تحلیلی زمینه‌یابِ پیکره‌ای استفاده شده است. پیکره‌ی مورد استفاده در این پژوهش فرهنگ شش جلدی معین بوده است.

ابتدا اشکالات مختلف املائی در زبان فارسی با استفاده از کتب گوناگون دستور زبان و زبانشناسی مورد بررسی قرارگرفت، سپس واژه‌های دارای هم‌آوایی و هم‌نویسه استخراج شده و جهت استفاده‌های الکترونیکی وارد نرم افزار صفحه‌ی گسترده‌ی excel گردید و سپس مورد تجزیه و تحلیل قرارگرفت و در نهایت پیکره‌ای واژه‌نامه‌ای هم‌آوا و هم‌نویسه استخراج گردید. لذا این پژوهش شامل مراحل زیر می‌باشد:

۱- ابتدا پیشینه و شکل‌گیری واژه‌های هم‌آوا و واژه‌های هم‌نویسه در زبان فارسی مورد تحقیق و بررسی قرار گرفت.

۲- سپس تمامی واژه‌هایی که دارای حروف «ت - ط - ث - س - ص - ح - ه - ذ - ز - ض - ظ - غ - ق - آ - ع و همزه (ء) و همچنین واژه‌هایی که دارای گروه همخوان (خو)» بوده‌اند و همچنین واژه‌های هم‌شکل و هم‌نویسه‌ی غیر هم‌معنا و واژه‌های هم‌آوا از کتاب فرهنگ معین (۱۳۷۱) استخراج گردید.

۳- در ادامه تمامی واژه‌های فارسی هم‌آوا و هم‌نویسه، مشخص شد.

۴- برای تأیید متداول بودن واژه‌های استخراج شده، فهرست این واژه‌ها در اختیار تعدادی گویشور باسواد فارسی زبان قرارگرفت تا آن‌ها واژه‌های متداول را از ناواژه‌ها (منظور از ناواژه شبه‌واژه‌هایی هستند که در زبان وجود ندارد و احتمالاً به دلیل اشتباه تایپ اشتباه تایپیست‌ها در لیست ما قرار گرفته‌اند مانند: «یاهاوژاو» و ... (non sense word)) متمایز سازند.

۵- سپس ورود واژه‌های متداول به رایانه (نرم افزار اکسل) انجام پذیرفت.

۶- در ادامه در اکسل برای واژه‌های هم‌آوا و هم‌نویسه داده پردازی (مجموعه‌سازی و مرتب‌سازی) انجام پذیرفت، پس از آن، لیست نهایی از واژه‌های هم‌آوا و هم‌نویسه در نرم افزار اکسل تهیه گردید. برای مدیریت و کنترل واژگی این حجم از داده‌ها می‌بایست از راهکارهای خودکار کمک گرفت، ما نیز برای این کار از نرم‌افزار اکسل استفاده نمودیم.

## فصل چهارم: تجزیه و تحلیل داده‌ها

### ۴- روش و ابزار گردآوری داده‌ها

واژه‌های این فرهنگ‌ها از بخش اول (لغات) فرهنگ شش جلدی معین که شامل ۵۲۷۷ صفحه می‌باشد استخراج شده است.

مرحله اول روش استخراج واژه‌ها توسط ما، در بندهای ۲ و ۳ و ۵ مراحل پژوهش شرح داده شد، تعداد واژه‌های استخراج شده ۱۳۷۷۰ واژه را شامل گردید و همراه با معانی در نرم‌افزار اکسل وارد شد، از این تعداد پس از بررسی، تعداد ۴۱۹۰ مدخل تحت عنوان «پیکره‌ی هم‌آوا» مشخص و در اکسل ثبت گردید. از این پیکره با بررسی‌های به عمل آمده، ۵۶۵۴ ردیف (مدخل همراه با معانی) «هم‌نویسه» و ۲۹۱۲ ردیف (مدخل همراه با معانی) «هم‌نام» و ۲۱۸۸ ردیف نیز شامل «امکان خطا در املاء» می‌باشند. چنانچه مشاهده می‌شود مجموع واژه‌های این چهار واژه‌نامه به ۱۴۹۴۴ واژه می‌رسد، این تفاوت از آنجا ناشی می‌شود که تعدادی از واژه‌ها در بعضی از واژه‌نامه‌ها تکرار شده‌اند مثلاً در واژه‌نامه‌ی «غلط‌های املائی» برخی از واژه‌های هم‌آوا نیز آورده شده‌اند. این واژه‌نامه‌ها به صورت جدول در اکسل برای استفاده در سایر نرم‌افزارها، استخراج گردید. که همراه با این گزارش در اختیار دفتر معاونت محترم پژوهشی مرکز منطقه‌ای اطلاع‌رسانی علوم و فناوری قرار گرفت.

### ۴-۱- خروجی‌های این طرح عبارتند از:

دو پیکره موقعیتی هم‌آوا و هم‌نویسه املائی الکترونیکی فارسی در Excel با اندازه محدود. و دو واژه‌نامه‌ی (پیکره‌ی) هم‌نام‌ها (چند معنایی) و واژه‌نامه‌ی (پیکره‌ی) «امکان خطا در املاء» براساس فرهنگ معین (این واژه‌نامه هم شامل هم‌آواها مانند «هما و حمی» و هم شامل واژه‌هایی است که نه هم‌آوا هستند، نه هم‌نویسه و نه هم‌معنا مانند «مزارع و مضارع / ادات و عُدات / تَب و طِب». به عبارت دیگر، واژه‌هایی که هنگام نوشتن ایجاد تردید می‌کنند).

### ۴-۲- این طرح می‌تواند در موارد زیر مورد استفاده قرار گیرد.

در مرکز منطقه‌ای اطلاع‌رسانی علوم و فناوری، پایگاه استنادی علوم جهان اسلام. بخش زبان و ادب فارسی دانشگاه‌های مختلف کشور بخش زبان‌شناسی و آموزش زبان فارسی به غیر فارسی زبانان دانشگاه‌های کشور، کلیه مراکز اطلاع‌رسانی، کلیه کتابخانه‌های کشور، کلیه آموزشگاه‌های دوره ابتدایی در مدارس دولتی و غیر انتفاعی، کلیه دانشگاه‌های جهان که دارای کرسی آموزش زبان فارسی می‌باشند و ....

برای تصحیح واژه تهیه‌ی فهرستی از واژه‌های کاندید جایگزینی واژه‌ی غلط و تهیه و دسترسی به قواعد و قوانین زبان مورد نظر فرایند پرچالشی است (دستغیب؛ ۱۳۹۵). لازم به توضیح است که تصحیح واژه‌های نادرست را می‌توان به سه دسته تقسیم کرد:

الف) تشخیص غیرواژه‌ها (ناواژه‌ها) (non sense word).

ب) تصحیح واژه‌ها به صورت منفرد (Isolated word, Error correction).

ج) تصحیح واژه‌های مبتنی بر متن (Error Detection).

همچنین بکارگیری روش‌ها و الگوریتم‌های تشخیص و تصحیح املائی خودکار متن، غلط یاب ماشینی، خطایاب‌ها و سامانه‌های ویراستاری متن بدون بکارگیری چنین واژه‌نامه‌هایی به نتیجه‌ی مطلوب نخواهد رسید، زیرا این گونه فرهنگ‌ها، ابزارهای بنیادین برای بازدهی مناسب و بادقت بالای چنین نرم‌افزارهایی می‌باشند.

از قابلیت‌ها و کاربردهای دیگر این پیکره می‌توان به اصلاح خطاهای املائی، و نیز استانداردسازی متون

فارسی اشاره کرد.

## فصل پنجم: بحث و نتیجه‌گیری

### ۵- روابط واژگانی

برای درک بهتر، از روابط واژگانی و بین کلمات هم‌آوا، هم‌نام و هم‌نویسه، ویژگی‌های هر کدام در جدول ذیل آمده است.

جدول ۱: ویژگی‌های هم‌آوا، هم‌نام، هم‌نویسه و هم‌معنا

ویژگی‌ها و نمونه‌ها	تلفظ	معنا	نوشتار (نویسه)
هم‌آوا (غیرهم‌معنا، غیرهم‌نویسه، هم‌آوا) سَفَر؛ صَفَر هُما (پرنده)؛ حُمی (تب) خوار؛ خار	+	-	-
هم‌نام (غیرهم‌معنا، هم‌نویسه، هم‌آوا) شیر؛ شیر	+	-	+
هم‌نویسه (غیرهم‌معنا، غیر هم‌آوا) صَفَر؛ صِفر کَرَم؛ کَرِم	-	-	+
هم‌معنا (غیرهم‌نویسه، غیر هم‌آوا) زهی؛ آفرین	-	+	-
هم‌معنا، هم‌نویسه، غیر هم‌آوا سَفَر؛ سِفر (تفاوت گویشی و لهجه)	-	+	+
هم‌معنا، هم‌آوا، غیر هم‌نویسه طهران؛ تهران	+	+	-
خطاهای املائی (امکان خطا در املاء) غیر هم‌آوا، غیر هم‌نویسه، غیر هم‌معنا مزارع؛ مضارع / ادات؛ عُدات (دشمنان)	-	-	-

## ۱-۵- هم‌ها

۱- هم‌آوا (غیرهم معنا، غیرهم نویسه، هم‌آوا): واژه‌های هم‌آوا واژه‌هایی هستند که، فقط از نظر آوایی شبیه هم تلفظ می‌شوند ولی از نظر معنایی و شکل نوشتاری با هم متفاوت هستند. مانند: سَفَر؛ صَفَر هُما (پرنده)؛ حُمی (تب) خوار؛ خار

۲- هم‌نام (غیرهم معنا، هم نویسه، هم‌آوا): هم‌نام‌ها واژه‌هایی هستند که، هم شبیه هم نوشته می‌شوند و هم شبیه هم تلفظ می‌گردند اما از نظر معنا با هم کاملاً متفاوت هستند. مانند: شیر؛ شیر شانه؛ شانه و ...

۳- هم‌نویسه (غیرهم معنا، غیر هم‌آوا، هم نوشتار): هم‌نویسه‌ها واژه‌هایی هستند که، شبیه هم نوشته می‌شوند، اما از نظر معنا و شکل تلفظ باهم کاملاً متفاوت هستند. مانند: صَفَر؛ صِفَر کَرَم؛ کِرَم

۴- هم معنا (غیرهم نویسه، غیر هم‌آوا، هم معنا): هم‌معناها واژه‌هایی هستند که فقط هم معنا هستند ولی از نظر تلفظ و نوشتار کاملاً متفاوت می‌باشند. مانند: زهی؛ آفرین

۵- هم معنا، هم‌نویسه، غیر هم‌آوا نیز واژه‌هایی هستند که از نظر آوایی فقط باهم تفاوت دارند ولی از نظر معنا و نوشتار کاملاً شبیه همدیگر می‌باشند. مانند: سَفَر؛ سِفَر (معمولاً اینگونه واژه‌ها بیشتر در تفاوت‌های گویشی و لهجه‌ها مشاهده می‌شوند)

۶- هم‌معنا، هم‌آوا، غیر هم‌نویسه نیز واژه‌هایی هستند که مانند هم تلفظ می‌شوند و هم‌معنا هستند ولی با شکل نوشتاری متفاوت. مانند: طهران؛ تهران تشت؛ طشت

۷- غیر هم‌آوا، غیر هم‌نویسه، غیر هم‌معنا که ما در اینجا آن‌ها را خطاهای املائی (امکان خطا در املاء) نامیدیم، و هدف ما از مطرح کردن اینگونه واژه‌ها، نشان دادن واژه‌هایی است که در هنگام نوشتار ممکن است نویسنده بخاطر شباهت تلفظ مشترک بعضی از حروف فارسی مرتکب خطا گردد. مانند: مزارع؛ مضارع / ادات؛ عُدات (دشمنان).

اگر چه برای هم‌معنایی در این پژوهش پیکره‌ای تهیه نشده اما بنا به ضرورت اجمالاً مورد بررسی قرار می‌دهیم.

## ۲-۵- هم‌معنایی

دو صورت زبانی متفاوت با معنای یکسان هم‌معنا تلقی می‌شوند. عاصی (۱۳۸۳) معتقد است که هم-معنایی (Synonymy) مطلق کمتر وجود دارد زیرا کلمات دارای کاربرد یکسان نیستند. نمونه‌ی بارز هم-معنایی مطلق دو کلمه کامپیوتر و رایانه در زبان فارسی است. به اعتقاد پالم (۱۹۷۶) با مطالعه‌ی واژه‌های هم-معنا می‌توان پنج امکان تمایز میان آن‌ها در نظر گرفت. نخست اینکه گروهی واژه‌های هم‌معنا، متعلق به گونه‌های خاص زبانی هستند مانند: «شدن، در گویش گیلکی و رفتن در گویش مشهدی» و «پاییز و خزان». دوم اینکه چند واژه ممکن است در سبک‌ها یا کاربردهای مختلف هم‌معنا باشند مانند: آقا، مرد، جناب. سوم بعضی از واژه‌ها تنها از نظر معنایی عاطفی یا سنجش معنایی از یکدیگر متمایز هستند و جنبه‌های دیگر معنایی یا معناساختی آن‌ها یکسان است. برای مثال دولتمرد / سیاستمدار، نهان / مخفی یا صورت‌های سه‌گانه‌ی صرفه-جو / مقتصد / خسیس را می‌توان نام برد. چهارم، بعضی از واژه‌ها از نظر هم‌نشینی در محدودیت هستند، یعنی ارائه‌ی آن‌ها تنها در ارتباط با واژه‌هایی خاص میسر است. مثل آشپزخانه / آشپزمنزل. پنجم، اینطور به نظر می‌رسد که معنای بعضی از واژه‌ها نزدیک به هم یا منطبق بر هم هستند. این هم‌معنایی از آن نوعی است که مورد استفاده‌ی فرهنگ‌نویسان قرار می‌گیرد. برای نمونه هم‌معناهای ممکن برای صفت بالغ عبارتند از: به حد رشد رسیده، جاافتاده، کامل و شایسته.

مدرسی (۱۳۸۷: ۴۶۵) می‌نویسد: واژه‌ها، یا عباراتی که دارای معنای یکسانی هستند، هم‌معنا یا مترادف خوانده می‌شوند. هم‌معنایی وقتی ممکن است که بتوان به جای جزء یا اجزایی از یک زنجیره‌ی زبانی، جزء یا اجزای دیگر نهاد، بدون اینکه معنی زنجیره تغییر یابد، باید توجه داشت که دو واژه هنگامی هم‌معنا تلقی می‌شوند که مجموعه‌ی مشخصه‌های آن‌ها یکسان باشد. بنظر می‌رسد یافتن دو هم‌معنای کامل دشوار است. مثلاً واژه «ثابت» را همه کس به کار می‌برد، اما واژه «پایا» یا «برجای» کاربرد خاص و ادبی دارد، بنابراین، با مطالعه در محتوا و مشخصه‌های واژه، معلوم می‌شود که دو واژه هیچ‌گاه به‌طور دقیق هم‌معنا نیستند، زیرا ممکن است بار محتوایی و معنایی آن‌ها به مقدار کمی با هم تفاوت داشته باشد.

به نظر می‌رسد که امکانات زبانی، اجازه نمی‌دهد که دو واژه، معنای واحدی را بطور کامل داشته باشند. فقط گاهی واژه‌ها در برخی هم‌نشینی‌ها روی زنجیره‌ی گفتار، می‌توانند معنای واحدی بیابند. مثلاً «تر» و



«خیس» در دوجمله: الف) این لباس خیس است. ب) این لباس تر است. می‌توانند در این زنجیره جانشین یکدیگر شوند، اما در جمله «او تر دامن است» نمی‌توان واژه «خیس» را جانشین «تر» کرد، زیرا این بافت تنها با واژه «تر» می‌تواند معنای «آلوده» را برساند و قابل جانشینی با «خیس» نیست.

### ۳-۵- هم معنا هم شکل، غیر هم آوا

غالباً واژه‌های هم‌معنا هم‌شکل و غیرهم‌آوا در تلفظ واژگان در گویش‌های متفاوت مشاهده می‌شود، مانند: آلوار – الوار (گلیپایگانی)، الکل – اَلکل. مدرسی این هم‌معنایی را هم‌نویسه می‌نامد و می‌نویسد: هم‌نویسه (هم-نامی ناقص) دو واژه، زمانی هم‌نویسه‌اند که صورت نوشتاری یکسان، اما صورت‌های آوایی متفاوت داشته باشند. مانند شور (Šour / Šur) و دور (Dour / Dur). صورت‌های نوشتاری و معنایی این واژه‌ها همسان، اما صورت‌های تلفظی آن‌ها جداگانه است (مدرسی؛ ۱۳۸۷: ۴۶۸).

### ۴-۵- چند معنایی

یک واژه که دارای چندین معنای به هم وابسته و مرتبط با یکدیگر هست چند معنایی (polysemy) نام دارد. مانند دل به معنای قلب، مرکز، میان، جرأت، و شکم (عاصی، ۱۳۸۳). اغلب در انتخاب میان واژه‌های چندمعنا و واژه‌های هم‌معنا با مشکل مواجه می‌شویم ولی به عقیده پالمر (۱۹۷۶) عاملی که موجب انتخاب مقوله‌ی چندمعنایی به جای هم‌نامی می‌شود، در نظر گرفتن معنای مرکزی یا یک مرکز ثقل معنایی است. (رحیمیان و رزمجو، ۱۳۸۹: ۱۸۴).

### ۵-۵- هم آوا هم معنا، غیر هم شکل

نوشتار برخی از واژه‌ها در طول زمان به دلایل مختلف از جمله سلیقه‌ها، متفاوت نوشته می‌شوند، به عبارت دیگر نگارش آن‌ها به دو شکل می‌باشد که هر دو شکل نیز صحیح محسوب می‌شوند، مانند: طهران – تهران، اطاق – اتاق، طشت – تشت و ... اینگونه واژه‌ها در لیست واژه‌های هم‌آوا قرار می‌گیرند.

### ۲-۵- نتیجه گیری

با بکارگیری تصمیم‌گیری و منطق انسانی در مراحل مختلف تهیه واژه‌نامه‌ها و پیکره‌های هم‌آوا و هم‌نویسه، توانستیم دو واژه‌نامه‌ی دیگری نیز تهیه نماییم، چهار واژه‌نامه‌ی زیر را بر اساس فرهنگ شش جلدی معین تهیه

و همراه با این گزارش به دفتر معاونت محترم پژوهشی مرکز منطقه‌ای اطلاع‌رسانی علوم و فناوری تحویل  
نماییم که عبارتند از:

۱- واژه‌نامه‌ی (پیکره‌ی) هم‌آوای زبان فارسی براساس فرهنگ معین.

۲- واژه‌نامه‌ی (پیکره‌ی) هم‌نویسه‌ی زبان فارسی براساس فرهنگ معین.

۳- واژه‌نامه‌ی (پیکره‌ی) هم‌نام‌ها براساس فرهنگ معین.

۴- واژه‌نامه‌ی (پیکره‌ی) امکان خطا در املاء براساس فرهنگ معین.

این پیکره‌های زبانی به منظور بکارگیری در آموزش زبان فارسی، خطایاب‌های رایانه‌ای، تصحیح متون  
خودکار و ..... تهیه گردیده‌اند. و می‌توانند مورد استفاده‌ی پژوهشگران سامانه‌های غلط‌یاب، تشخیص متن و  
..... قرار گیرند.

نتایج این پژوهش نشان داد که کل واژه‌های استخراج شده ۱۴۹۴۴ مدخل همراه با معانی واژه‌ها می‌باشند  
که از این تعداد ۲۸٪ را «هم‌آواها» و ۳۸٪ را «هم‌نویسه‌ها» و ۱۵٪ را «امکان خطا در املاء» و ۱۹٪ را نیز  
«هم‌نام‌ها تشکیل می‌دهند. همچنین معادل ادبیات سنتی هم‌نویسه‌ها «جناس خطی» هم‌نویسه‌های غیر هم‌آوا  
«جناس ناقص»، هم‌نام‌ها «جناس تام»، هم‌آواها «جناس لفظی»، و برخی واژه‌ها که با پیشوندها یا پسوندها با  
واژه‌های دیگر هم‌نویسه می‌شوند «جناس مرکب» (مانند کمند و کمند) می‌باشند.

در این تحقیق برای اولین بار واژه‌نامه‌ی «امکان خطا در املاء» که شامل هم‌آواها و واژه‌های دارای واج‌های  
مشابه بودند، استخراج گردید تا بتوانید در اختیار آموزگاران و مدرسین دوره‌های ابتدایی برای تدریس واژه‌های  
دشوار قرار گیرد.

در انتها پیشنهاد می‌شود که این تحقیق در مورد سایر واژه‌نامه‌ها از جمله: لغت‌نامه‌ی دهخدا، برهان قاطع  
و ... نیز انجام پذیرد و در نهایت با یکدیگر مقایسه گردند.

## منابع

- برنجیان، شاپوررضا؛ (۱۳۸۹)؛ "نظری بر نواقص خط فارسی"، اباختر، شماره ۱۹ و ۲۰، اسفند ۱۳۸۹
- دستغیب، محمد باقر (۱۳۹۵). تولید مجموعه‌ای برای مقایسه اثر طول واژه و محل رخداد خطا در تصحیح خطاهای لغوی برای واژگان زبان فارسی، طرح تحقیقاتی، مرکز منطقه‌ای اطلاع رسانی علوم و فناوری.
- دهخدا، علامه، " لغت نامه دهخدا" (۱۳۷۳)، انتشارات موسسه لغت نامه دهخدا، دانشگاه تهران.
- رحیمیان، جلال و رزمجو، سید آیت اله (۱۳۸۹). *مبانی زبان‌شناسی* ۱. مرکز نشر دانشگاه شیراز، شیراز.
- سمیعی، احمد (۱۳۶۷). واژه‌های فریبکار، ناشناسهای آشنا. *نشر دانش*، بهمن و اسفند ۱۳۶۷، شماره دوم، سال نهم، شماره مسلسل ۵۰.
- صادقی، علی اشرف؛ زندی مقدم، زهرا؛ (۱۳۸۵). فرهنگ املایی خط فارسی (براساس دستور خط فارسی مصوب فرهنگستان زبان و ادب فارسی) نسخه رقومی. فرهنگستان زبان و ادب فارسی، تهران.
- صفوی، کورش؛ (۱۳۹۲) " در آمدی بر معنانشناسی" انتشارات پژوهشکده فرهنگ و هنر اسلامی، تهران.
- عاصی، م (۱۳۸۳). پردازش دستوری زبان فارسی با رایانه، دستور، ویژه نامه فرهنگستان، شماره ۱. تهران.
- عمید، حسن، (۱۳۸۰) " فرهنگ فارسی عمید" انتشارات امیر کبیر، تهران ۱۳۸۰
- فرهنگستان زبان و ادب فارسی، (۱۳۸۱) " دستور خط فارسی"، خرداد ۱۳۸۱
- مدرسی، فاطمه؛ (۱۳۸۷)؛ "از واج تا جمله" انتشارات نشر چاپار چاپ دوم؛ تهران
- معین، محمد؛ (۱۳۷۱)؛ فرهنگ فارسی معین (فرهنگ لغت ۶ جلدی)؛ انتشارات امیر کبیر. تهران.
- همایی، جلال الدین (۱۳۸۳). *فنون بلاغت و صناعات ادبی*. انتشارات نشر هما. تهران ۱۳۸۳.

Damerau. F. J and. Mays. E. (1989). "An examination of undetected. Typing errors." *Inf process. Manag*, vol 25 , no 6.

Dastgheib, Mohammad Bagher . Fakhrahmad, seyed Mostafa. Zolghadri jahromi, Mansoor (2016) *Perspell: A New Persian Semantic Based Spelling*

Correction System . Digital Scholarship in the HUMANITIES Advance Access  
Published March27,2016.

Faili, H., Ehsan, N ., Montazery,.M., Pilehvarm. M. T (2014).Vafa Spell –  
checker for detecting spelling, grammatical, and real – word errors of Persian  
language . literary and linguistic computing, fgu 043.

Francis, W N and H. Kucera (1964) *Manual of Information to Accompany  
Standard Sample of Present-Day Edited American English, for Use with Digital  
Computers"* (revised 1979), Providence, RI: Department of Linguistics,  
BrownUniversity.

Kashefi, o, Sharifi, M., Minaie, B. (2012). A nove 1 string distance metric for  
ranking Persian respelling suggestions . natural language Engineering, 19(02)

Kennedy, Graeme (1998) *An Introduction to Corpus Linguistics*, London:  
Longman.

Leech, Geoffry (1992) "Corpora and theories of linguistic performance" in:  
Svartvik (1992): 105-122.

Mitton, R (1987). "Spelling chechers, Spelling correctors and the misspellings  
of poor Spellers" inf. Process. Manag, vol. 23, no 5.

Palmer, F.R. (1976). *Semantics: a new outline*, Cambridge: Cambridge  
University press. Pollock. j. j and A.Zamora, (1984). Automatic. Spelling  
correction in Scientific and scholarly. Text" commun. ACM. Vol, 27, no 4.

Pollock. j. j and A.Zamora, (1984). Automatic. Spelling correction in Scientific  
and scholarly. Text" commun. ACM. Vol, 27, no 4.

Richards. J. C. , platt, J and platt.H (1992). *Longman dictionary of Language  
teaching & applied Linguistics* . Essex: Longman.

Richards. J. C. and Schmidt, R. (2002). *Longman dictionary of Language  
teaching & applied linguistics*: Essex: Longman.

Svartvik, Jan (ed.) (1992) *Directions in Corpus Linguistics*, Berlin: Mouton de  
Gruyter.

Teubert, Wolfgang (1999) "Corpus linguistics: A partisan view" in:  
*International Journal of Corpus Linguistics*, Vol.4, No.1.

Thomas, J. (1995). *Meaning in interaction: an introduction to pragmatics*.  
New York: longman Group.

Yule, G. (2006). *The study of language*. Cambridge: Cambridge University  
Press.



								هم گویند.	
								ابهر (abhar)	واژه
							نرگس .	رگی است در پشت ، رگ پشت که به دل پیوسته است ، رگ جان ، اورتی ، ام الشرایین .	معنی
								ابيض (abiz)	واژه
							جرقه ، شراره آتش .	۱ - سفید، سپیدرنگ . ۲ - شمشیر . ۳ - جوانی . ۴ - مرد پاک ناموس .	معنی
								atamm اتم	واژه
							زیتون .	تمامتر، کاملتر.	معنی
							اسلم (aslam)	اِثْم (aslam)	واژه
							سالم تر، درست تر، بی گزندتر.	رخنه دار، رخنه یافته .	معنی
								اصواب (asvāb)	واژه
							صواب تر، درست تر، راست تر.	ج ثوب ؛ جامه ها.	معنی
								اجل (ajal)	واژه
							گل و لای سیاه و بدبو. ج . عجله .	جلیل تر، بزرگوارتر.	معنی
								اجم (ajam)	واژه
							۱ - غیر عرب . ۲ - ایرانی .	نیستان ، بیشه ، انبوه درختان .	معنی
								اجمه (ajame)	واژه
							۱ - یک هسته خرما . ۲ - یک تکسک انگور . ۳ - خرماین که از هسته روید؛ ج . عجمات .	بیشه ، نیستان . ج . اجم .	معنی

								واژه	احباط (ahbāt)	احباط (ahbāt)	۱۳
								معنی	فرود آوردن ، هبوط دادن .	۱ - باطل گردانیدن . ۲ - دوری کردن از کسی .	
								واژه	احسان (ehsān)	احسان (ehsān)	۱۴
								معنی	۱ - نیکی کردن . ۲ - بخشش کردن . بخشکاری ، بخشش .	۱ - استوار و محکم کردن . ۲ - نگه داشتن نفس از انجام کار بد . ۳ - شوی کردن زن . ۴ - زن گرفتن مرد . ۵ - زن و مردی که به عقد دائم در آمده باشند که مرد «محسن» و به زن «محسنه» گویند.	
								واژه	اخص (axass)	axass اخص	۱۵
								معنی	خاص تر، گزیده تر.	خسیس تر، زبون تر، فرومایه تر، خوارتر.	
								واژه	ادیم (adim)	ادیم (adim)	۱۶
								معنی	نیست شده ، نابود.	۱ - چرم دباغی شده . ۲ - پوست خوشبوی سرخ رنگ . ۳ - روی زمین . ۴ - سفرة غذا.	
								واژه	اضاعه azāe	azāe اذاعه	۱۷
								معنی	نک اضاعت .	آشکار ساختن ، فاش کردن .	
								واژه	ازالت (ezālat)	اذالت (ezālat)	۱۸



								۱ - طرد کردن ، دور کردن . ۲ - زایل کردن ، از بین بردن .	فروهستن دامان ، دراز کردن دامن .	معنی	
								(azal) ازل	(azal) اذل	واژه	۱۹
								زمان بی ابتداء.	ذلیل تر، خوارتر.	معنی	
								(azeme) ازمه	ازمه (azeme)	واژه	۲۰
								ج . زمام ؛ مهارها، افسارها.	ج زمام و جج ذمه ؛ حق و حقوق .	معنی	
								(arab) عرب	(arab) ارب	واژه	۲۱
								تازی ، از نژاد عرب . ؛به عرب عجمی بند نبودن کنایه از: هیچ پشت و پناهی نداشتن . ؛ از بیخ عرب بودن کنایه از: الف - مطلقاً انکار کردن . ب - هیچ ندانستن .	۱ - حاجت . ۲ - مقصود. ج . آراب .	معنی	
							عرض (arz)	(arz) ارض	(arz) ارز	واژه	۲۲
								زمین .	۱ - بها، قیمت ، ارزش . ۲ - پول خارجی ، پول بیگانه . ؛ ارز تهاتری ارزی که در قراردادهای پایاپای مبنای محاسبه قرار می گیرد. ؛ ارز یوزانس ارزی که پس از دریافت کالا حواله می شود. ؛ ارز دولتی ارزی که دولت از طریق بانک های مجاز و به نرخ دولتی می فروشد. ؛ ارز دانشجویی ارزی که	معنی	

									دولت به دانشجویان خارج از کشور برای ادامه تحصیل می دهد. ؛ ارز شناور ارزی که بهای آن ثابت نیست و براساس عرضه و تقاضا تعیین می شود. ؛ ارز رقابتی ارزی که از سوی دولت در رقابت با بازار آزاد عرضه می شود. ؛ ارز صادراتی ارزی که از طریق فروش کالای صادراتی تأمین می شود.		
							ارزه	ارزه (arze)	ارزه (arze)		
							۲۳	معنی	۱ - موریانه ، چوب خواره ، دیوچه ، دیوک . ۲ - زنگ آهن .	۱ - به نمایش گذاشتن . ۲ - نمایش ، ارائه .	
							۲۴	معنی	اراق (arag)	عرق بدن	
							۲۵	معنی	ارقه (arage)	نک ارقه .	

							اروس (arous)	اروس (arous)	واژه
							۱ - زن تازه شوهر کرده . ج . عرائس . ۲ - در فارسی : زن پسر . ۳ - بهترین ، زیباترین ؛ عروس هزار داماد کنایه از : دنیا و بی وفایی آن .	کالا ، متاع .	معنی ۲۶
							عریش (ariš)	اریش (ariš)	واژه
							۱ - کجاوه ، هودج . ۲ - کلبه ، کومه . ۳ - چفت انگور .	زیرک ، عاقل .	معنی ۲۷
							عریکه (arike)	اریکه (arike)	واژه
							خلق و خوی .	تخت ، سریر .	معنی ۲۸
							عض (az)	از (az)	واژه
							۱ - علامت مفعول غیر - صریح یا باواسطه . ۲ - علامت ابتدا و آغاز . ۳ - در ، اندر . ۴ - برای ، بهر ، به سبب . ۵ - نسبت به ، در مقایسه با . ۶ - به دلیل ، به علت . ۷ - در ، اندر . ۸ - از سویی ، از طرف . ۹ - به جای ، در عوض .	۱ - علامت مفعول غیر - صریح یا باواسطه . ۲ - علامت ابتدا و آغاز . ۳ - در ، اندر . ۴ - برای ، بهر ، به سبب . ۵ - نسبت به ، در مقایسه با . ۶ - به دلیل ، به علت . ۷ - در ، اندر . ۸ - از سویی ، از طرف . ۹ - به جای ، در عوض .	معنی ۲۹
							عذار (ozār)	ازار (ozār)	واژه
							۱ - خط ریش . ۲ - در فارسی به معنی رخسار ، چهره . ۳ - لگام اسب .	۱ - زیرجامه ، شلوار . ۲ - دستار ، فوطه .	معنی ۳۰
							عزیز (aziz)	ازیز (aziz)	واژه
							۱ - گرمی ، محبوب . ۲ - ارجمند ، بزرگواری . ۳ - کمیاب ، نادر .	۱ - به جوش آمدن . ۲ - غلغل ، صدای جوشیدن دیگ	معنی ۳۱



								اختیار و آزادی بدون مداخله کشورهای بیگانه .		
								استلام (estelām)	واژه	
								از بیخ برکندن چیزی را، از بن برکندن .	معنی	۳۸
								۱ - لمس کردن ، دست کشیدن به چیزی . ۲ - بوسه دادن .		
								اسم (esm)	واژه	
								۱ - کلمه ای که به وسیله آن چیزی یا کسی را می خوانند، نام . ۲ - عنوان . ۳ - شهرت ، آوازه . ، سفیدی ، هوش .	معنی	۳۹
								۱ - گناه ، خطا . ۲ - کاری که کردن آن روا نباشد .		
					عسیر (asir)	عسیر (asir)	عسیر (asir)	واژه		
								به باور قدما کره آتش که بالای کره هواست ، سیالی رقیق و بی وزن است .	معنی	۴۰
								۱ - گرفتار . ۲ - برده ، بنده . ج . اسراء .		
								اسطرلاب (ostorlāb)	واژه	
								استاره سنج ، ابزاری است که برای اندازه گیری محل و ارتفاع ستارگان و دیگر اندازه گیری های نجومی بکار می رود .	معنی	۴۱
								نک اسطرلاب .		
								اصعاد (esād)	واژه	
								۱ - نیکبخت گردانیدن . ۲ - مبارکی . ۳ - یاری ،	معنی	۴۲
								بالا رفتن ، صعود کردن		



								۱ - وام خود را بازخواستن . ۲ - درخور بودن . ۳ - خواهش ، درخواست . ۴ - لزوم .	غذا گرفتن .	معنی	
								اغتراب (egterāb)	اغتراب (egterāb)	واژه	
								نزدیک شدن .	۱ - از دیوار خویش دور شدن . ۲ - با بیگانگان ازدواج کردن .	معنی	۵۱
								اقتراف (egterāf)	اغتراف (egterāf)	واژه	
								۱ - کسب کردن ، کسب معاش کردن . ۲ - گناه کردن . ۳ - به جا آوردن . ۴ - وزیدن .	با کف دست آب خوردن .	معنی	۵۲
							عقرب agrab	اقراب agrab	اغرب agrab	واژه	
							۱ - کژدم ؛ ج عقارب . ۲ - نام صورتی فلکی در نیم کره جنوبی آسمان و نام هشتمین برج از بروج دوازده گانه که خورشید در حرکت ظاهری خود، آبان ماه در این برج قرار می گیرد.	نزدیکتر .	شگفت تر، غریب تر.	معنی	۵۳
								اقفال (egfāl)	اغفال (egfāl)	واژه	۵۴
								۱ - قفل کردن . ۲ - حرکت کردن .	غافل کردن ، گول زدن .	معنی	
								اغلاق (aglāg)	اغلاق (aglāg)	واژه	
								۱ - بی آرام کردن ، آرام بودن . ۲ - جنبانیدن .	۱ - دریستن . ۲ - سخن را پیچیده گفتن .	معنی	۵۵







										آب یا شراب . ۳ - آفتابه ، لولهین . ۴ - مطهره . ۵ - وزنی معادل دو من .	
										واژه	۵
									ابی (abi)	ابی (abi)	معنی
									پدری ، صلبی .	بی ، بدون .	واژه
									احدی (ahadi)	احدی (ahadi)	معنی
									یک تن ، هیچکس ، کسی .	۱ - منسوب به احد . ۲ - مربوط به خدای یگانه . ۳ - فرقه ای از سپاهیان پادشاه هند .	۶
									احسن (ahsan)	احسن (ahsan)	واژه
									نیکوتر ، بهتر . ؛ به نحو احسن به بهترین شیوه و طرز . ؛ احسن التقویم بهترین شکل ، بهترین صورت .	آفرین ، احسنت ، مرحبا .	۷
									اختر (axtar)	اختر (axtar)	واژه
									شمار ستاره شناس ، منجم .	ستاره ، کوکب ، نجم .	معنی
									ادا (adā)	ادا (adā)	واژه
									اصول ناز و کرشمه ، نمودن کراهت و جز آن .	۱ - به جا آوردن ، پرداختن دینی که بر شخص فرض و لازم است . ۲ - ناز ، کرشمه . ۳ - رمز ، اشاره . ۴ - حرکات مسخره و بیهوده . ۵ - تقلید .	معنی
									ارتفاع	ارتفاع	واژه
											۱۰

								(ertefā)	(ertefā)		
								گرفتن بالا رفتن ، بلند شدن .	۱ - برخاستن ، بلند شدن . ۲ - جمع آوری محصول . ۳ - بلندی ، اوج . فاصله بین رأس تا ضلع روبرو . ۴ - حاصل زراعت . ۵ - بلندی سطح زمین نسبت به سطح دریا . ۶ - فاصله ستاره از افق .	معنی	
								ارداف (ardāf)	ارداف (ardāf)	واژه	
								وزیران و نزدیکان پادشاه ، بزرگان دربار .	۱ - از پی درآمدن ، پیروی کردن ، در پی کسی رفتن . ۲ - از پی درآوردن ، سپس نشانیدن ، به ترک نشانیدن . ؛ ارداف نجوم : از پس یکدیگر برآمدن ستارگان .	معنی	۱۱
								اردو (ordou)	اردو (ordou)	واژه	
								زبان مردم پاکستان و بخشی از هندوستان ، که مربک از فارسی ، عربی و هندی است .	۱ - گروهی سپاهیان با تمام لوازم که به سویی فرستاده شوند . ۲ - لشکرگاه ، محل لشکر . ۳ - محلی که ورزشکاران یا دانش آموزان برای تمرین یا تفریح مدتی	معنی	۱۲

									معین به آنجا می روند.		
								ارژنگ (aržang)	ارژنگ (aržang)	واژه	
								نام کتاب مانی که دارای انواع نقش و نگار بوده و به موجب آن مانی ادعای پیغمبری داشت .	نام چاهی که افراسیاب ، بیژن را در آن زندانی کرد.	معنی	۱۳
								ارمک (armak)	ارمک (armak)	واژه	
								۱ - پارچه پشمینه ، صوف . ۲ - کلاه پشمین . ۳ - پارچه ای پنبه ای به رنگ خاکستری .	نام چند گونه درختچه از تیره ریش بزها است .	معنی	۱۴
								ارواح (arvāh)	ارواح (arvāh)	واژه	
								ج ریح ؛ بادها.	ج روح ؛ روح ها، روان ها.	معنی	۱۵
								اروند (arvand)	اروند (arvand)	واژه	
								۱ - شأن و شوکت . ۲ - حسرت ، آرزو.	۱ - افسون و نیرنگ . ۲ - رود دجله .	معنی	۱۶
								از بر (azbar)	از بر (azbar)	واژه	
								۱ - از حفظ ، از حافظه . ۲ - به یاد سپرده شده .	بالای ، بر فراز .	معنی	۱۷
								ازار (azāre)	ازار (azāre)	واژه	
								= ازاره . ایزاره . هزاره : پایاب ، قعر	۱ - زیرجامه ، شلوار . ۲ - دستار ، فوطه .	معنی	۱۸

									آب .			
									اسپک (espak)	اسپک (espak)	واژه	
									= اسپک : ۱ - اسب چوبین یا گلین که کودکان برای بازی می سازند. ۲ - یکی از اندام های اسطربلاب . ۳ - خیمه بزرگ .	ضربه محکم و سریع به توپ برای رد کردن توپ به محوطه حریف از روی تور، اسبک ، آبشار .	معنی	۱۹
									است (ast)	است (ast)	واژه	
									استر .	۱ - استخوان . ۲ - هسته میوه .	معنی	۲۰
									استامبولی (estamboul y)	استامبولی (estambou ly)	واژه	
									= استانبولی . اسلامبولی : ۱ - آن چه منسوب به استامبول باشد . ۲ - از مردم استامبول ، اهل استامبول .	ظرفی فلزی که برای بردن گل و گچ در بنایی به کار برند .	معنی	۲۱
									استحیاء estehyā	استحیاء estehyā	واژه	
									شرم داشتن ، شرم کردن .	زنده نگه داشتن .	معنی	۲۲
									استیل (estil)	استیل (estil)	واژه	
									فولاد ضدزنگ .	۱ - شکل ، فرم ، اندازه . ۲ - نوعی میل به سبک قدیم فرانسه . ۳ - شیوه ای در انجام مهارت های ورزشی .	معنی	۲۳
									اسحار (eshār)	اسحار (eshār)	واژه	
											۲۴	





									جامه ای که وارو به تن کرده باشند.	نوباوه .		
									اشنا (āšenā) اشنا	واژه		
									گوهر گرانبها، گوهر گرانمایه .	۳۴ معنی		
									اعراض (e,arāz) اعراض	واژه		
									شیر پستان زن هنگامی که عصبی یا وحشت زده باشد یا غم و اندوه شدیدی به او دست داده باشد. خوردن شیر اعراض باعث ناخوشی کودک می شود.	۳۵ معنی		
									اعراض (E,arāz) اعراض	واژه		
									ج عَرَض ؛ بیماری ها.	۳۶ معنی		
									افرنگ (afrang) افرنگ	واژه		
									تخت پادشاهی .	۳۷ معنی		
									افسر (afsar) افسر	واژه		
									تاج ، دیهیم .	۳۸ معنی		
									اقسام (agsām) اقسام	واژه		
									کسی که در ارتش درجه اش از ستوان به بالا باشد، صاحب منصب .	۳۹ معنی		



معنی	ج قسم ؛ سوگندها .	سوگند دادن ، قسم دادن .							
واژه	(al) ال	(al) ال							
معنی	حرف تعریف است در عربی و آن چون بر اسمی نکره درآید، آن را معرفه سازد.	درختی از تیره زغال اخته ها که گاهی بعضی گونه هایش به صورت درختچه می باشند. گل هایش سفید یا زرد و میوه اش سفت و شامل یک هسته است .							۴۰
واژه	(alfanj) الفنج	(alfanj) الفنج							
معنی	چسبیده ، چسبناک .	گردآوری ، اندوختگی .							۴۱
واژه	(alo) الو	(alo) الو							
معنی	لفظی است که در شروع برقراری ارتباط تلفنی گفته می شود.	شعله آتش ، زبانة آتش .							۴۲
واژه	(emārāt) امارات	(emārāt) امارات							
معنی	ج اماره . ۱ - ولایت ها . ۲ - فرمانفرمایی ها، سرداری ها .	ج اماره .							۴۳
واژه	(anbā) انباء	(anbā) انباء							
معنی	ج نباء؛ خبرها، آگاهی ها .	خبر دادن ، آگهی دادن .							۴۴
واژه	(anboudan) انبودن	(anboudan) انبودن							
معنی	آفریدن .	چیدن ، روی هم گذاشتن .							۴۵
واژه	(andi) اندی	(andi) اندی							
معنی	۱ - امیدواری . ۲ - شگفت ، تعجب . ۳ -	۱ - آن گاه ، آن لحظه . ۲ - از این زمان ، از این لحظه .							۴۶



										شوکت .				
										اورنگ (orang)	اورنگ (orang)	اورنگ (orang)	واژه	۵۵
										مکر، فریب، حيله .	تخت پادشاهی ، سریر .	فر، شکوه ، شان .	معنی	
											(igā) ایقاد	(igā) ایقا	واژه	۵۶
											فرستادن ، روانه کردن .	برافروختن ، آتش افروختن .	معنی	
											(ilām) ایلام	(ilām) ایلام	واژه	۵۷
											به درد آوردن ، دردمند کردن .	به درد آوردن ، به رنج افکندن .	معنی	
											آ	(ā) آ	واژه	۵۸
											هان ! هلا! آی !	«آ» یا «الف ممدوده» نخستین حرف از الفبای فارسی ؛ اولین حرف از حروف ابجد، برابر با عدد «۱» .	معنی	
											(āb) آب	(āb) آب	واژه	۵۹
											۱ - یازدهمین ماه از سال سریانی برابر با «مرداد ماه ». ۲ - نام ماه یازدهم سال یهود.	مایعی است شفاف ، بی طعم و بی بو، مرکب از دو عنصر اکسیژن و نیدروژن ؛ در $O_2 H$ ، باور قدما یکی از چهار عنصر «آب ، آتش ، باد، خاک» محسوب می شده .	معنی	
											(ābād) آباد	(ābād) آباد	واژه	۶۰





									در ترکیبات به جای «آراینده» آید: بزم آرا، جهان آرا، صف آرا.	معنی	
						آرنگ (ārang)	آرنگ (ārang)	آرنگ (ārang)	واژه		
						آرنج .	۱ - رنج ، اذیت ، آزار . ۲ - مکر، حيله .	۱ - رنگ ، لون . ۲ - سُرخاب . ۳ - گونه ، روش .	معنی	۷۱	
						آریا (āryā)	آریا (āryā)	آریا (āryā)	واژه		
						مهم ترین شعبه نژاد سفید .	قسمت اصلی ملودیک اپرا .	شفایق وحشی خودرو از نوع خشخاش .	معنی	۷۲	
							آس (ās)	آس (ās)	واژه		
							۱ - تک خال ۲ - یکی از ورق - های بازی که یک خال بر آن نقش شده .	دو سنگ گرد و مسطح بر هم نهاده که به وسیله آن غلات را آرد کنند . سنگ زیرین ساکن و سنگ بالایی متحرک می باشند .	معنی	۷۳	
							آسا (āsā)	آسا (āsā)	واژه		
						خمیازه ، دهان دره .	۱ - زیب ، زینت . ۲ - وقار، تمکین . ۳ - طرز، روش . ۴ - پسوندی است که شباهت را می رساند مانند: رعدآسا .		معنی	۷۴	
							آستان (āstān)	آستان (āstān)	واژه		
						ستان ، به پشت خوابیده .	۱ - درگاه ، پیشگاه . ۲ - کفش کن .		معنی	۷۵	
							آستر (āstar)	آستر (āstar)	واژه		
						۱ - پارچه ای که زیر لباس می	آن سوی تر، زاستر .		معنی	۷۶	











																			۱ - ظرف . ۲ - کوزه آب یا شراب . ۳ - لوله هایی که شیره خام را از ریشه به برگ ها می رساند .	دلیل ، برهان	معنی	
																			<b>آهار āhār</b>	<b>آهار āhār</b>	واژه	
																			گیاهی از تیره مرکبان جزو دسته پیوسته گلبرگ ها که اصل آن از مکزیک است و دارای گونه های متعدد زینتی است .	مایعی که از نشاسته یا کتیرا درست می کنند و به پارچه می زنند تا سفت و براق شود .	معنی	۹۱
																			آهو (āhou)	آهو (āhou)	واژه	
																			جانوری از خانواده تهی - شاخان ، جزو راسته نشخوارکنندگان که اقسام مختلف دارد و عموماً دونده بسیار سریع و چابک و دارای دست و پای بلند و چشمان زییاست ، غزال ، مارال . ؛ پشت آهو بسته بودن کنایه از: دور از دسترس بودن .	۱ - عیب ، نقص . ۲ - بیماری ، مرض . ۳ - بد، ناپسند .	معنی	۹۲
																			بره بچه آهو .			

همانگونه که در صفحه‌ی ۳۰ ذکر شده است، هم‌نویسه (غیرهم معنا، غیر هم‌آوا، هم نوشتار): هم‌نویسه‌ها واژه-

هایی هستند که، شبیه هم نوشته می‌شوند، اما از نظر معنا و شکل تلفظ باهم کاملاً متفاوت هستند. مانند: صَفَر؛

صِفَر کَرَم؛ کِرِم

تفاوت اکثر این واژه‌ها در واکه‌های کوتاه می‌باشد

واژه	(abâ) ابا	(ebâ) ابا	ابا (âbâ)							
۱	معنی	با، همراه .	۱ - سرباز زدن ، سر - پیچیدن . ۲ - خودداری کردن . ۳ - سرکشی ، نافرمانی . ۴ - نخوت ، تکبر .	پدر .						
۲	معنی	ابدال (abdâl)	(ebdâl) ابدال							
۳	معنی	ج بدل یا بدیل . ۱ - - نیکان ، صالحان ، که جهان به برکت وجود ایشان برپاست . ۲ - نجیبان ، شریفان .	۱ - عوض و بدل کردن . ۲ - قرار دادن حرفی به جای حرفی دیگر برای دفع ثقل و سنگینی . ۳ - یکی از اقسام نه گانة وقف مستعمل چون تبدیل تاء به هـاء در رحمت و رحمه .							
۴	معنی	ابرو (abrû)	(âbrû) ابرو							
۳	معنی	مجموع موهای روییده بر ظاهر استخوان قوسی شکل بالای کاسه چشم زیر پیشانی . ؛	عرق، خوی							
۴	واژه	ابره abra	(obra) ابره							

									لیاس ، بخش بیرونی لیاس .	معنی								هویره ، آهویره .	
									ابریق ebrig	واژه								abrig ابریق	
									۱ - کوزه . ۲ - ظرف - سفالین با دسته و لوله برای آب یا شراب . ۳ - آفتابه ، لولهین . ۴ - مطهره . ۵ - وزنی معادل دو من .	معنی	۵								۱ - شمشیر بسیار درخشانده . ۲ - زن صاحب جمال .
									ابطال abtâl	واژه								ebtâl ابطال	
									۱ - باطل کردن ، لغو کردن ، بیهوده کردن ، ناچیز کردن . ۲ - دروغ و باطل گفتن .	معنی	۶								
									ابعاد (abâd)	واژه								(ebâd) ابعاد	
									ج بعد؛ دوری ها. ؛ ابعاد هندسی طول ، عرض و ارتفاع .	معنی	۷							۱ - دور کردن ، راندن . ۲ - دور رفتن .	
									ابکار abkâr	واژه								(ebkâr) ابکار	
									ج بکر؛ دوشیزگان ، دختران دوشیزه .	معنی	۸							۱ - پگاه برخاستن . ۲ - بامداد از خواب بیدار کردن . ۳ - بامداد.	
									ابل abol	واژه	۹							ebel ابل	



										دیده نمی شود.			
										اثبات (esbât)	اثبات (asbât)	واژه	
										۱ - ثابت گردانیدن . ۲ - نام نویسی در دفتر لشکر و سپاه .	ج ثَبِت ؛ معتمدان .	معنی	۱۴
										انتقال (esgâl)	انتقال (asgâl)	واژه	
										۱ - گرانبار کردن . ۲ - گرانبار شدن .	ج ثَقَلَ ، ثَقَلَ . ۱ - بارهای گران ، گران بها، اسباب . ۲ - اشیاء نفیس ، چیزهای گرانبها . ۳ - رخت های مسافر .	معنی	۱۵
										احباب (ehbâb)	احباب (ahbâb)	واژه	
										دوست داشتن .	ج حَبِيب ؛ دوستان ، یاران .	معنی	۱۶
										احداث (ehdâs)	احداث (ahdâs)	واژه	
										چیزی نو به وجود آوردن ، ساختن و برقرار کردن .	ج حَدَث ۱ - نوها، تازه ها، هر چیز تازه و نو پدید آمده . ۲ - جوانان . ۳ - نوعی حقوق دیوانی . ؛ احداث	معنی	۱۷

									اربعه : حدث های چهارگانه قتل ، ازالة بکارت ، شکستن دندان و کور کردن .	
									احرام (ahrâm )	واژه
									۱ - آهنگ حج کردن ، در حرم درآمدن . ۲ - مجازاً دو تکه لباس نادرخته که در ایام حج یکی را به کمر بندند و دیگری را بر دوش اندازند.	۱۸ معنی
									احکام (ehkâm )	واژه
									۱ - ج حکم رأی ها، دستور - ها . ۲ - مجموعه دستور العمل های شرعی . ۳ - آداب ، رسم ها . ۴ - مجموعه قوانین و مقرراتی که به ارادة محکوم علیه قابل تغییر است .	۱۹ معنی
									احیاء (ahyâ)	واژه
									۱ - زنده کردن ۲ - آباد کردن زمین . ۳ - شب را به عبادت گذرانیدن . ۴ - شب زنده داری کردن . ۵ - زندگی .	۲۰ معنی
									آخ (âx)	واژه
									اخ (ax)	۲۱



									۱ - صوتی است که هنگام درد و سوزش بر زبان رانند. ۲ - صوتی است برای نمودن خوشی و لذت .	برادر. ج . اخوان .	معنی	
									اخبار (axbâr)	اخبار (exbâr)	واژه	
									خبر دادن ، آگاه ساختن .	ج خبر؛ آگاهی ها، خبرها.	معنی	۲۲
									اخطار (extâr)	اخطار (axtâr)	واژه	
									۱ - در خطر افکندن . ۲ - ابلاغ و اعلام کردن . ۳ - آگهی ، اعلام .	ج خطر؛ بلاها، سختی ها.	معنی	۲۳
									اخلا (exlâ)	اخلا (axellâ)	واژه	
									خالی کردن	خلیل، دوستان	معنی	۲۴
									اخوان (exvân)	اخوان (axavân n)	واژه	
									ج آخ ؛ برادران ، دوستان ، برادرخواندگان .	تثنيه اخ ، دو برادر.	معنی	۲۵
									ادا (edâ)	ادا adâ	واژه	
									اصول ناز و کرشمه ، نمودن کراهت و جز آن .	۱ - به جا آوردن ، پرداختن دینی که بر شخص فرض و لازم است ۲ - ناز، کرشمه . ۳ - رمز، اشاره . ۴ - حرکات مسخره و بیهوده . ۵	معنی	۲۶



										خانواده .			
										ارداف (ardâf)	ارداف (erdâf)	واژه	
										وزیران و نزدیکان پادشاه ، بزرگان دربار .	۱ - از پی درآمدن ، پیروی کردن ، در پی کسی رفتن . ۲ - از پی درآوردن ، سپس نشانیدن ، به ترک نشانیدن . ؛ ارداف نجوم : از پس یکدیگر برآمدن ستارگان .	معنی	۳۲
										ارز (araz)	ارز (arz)	واژه	
										درخت سلیمان ، سرو	برنج ، دانه شلتوک	معنی	۳۳
										ارس (oros)	ارس (ors)	واژه	
										نام چند گونه سرو کوهی جزو تیره ناژویان که در اغلب نقاط استپی و خاتمة جنگل های مرطوب پراکنده	نک اروس .	معنی	۳۴

									اند، ارساء، ارجه .												
															ارذال (arzâl)	ارذال (arzal)	واژه	۳۵			
															رذل فرومايه.	خوارتر	معنی				
															ارگ (arg)	ارگ (org)	واژه	۳۶			
															قلعة کوچک میان قلعة بزرگ .	۱ - ارغنون . ۲ - نوعی ساز شبيه به پیانو اما کوچکتر از آن .	معنی				
															ارمک (armak )	ارمک (ormak)	واژه	۳۷			
															نام چند گونه درختچه از تیره ریش بزها است .	۱ - پارچه پشمینه ، صوف . ۲ - کلاه پشمین . ۳ - پارچه ای پنبه ای به رنگ خاکستری .	معنی				
															اروس (arūs)	اروس (orūs)	واژه	۳۸			
															کالا، متاع .	۱ - روس . ۲ - روسی ، از مردم روسیه .	معنی				
															اروند (arvand)	اروند (ervand)	واژه	۳۹			
															۱ - افسون و نیرنگ . ۲ - رود دجله .	۱ - شأن و شوکت . ۲ - حسرت ، آرزو.	معنی				
															اریب (arib)	اریب (orib)	واژه	۴۰			
															خردمند، زیرک ، دانا.	۱ - منحرف ، کج . ۲ - کجی .	معنی				
															ازار (ezâr)	ازار (ezâr)	واژه	۴۱			
															۱ - زیرجامه ، شلوار. ۲ -	= ازاره . ایزاره . هزاره : پایاب ، قعر آب .	معنی				

									دستار، فوطه .		
								اسپری (esperay)	اسپری (espari )	واژه	
								۱ - وسیله ای حاوی مایع که با فشار دکمه آن مایع درونش به صورت ذرات ریز خارج می شود، افشانه . ۲ - مایعی که به شکل پودر از این ظرف بیرون پاشیده می شود.	۱ - اسپری ، پایان یافته ۲ - نابود شده ، معدوم .	معنی	۴۲
								استار (estâr)	استار (astâr)	واژه	
								وزنی برابر چهار متقال ، یا چهار متقال و نیم ؛ استیر .	ج ستر؛ پرده ها .	معنی	۴۴
								استام (ostâm)	استام (astâm )	واژه	
								سیخی که در تون حمام و تنور نانوایی به کار می رود؛ کفچه آتشدان ، آتش کش ، بیلچه ، خاک انداز .	زین و یراق اسب .	معنی	۴۵
								استپ (stop)	استپ (estep)	واژه	
								فرمان ایست ، بایست .	جلگه وسیع و بی درخت ، علفزار .	معنی	۴۶
								استحیاء (stahyâ)	استحیاء (estehy â)	واژه	
								شرم داشتن ، شرم کردن .	زنده نگه داشتن .	معنی	۴۷
								استه (este)	استه (aste)	واژه	
								کفل ، سرین .	۱ - دانه	معنی	۴۸

									و هسته میوه ها. ۲ - استخوان .	
								اسحار (ashâr)	اسحار (ashâr)	واژه
								ج سحر؛ افسون ها، سحرها.	ج سحر؛ پاس آخر شب ، بامداد.	معنی
								اسراء (osarâ)	اسراء (asrâ)	واژه
								ج اسیر؛ اسیران ، گرفتاران .	۱ - در شب سیر کردن . ۲ - به سیر درآوردن کسی در شب . ۳ - معراج پیامبر اکرم .	معنی
								اسرار (esrâr)	اسرار (asrâr)	واژه
								۱ - نهفتن . ۲ - پوشیده با کسی سخن گفتن .	۱ - ج سیر؛ رازها، نهانی ها، سرها . ۲ - ج سُرر و سُرر؛ خط ها و شکن های کف دست .	معنی
							اسفار	اسفار (esfâr)	اسفار (asfâr)	واژه
							ج سفر؛ نامه ها، کتاب ها.	۱ - به روشنایی صبح درآمدن . ۲ - روشن شدن صبح ۳ - آشکار شدن .	ج سَفَر؛ سفرها.	معنی
								اشباه (ešbâh)	اشباه (ašbâh)	واژه
										۵۳







									۱ - مطابقه کردن . ۲ - برهم نهادن ، جمع شدن برای کاری .	معنی	
								اطراد (etterâ)	اطراد (etrâd)	واژه	۶۵
								روان شدن کار، راست آمدن .	کسی را طرد کردن .	معنی	
								اطناب (etnâb)	اطناب (etnâb)	واژه	۶۶
								۱ - بسیار گفتن ۲ - پرگویی . ۳ - تطویل کلام و مبالغه در آن	ج طناب ؛ ریسمان ها، بندها.	معنی	
								اعداد (e,dâd)	اعداد (a,dâd)	واژه	۶۷
								۱ - آماده کردن ، بسیجیدن . ۲ - بسیج .	ج عدد؛ ارقام ، شماره ها.	معنی	
								اعراض (e,râz)	اعراض (a,râz)	واژه	۶۸
								۱ - روی گردانیدن . ۲ - برگشتن . ۳ - نفرت داشتن . ۴ - نفرت ، کراهت	ج عرض . ۱ - آبروها . ۲ - خواسته ها.	معنی	
								اعصار (e,sâr)	اعصار (a,sâr)	واژه	۶۹
								۱ - درآمدن در عصر . ۲ - گردباد.	ج عصر؛ روزگاران ، زمان ها، دورها.	معنی	
								اعلام (e,lâm)	اعلام (a,lâm)	واژه	۷۰
								۱ - آگاه کردن ، هشدار دادن . ۲ - آگاهی .	ج عَلم . ۱ - نشانه ها . ۲ - ببیرق ها.	معنی	

واژه	اعمال (a,mâl)	اعمال (e,mâl)
۷۱	معنی	به کار بستن .
	ج عمل ۱ - کارها، کرده ها. ۲ - شغل ها، پیشه ها. ۳ - نواحی .	

#### پیوست ۴

همانگونه که در صفحه‌ی ۳۰ ذکر شده است، غیر هم‌آوا، غیر هم‌نویسه، غیر هم‌معنا که ما در اینجا آن-ها را خطاهای املائی (امکان خطا در املاء) نامیدیم، و هدف ما از مطرح کردن اینگونه واژه‌ها، نشان دادن واژه-هایی است که در هنگام نوشتار ممکن است نویسنده بخاطر شباهت تلفظ مشترک بعضی از حروف فارسی مرتکب خطا گردد. مانند: مزارع؛ مضارع / ادات؛ عُدات (دشمنان).

واژه	(abâz) اباز	(ebâz) اباض
۱	معنی	باز، دوباره
		۱ - ریسمانی که به وسیله آن خرده دست شتر بربندند تا دست از زمین برداشته دارد، بند. ۳ - نام رگی است در پای .
واژه	(abhor) ابجر	(abhar) ابهر
۲	معنی	نرگس .
	ج بحر؛ دریاها.	رگی است در پشت، رگ پشت که به دل پیوسته است، رگ جان ، آورتی، ام

			الشرایین .			
			عبد (abd)	ابد (abad)	واژه	
			۱ - عبور کردن ، گذشتن . ۲ - درگذشتن ، مردن . ۳ - جاری کردن اشک .	۱ - زمانی که آن را نهایت نباشد، همیشه جاوید. مق ازل . ۲ - قدیم ، ازلی . حیات ابد زندگی جاوید	معنی	۳
			عبر (abar)	ابر (abr)	واژه	
			۱ - عبور کردن ، گذشتن . ۲ - درگذشتن ، مردن . ۳ - جاری کردن اشک .	۱ - بر، به ، با . ۲ - بالای ، روی . ۳ - در . ۴ - بر سر .	معنی	۴
			عبره (abre)	ابره (obare)	واژه	
			عبور کردن .	هوبره ، آهوبره .	معنی	۵
			عبری (ebri)	ابری (abry)	واژه	
			نک عبرانی .	۱ - پوشیده از ابر . ۲ - با نقشی چون موج آب یا ابرهای بریده از یکدیگر	معنی	۶
			ابیز (abiz)	ابیض (abyaz)	واژه	
			جرقه ، شرارة آتش .	۱ - سفید، سپیدرنگ . ۲ - شمشیر . ۳ - جوانی . ۴ - مرد پاک ناموس .	معنی	۷
			عطا (atâ)	اتا (âtâ)	واژه	
			۱ - بخشیدن . ۲ - بخشش ، دهش . ۳ - آن چه که بخشیده شود؛ ج . عطیه .	پدر .	معنی	۸
			عطر (atr)	اتر (eter)	واژه	۹

				۱ - جسمی سیال و رقیق که قسمت فوقانی کره زمین را فراگرفته . ۲ - نمک فرار که از ترکیب یکی از اسیدهای معدنی یا آلی با الکل به دست می آید.	معنی	
			اطراب (etrâb)	اطراب (atrâb)	واژه	
			۱ - کسی را به طرب آوردن ۲ - سرود گفتن .	۱ - ج ترپ . ۱ - همسالان ، همزادان . ۲ - دوشیزگان .	معنی	۱۰
		اتم (atom)	عتم (atam)	اتم (atamm)	واژه	
		کوچکترین جزء یک جسم	زیتون .	مئامتر، کاملتر .	معنی	۱۱
			عصاره (osâre)	اثاره (asâre)	واژه	۱۲
			شیره ، چکیده ، فشرده .	۱ - انتقام . ۲ - برانگیختن .	معنی	
		اسلم (aslam)	اصلم (aslam)	اثلم (aslam)	واژه	۱۳
		سالم تر، درست تر، بی گزندتر .	گوش بریده .	رخنه دار، رخنه یافته .	معنی	
			اصواب (asvâb)	اثواب (esvâb)	واژه	۱۴
			صواب تر، درست تر، راست تر .	ج ثوب ؛ جامه ها .	معنی	
			عجر (ojar)	اجر (ajr)	واژه	۱۵
			۱ - عیب . ۲ - - حزن ، اندوه .	۱ - مزد، اجرت . ۲ - ثواب . ۳ - مزد دادن .	معنی	
			عجل (eji)	اجل (ajal)	واژه	۱۶
			گل و لای سیاه و بدبو . ج .	جلیل تر، بزرگوارتر .	معنی	

			عجله .		
			عجله (ajale)	واژه	
			۱ - شتاب کردن . ۲ - تعجیل ، تندی . ۳ - شتاب ، سرعت .	معنی	۱۷
			عجم (ajam)	واژه	
			۱ - نقطه گذاشتن . ۲ - نقطه .	معنی	۱۸
			عجمه (ajama)	واژه	
			۱ - یک هسته خرما . ۲ - یک تکسک انگور . ۳ - خرماين که از هسته روید؛ ج . عجمات .	معنی	۱۹
			اهباط (ehbât)	واژه	
			فرود آوردن ، هبوط دادن .	معنی	۲۰
			عهد (ahd)	واژه	
			۱ - شناختن امری . ۲ - حفظ کردن . ۳ - پیمان بستن . ۴ - پیمان ، ميثاق . ۵ - روزگار ، عصر ، دوره . ۶ - به گردن گرفتن و ملتزم شدن امری ، ضمان . ؛ عهد بوق یا دقیانوس کنایه از: زمان بسیار دور .	معنی	۲۱
			احسان (ehsân)	واژه	
			۱ - نیکي کردن . ۲ - بخشش کردن . ۳ - نیکوکاری	معنی	۲۲
			احسان (ahsân)	واژه	
			۱ - استوار و محکم کردن . ۲ - نگه داشتن نفس از انجام	معنی	

			بخشش .	کار بد. ۳ - شوی کردن زن ۴ - زن گرفتن مرد. ۵ - زن و مردی که به عقد دائم در آمده باشند که مرد «محسن» و به زن «محسنه » گویند.		
			اخص (axass)	اخس (axas)	واژه	
			خاص تر، گزیده تر.	خسیس تر، زبون تر، فرومایه تر، خوارتر.	معنی	۲۳
			عدات (odât)	ادات (adât)	واژه	
			ج عادی ؛ دشمنان .	افزار، ابزار، آلت ، وسیله .	معنی	۲۴
			عدله (adle)	ادله (adelle)	واژه	
			ج عادل ؛ کسانی که برای شهادت دادن شایسته باشند.	ج دلیل ؛ برهان ها، حجت ها.	معنی	۲۵
			ادیم (adim)	ادیم (adim)	واژه	
			نیست شده ، نابود.	۱ - چرم دباغی شده . ۲ - پوست خوشبوی سرخ رنگ . ۳ - روی زمین . ۴ - سفره غذا.	معنی	۲۶
			اضاعه (ezâe)	اذاعه (ezâ,a)	واژه	
			نک اضاعت .	آشکار ساختن ، فاش کردن .	معنی	۲۷
			ازالت (ezâlat)	اذالت (azâlat)	واژه	
			۱ - طرد کردن ، دور کردن . ۲ - زایل	فرو هشتن دامان ، دراز کردن دامن .	معنی	۲۸

			کردن ، از بین بردن .		
	عزل (azl)	اضل	ازل (azal)	اذل (azel)	واژه
۲۹	برکنار کردن ، از کار بازداشتن .	گمراه تر، به ضلالت تر.	زمان بی ابتداء.	ذلیل تر، خوارتر.	معنی
			عضله (azole)	اذله (azelle)	واژه
۳۰			ماهیچه .	۱ - ج ذلیل ؛ ذلیل شدگان . ۲ - ج ذلول ؛ نرم دلان .	معنی
			ازمه (azemma)	ازمه (azemma)	واژه
۳۱			ج . زمام ؛ مهارها، افسارها.	ج زمام و جج ذمه ؛ حق و حقوق .	معنی
			ازن (ozon)	اذن (ezn)	واژه
۳۲			ترکیبی است از اکسیژن به صورت ۳ خاصیت اکسیدکنندگی آن بسیار زیادتیر از اکسیژن است به طوری که نقره را اکسید می کند. به میزان کم در هوا موجود است و به صورت متراکم در لایة بالایی جو زمین وجود دارد. باکتری کش و رنگ زداست .	گوش .	معنی
			عزی (ozâ)	اذی (azâ)	واژه
۳۳			یکی از دو بت معروف طایفة قریش در عهد جاهلیت و دیگری «لات» نام داشته و اعراب بت پرست آن ها را دختران خدا می	۱ - آزار، رنج . ۲ - خس و خاشاک . ۳ - رنجش . ۴ - رنجه کردن ، رنجیدن . ۵ - پلیدی .	معنی

			دانستند.			
			عراده (arrâde)	erâde) اراده	واژه	
			۱ - از ابزارهای جنگ شبیبه به منجنیق که در قدیم برای پرتاب سنگ از آن استفاده می کردند. ۲ - واحد شمارش توپ .	۱ - خواستن ۲ - خواست ، میل ، آهنگ .	معنی	۳۴
			عرب (arab)	arab) ارب	واژه	
			تازی ، از نژاد عرب . ؛ به عرب عجمی بند نبودن کنایه از: هیچ پشت و پناهی نداشتن . ؛ از بیخ عرب بودن کنایه از: الف - مطلقاً انکار کردن . ب - هیچ ندانستن .	۱ - حاجت . ۲ - مقصود. ج . آراب .	معنی	۳۵
			عرج (araj)	arj) ارج	واژه	
			۱ - لنگ شدن ، لنگیدن . ۲ - لنگی .	ارز، ارزش ، رتبه ، مقام .	معنی	۳۶
		عرس (ors)	ers) ارث	aras) ارس	واژه	
		۱ - ستون ، ستون میان خیمه . ۲ - ریسمان . ۳ - گُره شتر خردسال .	۱ - سهم بردن از اموال شخص مرده . ۲ - آنچه از مال مرده که به بازماندگانش رسد، مرده ریگ ، مُردَر، ارثیه . ؛ ارث بابای کسی کنایه از: مال و ثروت شخصی کسی .	نک اروس .	معنی	۳۷
			عرش (arš)	aras) ارش	واژه	
			۱ - تخت ، تخت پادشاه . ۲ - خیمه ، چادر . ۳ - سقف ، سایبان . ۴ -	واحدی است از آرنج تا سر انگشت ؛ ذراع .	معنی	۳۸



			رکن . ۵ - آسمان نهم ، فلک الافلاک . ج . اعراش . عروش . عرشه .		
			عرضه (arze)	واژه (erze) ارضه	واژه
			۱ - به نمایش گذاشتن . ۲ - نمایش ، ارائه .	۱ - موریانه ، چوب خواره ، دیوچه ، دیوک . ۲ - زنگ آهن .	معنی
					۳۹
	عرق (arag)	عرق (erg)	ارغ (org)	اراق (arag)	واژه
	عرق بدن	۱ - رگ . ۲ - اصل ، ریشه . ج . اعراق و عروق .	نک آروغ .	رقیق تر ، تنگ تر ، شفاف تر ، باریک تر .	معنی
			ارغام (ergâm)	ارقام (argâm)	واژه
			۱ - به خاک مالیدن ، بینی کسی را به خاک مالیدن . ۲ - خوار کردن . ۳ - خشم گرفتن .	۱ - ج رقم ؛ خط ها ، نوشته ها . ۲ - اجناس ، مجموع چند بخش از کالا .	معنی
			عرم (arem)	ارم (eram)	واژه
			ج عرمه ؛ سدها ، بندها . باران های شدید .	نام باغی که شداد پسر عاد به تقلید از صفات بهشت درست کرده و دعوی خدایی نمود ، اما هنگام	معنی
					۴۲

			داخل شدن در آن جان باخت .		
			اروس (orūs)	اروس (arūs)	واژه
۴۳	معنی	کالا، متاع .	۱ - زن تازه شوهر کرده . ج . عرائس . ۲ - در فارسی : زن پسر . ۳ - بهترین ، زیباترین . ؛ عروس هزار داماد کنایه از: دنیا و بی وفایی آن .		
			اروق (arūg)	عروق (ūrūg)	واژه
۴۴	معنی	= اروغ . اوروغ . اوروق : خانواده ، دودمان ، خویشان .	ج عرق . ۱ - رگ ها . ۲ - ریشه ها .		
			اریش (ariš)	عریش (ariš)	واژه
۴۵	معنی	زیرک ، عاقل .	۱ - کجاوه ، هودج . ۲ - کلبه ، کومه . ۳ - چفت انگور .		
			اریکه (arika)	عریکه (arika)	واژه
۴۶	معنی	تخت ، سریر .	خلق و خوی .		
			از (az)	عز (ezz)	واژه
۴۷	معنی	۱ - علامت مفعول غیر - صریح یا باواسطه . ۲ - علامت ابتدا و آغاز . ۳ - در ، اندر . ۴ - برای ، بهر ، به سبب . ۵ - نسبت به ، در مقایسه با . ۶ - به دلیل ، به	۱ - ارجمند شدن ، عزیز شدن . ۲ - ارجمندی .		
			ازار (ezâr)	عذار (ezâr)	واژه
۴۸	معنی	۱ - زیرجامه ، شلوار . ۲ - دستار ، فوطه .	۱ - خط ریش . ۲ - در فارسی به معنی رخسار ،		

			چهره ۳ - لگام اسب .		
		عزت (ezat)	عزت (ezzat)	ازت (azot)	واژه
		۱ - نصیحت کردن ۲ - نصیحت .	۱ - سر بلند شدن ، گرامی شدن ۲ - سر بلندی ، سرافرازی .	نیتروژن ، گازی است بی رنگ و بی بو و بی مزه .	معنی
			ازهار (ezhâr)	ازهار (azhâr)	واژه
			۱ - پدیدار کردن ۲ - بیان کردن ۳ - آگاه کردن ۴ - قول ، گفتار . ج . اظهارات .	ج زهر و زهره ؛ شکوفه ها ، گل ها .	معنی
			عزیز (aziz)	ازیز (eziz)	واژه
			۱ - گرامی ، محبوب ۲ - ارجمند ، بزرگوار ۳ - کمیاب ، نادر .	۱ - به جوش آمدن ۲ - غلغل ، صدای جوشیدن دیگ . ۳ - بانگ رعد .	معنی
		عثار (esâr)	عصار (ossâr)	اسار (esâr)	واژه
		۱ - لغزیدن و افتادن ۲ - لغزش .	کسی که از دانه های روغنی روغن می گیرد .	۱ - اسیر کردن ۲ - چیزی را به بند کشیدن ۳ - اسیری ، بردگی ۴ - بند ، دوال .	معنی
			عصاری (ossâri)	اساری (esâry)	واژه
			روغن - گیری	ج اسیر ؛ اسیران .	معنی
			عصب (asab)	اسب (asb)	واژه
			پی ، رشته های سفید رنگ متصل به مغز که حس و	حیوانی است باهوش که برای سواری یا بارکشی به کار	معنی

			حرکت توسط آن ها صورت می گیرد.	گرفته می شود. ؛		
			(esbât) اثبات	(asbât) اسباط	واژه	
			۱ - ثابت گردانیدن . ۲ - نام نویسی در دفتر لشکر و سپاه .	ج سبط ؛ پسران پسر و پسران دختر.	معنی	۵۵
			استهلال (estehlâl)	استحلال (estehlâl)	واژه	
			۱ - ماه نو دیدن . ۲ - پدیدار شدن ماه نو.	حلالی طلبیدن .	معنی	۵۶
			استقص (ostogos)	استقص (estegas)	واژه	
			۱ - این کلمه در اصل یونانی است به معنی م اده و اصل هر چیزی ، عناصر چ ه ارگانه : آب ، خاک ، باد و آتش . ۲ - استخوان بندی هر چیز . ج . استقصات .	نک استقصات .	معنی	۵۷
			اصطلام (estelâm)	استلام (estelâm)	واژه	
			از بیخ برکندن چیزی را، از بن برکندن .	۱ - لمس کردن ، دست کشیدن به چیزی . ۲ - بوسه دادن .	معنی	۵۸
			اصهار (eshâr)	اسحار (eshâr)	واژه	
			به دامادی پیوستن .	ج سَحَر ؛ پاس آخر شب ، بامداد .	معنی	۵۹
			(esrâr) اصرار	(asrâr) اسرار	واژه	
			پافشاری کردن .	۱ - ج سِرُّ ؛ رازها، نهانی ها، سرها . ۲ - ج سَرَر و سُرَر خط ها و شکن	معنی	۶۰

				های کف دست	
			اتم (esm)	اسم (esm)	واژه
			۱ - گناه ، خطا. ۲ - کاری که کردن آن روا نباشد.	۱ - کلمه ای که به وسیله آن چیزی یا کسی را می خوانند، نام. ۲ - عنوان ۳ - شهرت ، آوازه . ، سفیدی ، هوش .	معنی
			اثمار (asmâr)	اسمار (esmâr)	واژه
			۱ - میوه آوردن درخت . ۲ - میوه دار شدن .	ج سَمَر.	معنی
			اثنان (esnân)	اسنان (esnân)	واژه
			۱ - دو، عدد دو. ۲ - دو مرد. ۳ - روز دوشنبه .	ج سین . ۱ - دندان ها. ۲ - دندانه داس . ۳ - تیزی مهرة پشت . ۴ - سال های زندگی .	معنی
	عسیر (asir)	عصیر (asir)	اثير (asir)	اسير (asir)	واژه
	سخت ، دشوار.	شیره و چکیده چیزی .	به باور قدما کره آتش که بالای کره هواست ، سیالی رقیق و بی وزن است .	۱ - گرفتار. ۲ - برده ، بنده . ج . اسراء.	معنی
			عشر (ašar)	اشر (ašr)	واژه
			یک دهم ، یک دهم چیزی .	شریرتر، بدتر، شرورتر.	معنی
			عشق (ešg)	اشق (ašag)	واژه
			۱ - به شدت دوست داشتن . ۲ - شیفتگی ، دلدادگی . ۳ - لذت ، کیف . ؛ عشق افلاطونی عشقی که با	دشوارتر، مشکل تر.	معنی

			گرایش های جنسی همراه نیست .			
			اشغر (ašgar)	اشقر (ašgar)	واژه	۶۷
			خارپشت .	۱ - سرخ موی . ۲ - اسبی که پال و دم آن سرخ باشد.	معنی	
			عصبه (asabe)	اصبح (asbah)	واژه	۶۸
			خویشاوندان شخص از سوی پدر	۱ - خوب رو، زیبارو. ۲ - مویی که سفید مایل به سرخ باشد	معنی	
			اسهاب (eshâb)	اصحاب (ashâb)	واژه	۶۹
			۱ - از اندازه گذشتن . ۲ - سخن را دراز گردانیدن . ۳ - پُرگویی ، اطناب ۴ - واگذاشتن چهارپایان سواری . ۵ - بسیار بخشش کردن . ۶ - شیر مکیدن بزغاله از مادر.	ج صاحب . ۱ - یاران ، دمسازان . ۲ - مالکان ، صاحبان .	معنی	
			اسعاد (es,âd)	اصعاد (es,âd)	واژه	۷۰
			۱ - نیکبخت گردانیدن . ۲ - مبارکی . ۳ - یاری ، کمک .	بالا رفتن ، صعود کردن .	معنی	
			عسل (asal)	اصل (asl)	واژه	۷۱
			مایع کم و بیش غلیظ و شربتی شکل و شیرینی که زنبوران عسل بر اثر جمع آوری نوش گل ها در کنند و به منظور تغذیه افراد کندو تهیه	۱ - ریشه ، بنیاد. ۲ - نژاد، گوهر.	معنی	

			کنند.			
			اصله (asle)	اصلاح (aslah)	واژه	
			۱ - یک درخت ، یک نهال . ۲ - واحد شمارش درختان .	۱ - نیکوتر، نکوکارتر. ۲ - شایسته تر.	معنی	۷۲
			اثیل (asil)	اصیل (asil)	واژه	
			محکم ، استوار.	۱ - پاک نژاد، نجیب . ۲ - شبانگاه . ج آصال .	معنی	۷۳
			عطالت (atâlat)	اطالت (etâlat)	واژه	۷۴
			بیکاری .	نک اطاله .	معنی	
			اتباع (atba,)	اطباع (etbâ,)	واژه	
			۱ - پیروی کردن ، در پی رفتن . ۲ - بازپس داشتن ، در پی فرستادن ۳ - واپس کردن ۴ - حواله کردن چیزی به کسی .	ج طبع ؛ سرشت ها، ذات ها.	معنی	۷۵
			اتلال (atlâl)	اطلال (atlâl)	واژه	
			ج تل ؛ توده های خاک و ریگ ، پشته ها.	ج طلل ؛ آنچه از عمارت و بنا، برپا مانده باشد.	معنی	۷۶
			عتیب (etib)	اطیب (atyab)	واژه	
			نک عتاب .	۱ - خوشبوتر. ۲ - حلال تر.	معنی	۷۷
	اذلال (ezlâl)	ازلال (azlâl)	اضلال (ezlâl)	اظلال (azlâl)	واژه	
	پیست شمردن ، خوار گرفتن .	لغزاندن ، لرزاندن .	گمراه ساختن ، به بیراهه افکندن .	ج ظل ؛ سایه ها.	معنی	۷۸
			ازهر (azhar)	اظهر (azhar)	واژه	
			۱ - روشن تر. ۲ - روشن ،	ظاهرتر، آشکارتر. ؛	معنی	۷۹

			درخشان . ۳ - ماه . ۴ - سفید رنگ .	اظهر من الشمس بسیار آشکار، کاملاً واضح .		
			اعظام (a,zâm)	اعزام (e,zâm)	واژه	
			۱ - بزرگ داشتن . ۲ - بزرگداشت .	در فارسی به معنای فرستادن ، روانه کردن .	معنی	۸۰
			اعصار (e,sâr)	اعصار (a,sâr)	واژه	
			۱ - تنگ دست شدن . ۲ - تنگ دستی .	ج عصر؛ روزگاران ، زمان ها، دورها.	معنی	۸۱
			اقتراب (agtarâb)	اغتراب (egterâb)	واژه	
			نزدیک شدن .	۱ - از دیار خویش دور شدن . ۲ - با بیگانگان ازدواج کردن .	معنی	۸۲
			عقور (agour)	اغور (ogor)	واژه	
			گزنده ، گاز گیرنده .	نک اُغر.	معنی	۸۳