



مقایسه میزان انطباق صورت مکتوب و ملفوظ نوشتار در زبان فارسی و انگلیسی

طاهره طارمی^۱

عادل رفیعی^۲

چکیده

در پژوهش حاضر، سعی بر تعیین میزان تطابق صورت مکتوب و ملفوظ نوشتار، از طریق بررسی تناظر نویسه- واج در زبان فارسی و مقایسه آن با زبان انگلیسی می‌باشد. برای نیل به این هدف، انواع روابط احتمالی ممکن میان واج و نویسه در شش گروه متمایز طبقه‌بندی و شماره‌گذاری می‌شوند. سپس، دو کتاب داستان کوتاه فارسی و انگلیسی را برگزیده و ۵۰۰۰ واژه از هر کتاب مورد بررسی قرار گرفته و هر نویسه در تناظر با واج معادلش قرار داده شده و با مقایسه آن‌ها، نوع رابطه نویسه و واج برای هر مورد تعیین می‌گردد. پس از تجزیه و تحلیل دقیق داده‌ها، بسامد و میزان فراوانی هر یک از روابط احتمالی در میان داده‌های زبان فارسی معین شده و با نتایج حاصل از داده‌های انگلیسی مقایسه می‌شوند تا مواردی که موجب عدم انطباق میان صورت مکتوب و ملفوظ نوشتار در زبان فارسی می‌شوند، مشخص شده و بسامد آن‌ها با موارد مشابه در زبان انگلیسی مورد سنجش قرار گیرد. در نهایت مشخص گردید که نارسایی‌هایی که برای نظام نوشتاری فارسی عنوان می‌شود، مختص این خط نیست و در نظام نوشتاری زبان انگلیسی هم موارد مشابهی وجود دارد و به طور کلی نبود رابطه یک به یک بین نویسه و واج در هر دو زبان مورد بررسی در میان واج‌های واکه‌ای به مراتب بیش از واج‌های همخوانی مشاهده شد و در هر زبانی این عدم تطابق به نحوی خود را می‌نماید. همچنین، میزان انطباق کامل صورت مکتوب و ملفوظ نوشتار نیز در زبان فارسی حدود یک و نیم برابر بیش از زبان انگلیسی رخ داده بود.

کلیدواژه‌ها: نظام نوشتاری، خط، نویسه، واج، نظام خوانشی

۱- ta_taremi@yahoo.com

۱- کارشناسی ارشد زبان شناسی همگانی (نویسنده مسئول)

۲- a.rafieei@fqn.ui.ac.ir

۲- استادیار گروه زبان شناسی دانشگاه اصفهان

۱. مقدمه

اصطلاح نظام نوشتاری (writing system)، به مجموعه‌ای از واحدهای نوشتاری و قراردادهای کاربرد آن‌ها برای نوشتن یک زبان اطلاق می‌شود (سمپسون^۱، ۱۹۸۵: ۲۱). به زبان دقیق‌تر، نظام نوشتاری یک زبان شامل مجموعه‌ای نشانه، مجموعه‌ای قاعدهٔ همنشینی نشانه‌ها و مجموعه‌ای از قواعد خواندن آن نشانه‌ها می‌باشد. طبقه‌بندی‌های گوناگونی برای نظام‌های نوشتاری مختلف ارائه شده است که دو نظام نوشتاری فارسی و انگلیسی در گروه نظام نوشتاری واجی واقع می‌شوند^۱. به این مفهوم که در این‌گونه نظام‌های نوشتاری هر یک از علائم نوشتاری بیانگر واج اعم از همخوان و واکه می‌باشد، با این تفاوت که نظام نوشتاری انگلیسی در زیر شاخهٔ الفبایی این گروه قرار گرفته و نظام نوشتاری فارسی که برگرفته از نظام نوشتاری عربی است، در زیر شاخهٔ همخوانی این گروه واقع می‌شود. شاید در ابتدای پیدایش نظام نوشتاری الفبایی فرض بر این بوده که هر نویسه معادل واجی واحد قرار گیرد و رابطه یک به یک میان نویسه و واج و در نتیجه، انطباق کامل میان صورت مکتوب و ملفوظ نوشتار وجود داشته باشد، اما در عمل شاهدیم که چنین وضعیتی حاکم نیست؛ به طوری که تناظر کامل میان نویسه و واج در نظام‌های نوشتاری وضعیتی نشان‌دار محسوب می‌شود. در دستور خط فارسی مصوب فرهنگستان زبان و ادب فارسی (۱۳۸۵)، عنوان می‌شود که خط چهره مکتوب زبان است و خط فارسی به موجب اصل پانزدهم قانون اساسی جمهوری اسلامی ایران، خط رسمی کشور می‌باشد و طبعاً این خط باید قواعد و اصولی معین و مدون داشته باشد تا همگان با رعایت آن‌ها هویت خط را تثبیت کرده و محفوظ دارند. در این دستور خط، هفت قاعدهٔ کلی برای خط ذکر می‌شود که تطابق صورت مکتوب و ملفوظ نیز یکی از این قواعد می‌باشد. هدف از تدوین این مقاله نیز تعیین میزان انطباق نویسه و واج در نظام‌های نوشتاری فارسی و انگلیسی و مقایسهٔ آنهاست. بدین منظور، در پی آنیم تا عوامل بر هم زنده انطباق بین صورت ملفوظ و مکتوب نوشتار در زبان فارسی و انگلیسی را معین کرده و میزان هر یک از این موارد در این دو زبان با هم مورد سنجش واقع شوند و از این رهگذر، درصد مواردی که انطباق کامل میان نویسه و واج وجود دارد، در نظام‌های نوشتاری فارسی و انگلیسی نسبت به یکدیگر تعیین گردد.

۲. پیشینهٔ پژوهش

اغلب کتاب‌های مقدماتی آموزش زبان‌شناسی از قبیل باقری (۱۳۸۱)، باطنی (۱۳۸۵، ۱۳۸۰) و حق‌شناس (۱۳۸۶) و همچنین کتاب‌هایی که مبحث تاریخ زبان را مطرح کرده‌اند، نظیر ناتل خانلری (۱۳۴۷) و ابوالقاسمی (۱۳۸۱)، به موضوع نظام نوشتاری و خط به صورت اجمالی و در کنار بحث اصلی زبان پرداخته‌اند

1. G.Sampson

اما شاید بتوان کتاب صفوی (۱۳۸۶) با عنوان «آشنایی با نظام‌های نوشتاری» را تنها کتاب مقدماتی به زبان فارسی دانست که صرفاً به معرفی نظام‌های نوشتاری، خط و اصطلاحات رایج در حوزه نظام نوشتاری می‌پردازد. صفوی، نویسه را واحدی از نظام نوشتاری در نظر می‌گیرد که به واحدی از زبان دلالت می‌کند و در تقابل با دیگر واحدهای نظام نوشتاری قرار می‌گیرد (همان: ۳۷). به اعتقاد صفوی "جدا از خطی که زبان‌شناسان برای نگارش واحدهای نظام آوایی زبان ساخته‌اند و الفبای آوانگار بین‌المللی نامیده می‌شود، هیچ نظام نوشتاری به شکل کامل واجی نیست" (همان: ۴۶). با این گفته، وی تلویحاً به وجود تناظر نسبی و نه یک به یک بین واحدهای نظام نوشتاری و واحدهای نظام آوایی در کلیه نظام‌های نوشتاری کاربردی امروزی اشاره می‌کند. برخی از آثار پژوهشگران ایرانی به بررسی دو جنبه مادی زبان، یعنی گفتار و نوشتار، به صورت مقایسه‌ای پرداخته‌اند. از این میان، می‌توان به عبدالتاج‌دینی (۱۳۸۴)، کلباسی (۱۳۸۰) و باطنی (۱۳۸۰ و ۱۳۸۵) اشاره کرد. عبدالتاج‌دینی زبان نوشتار را صورت‌های ملفوظی معرفی می‌کند که هنگام خواندن یک نوشته شنیده می‌شوند و کاملاً متفاوت از خط می‌باشند. این در حالی است که کلباسی (۱۳۸۰) زبان نوشتار را در مفهوم خط به کار می‌برد و باطنی (۱۳۸۰ و ۱۳۸۵) نیز خط و زبان نوشتار را از هم متمایز نکرده است. گروه دیگری از پژوهشگران ایرانی به آنچه نارسایی‌های خط فارسی نامیده می‌شود، پرداخته‌اند. در این میان، می‌توان از میرزا فتحعلی آخوندزاده (خامنه) (۱۲۷۳) به عنوان اولین فرد مدعی اصلاح الفبای خط فارسی نام برد (ذکاء، ۱۳۲۹) به نقل از ابوالقاسمی (۱۳۸۱: ۱۱۵). افراد دیگری نیز چون فرهاد میرزا معتمدالدوله (۱۲۹۲)، رحیم زاده صفوی (۱۳۲۸) و کسروی تبریزی (۱۳۵۷) تلاش زیادی در جهت تغییر خط فارسی اعمال کردند. در چارچوب زبان‌شناسی نوین، شاید بتوان مقاله باطنی (۱۳۸۵) را بارزترین اثر در میان طرفداران اصلاح یا تغییر خط فارسی دانست. با وجود این که افراد زیادی معایب خط فارسی را برشمرده و درصدد اصلاح خط برآمده‌اند اما برخی دیگر وجود چنین مشکلاتی را خاص خط فارسی ندانسته و آن را ویژگی مشترک تمام نظام‌های نوشتاری عنوان کرده‌اند. از میان این افراد می‌توان به ناتل خانلری (۱۳۴۷)، حق‌شناس (۱۳۸۶) و بهمنیار (۱۳۷۷) اشاره کرد. به اعتقاد خانلری "هیچ خط و نظام نوشتاری در دنیا وجود ندارد که صورت‌های گفتاری را به صورت کامل ثبت کند (صورت گفتار و نوشتار تطابق کامل داشته باشند)، به نحوی که خواننده از روی علامت‌های خط به صورت درست ملفوظ واژه دست یابد، زیرا در این صورت علامت‌های خط به اندازه‌ای فراوان می‌شد که آموختن و به کار بردن خط را دشوار می‌ساخت" (خانلری ۱۳۴۷: ۲۶۷). به باور حق‌شناس (۱۳۸۶) نیز "این نکته به صورت یک واقعیت عمومی پذیرفته شده است که همه الفباهای موجود عموماً دارای کمبودهایی هستند که با وجود آنها، از عهده ضبط تلفظ دقیق زبان‌ها بر نمی‌آیند" (حق‌شناس، ۱۳۸۶: ۲۳). بهمنیار (۱۳۷۷) نیز در پاسخ به عده‌ای که اساسی‌ترین اشکال نظام نوشتاری فارسی را عدم ظهور

مصوت‌های کوتاه می‌دانند، معتقد است "واژه‌های اندکی در زبان فارسی وجود دارد که درج مصوت‌های کوتاه باعث تمیز آنها از یکدیگر می‌شود".

توجه به خط و تمایز بین گفتار و نوشتار در آثار زبان‌شناسان غربی نیز مطرح شده است. از این میان می‌توان به سوسور^۱ (۱۹۵۹)، کولمس (۲۰۰۳) و راجرز (۲۰۰۵) اشاره کرد. با این حال، بیان دقیق ویژگی‌های نظام‌های نوشتاری در آثاری دیده می‌شود که اختصاصاً به خط و نظام‌های نوشتاری پرداخته‌اند. از این میان می‌توان به جلب (۱۹۶۳)، گاور (۱۹۸۵)، سمپسون (۱۹۸۵)، کولمس (۱۹۸۹)، اسپرات (۲۰۰۰)، فیشر (۲۰۰۱) و نانادسیکن (۲۰۰۹) اشاره کرد. در این میان جلب (۱۹۶۳) برای مدت زمان طولانی تنها کتاب پیرامون نظام‌های نوشتاری و ساختار آن‌ها بوده است (راجرز، ۲۰۰۵: ۸). کولمس (۱۹۸۹) تمام نظام‌های نوشتاری را دارای پیچیدگی خاصی می‌داند و علت این امر را وجود نشانه‌های کمتر در خط نسبت به واج‌های زبان ذکر می‌کند. به اعتقاد وی، هیچ یک از نظام‌های نوشتاری و خط‌های متداول امروزی نمی‌توانند گفتار را به صورت کامل و دقیق نشان دهند. راجرز (۲۰۰۵) به اصطلاحات مرتبط با نظام نوشتاری می‌پردازد. وی ایجاد نظام نوشتاری را به سه طریق ممکن می‌داند: روش اول، ابداع و اختراع است که بسیار به ندرت رخ می‌دهد و به نظر وی تاکنون تنها توسط سومری‌ها، چینی‌ها و اهالی مایا به کار گرفته شده است. روش دوم، قرض‌گیری از نظام نوشتاری دیگر است که بسیار متداول و مرسوم بوده و تقریباً تمام نظام‌های نوشتاری امروزی - به جز چینی - از آن استفاده کرده‌اند. روش سوم، ایده گرفتن از یک نظام نوشتاری موجود برای ساختن یک نظام نوشتاری جدید است. این روش در نظام‌های نوشتاری چروکی، کری و بلیس مشاهده می‌شود.

قوانین حرف - صدا که به عنوان قوانین نویسه - واج (grapheme phoneme rules) هم شناخته می‌شوند، سال‌های زیادی است که توجه محققین را به خود معطوف داشته است و حتی در دستورهای سنتی زبان‌های مختلف نظیر انگلیسی و فرانسوی یافت می‌شود. از جمله مطالعاتی که در این خصوص در زبان انگلیسی صورت گرفته، می‌توان به ونزکی (۱۹۷۰، ۱۹۹۹)، برنستین و نسلی (۱۹۸۱) و وایتل (۱۹۹۱) اشاره کرد. گونتيجو و همکاران (۲۰۰۳) نیز به تناظر واج - نویسه پرداخته و در انتهای مقاله خود جدول کاملی از نویسه‌های نظام نوشتاری انگلیسی و واج متناظر هر یک ارائه داده‌اند.

با توجه به این که مطالعات انجام شده توسط پژوهشگران ایرانی بر روی نظام نوشتاری زبان فارسی تنها از لحاظ نظری بوده و هیچ کدام از آن‌ها بر پایه داده و اطلاعات آماری صورت نگرفته است، به نظر می‌رسد انجام تحقیقی دقیق مبتنی بر تعداد قابل ملاحظه‌ای از داده‌ها در جهت تعیین میزان تناظر نویسه - واج از اهمیت بسزایی برخوردار باشد. بنابراین، در این مقاله سعی شده است با تجزیه و تحلیل رایانه‌ای، بسامد وقوع

۱. برای مطالعه بیشتر درباره رده‌شناسی خط ر.ک. (صفوی ۱۳۸۶)

واج‌های متناظر هر نویسه و نوع رابطه آن‌ها با هم در نظام نوشتاری زبان فارسی، تعیین شده و نسبت به نظام نوشتاری انگلیسی مورد سنجش واقع شود.

۳. شیوه تحقیق

هدف از پژوهش حاضر، بررسی میزان انطباق صورت مکتوب و ملفوظ نوشتار در زبان فارسی و انگلیسی و مقایسه آنها می‌باشد. بدین منظور، دو کتاب داستان کوتاه فارسی و انگلیسی به عنوان پیکره زبانی انتخاب شدند. از زبان فارسی کتاب «مدیر مدرسه» نوشته جلال آل احمد و از زبان انگلیسی کتاب «ناتوردشت»، نوشته سالینجر برگزیده شدند. جدیدترین نسخه الکترونیکی هر یک از کتاب‌ها ملاک عمل قرار گرفت و با توجه به فاصله‌ای که رایانه بین کلمات قرار می‌دهد، تعداد ۵۰۰۰ واژه ابتدایی هر کتاب استخراج شد. با توجه به تجزیه و تحلیل‌های مورد نیاز و امکانات برنامه اکسل (Excel)، این برنامه جهت ثبت داده‌های این تحقیق (۵۰۰۰ واژه) برگزیده شد و مورد استفاده قرار گرفت. درج داده‌ها نیز همان گونه که در جدول ۱ و ۲ نشان داده شده است، بدین صورت انجام پذیرفت که برای هر یک از کلمات، سه ردیف جداگانه از برنامه اکسل در نظر گرفته شد. در ردیف نخست، صورت نوشتاری کلمه به شکل حروف مجزا و جداگانه ثبت گردید. در ردیف بعدی، صورت ملفوظ کلمه درج گردید و دقت لازم به عمل آمد تا هر یک از واج‌ها دقیقاً متناظر نویسه‌اش قرار گیرد.

ک		چ	و	ک	صورت مکتوب
k	a	č	u	k	صورت ملفوظ
					انواع احتمال‌ها

جدول ۱: شیوه ثبت داده‌های زبان فارسی

Written form	l	i	t	t	l	e
Spoken form	l	I	t		l	
probabilities						

جدول ۲: شیوه ثبت داده‌های زبان انگلیسی

ردیف سوم از هر واژه نیز به احتمالاتی اختصاص یافت که نوع رابطه واج- نویسه را مشخص می‌سازند. در مورد تلفظ واژه‌های فارسی، در مواردی که تردید وجود داشت، فرهنگ ۸ جلدی سخن ملاک عمل قرار

گرفت. گفتنی است که در مورد برخی کلمات، چند تلفظ مختلف وجود داشت که در این صورت اولین تلفظ ذکر شده در فرهنگ لغت مد نظر قرار گرفت. در انگلیسی نیز در تمامی موارد فرهنگ لغت آکسفورد مورد استناد واقع گردید. نکته قابل توجه دیگر اینکه تلفظ هر کلمه به صورت مستقل و خارج از بافت (strong form) منظور شد. به عنوان نمونه، تلفظ واژه "and" در فرهنگ لغت آکسفورد به صورت /ænd/ واج‌نگاری شده است، اما صورت مستقل آن خارج از جمله /ænd/ می‌باشد که بدین صورت ثبت می‌شود. در گام بعدی، هر حرف با واج معادلش که در ردیف پائینی آن قرار گرفته، مورد بررسی قرار گرفت. در این مقایسه، حالت‌های متنوعی پیش می‌آید که این حالات در قالب شش احتمال مختلف به شرح زیر معرفی می‌شوند:

احتمال شماره ۱: یک نویسه بسیط (نویسه متشکل از یک واحد نوشتاری) معادل یک واج قرار می‌گیرد. این احتمال خود به سه زیر گروه تقسیم می‌شود که تحت عنوان خصیصه از آن‌ها نام می‌بریم:

خصیصه الف) یک نویسه بسیط همیشه و همه جا بیانگر یک واج است و آن واج معین، همیشه و همه جا تنها به وسیله همان نویسه مشخص، نمایش داده می‌شود. یعنی رابطه یک به یک میان نویسه و واج برقرار است. مانند نویسه <پ> در نظام نوشتاری زبان فارسی و نویسه <p> در نظام نوشتاری زبان انگلیسی که در همه مواضع در تناظر با واج /p/ قرار می‌گیرند.

خصیصه ب) یک نویسه بسیط معین، معادل یک واج قرار می‌گیرد، اما واج متناظر این نویسه در جایگاه‌های متفاوت یکسان نمی‌باشد. بدین مفهوم که یک نویسه خاص در مواضع مختلف به صورت واج‌های متنوعی تجلی‌آوایی می‌یابد. نویسه‌های <و> و <f> مثال‌هایی برای این خصیصه می‌باشند که در جایگاه‌های مختلف می‌توانند به ترتیب نشانگر واج‌های /v, /u و /f, /v باشند.

خصیصه ج) نویسه‌های مختلفی وجود دارند که تمام آن‌ها یک واج خاص را نشان می‌دهند. به زبان دیگر، یک واج معین در موضع‌های گوناگون توسط نویسه‌های متفاوتی نمایش داده می‌شود. مثال: نویسه‌های <ص>، <ث> که همگی مبین واج /s/ هستند. در نظام نوشتاری زبان انگلیسی نیز واج /k/ می‌تواند در جایگاه‌های مختلف به صورت نویسه‌های <q, k, c> تظاهر یابد.

احتمال شماره ۲: یک نویسه غیر بسیط (نویسه متشکل از چند واحد نوشتاری) معادل یک واج واحد است. به عنوان نمونه نویسه <ای> و <ch> که به ترتیب در تناظر با واج واحد /I/ و /č/ قرار می‌گیرند.

احتمال شماره ۳: حضور یک واج بدون این که نویسه‌ای در تناظر با آن قرار گیرد. مانند واکه‌های کوتاه در مواضع میانی کلمات در نظام نوشتاری زبان فارسی.

احتمال شماره ۴: وجود نویسه‌ای که نشانگر هیچ واجی نیست. نظیر نویسه <e> در موضع پایانی برخی از کلمات نظام نوشتاری زبان انگلیسی.

احتمال شماره ۵: یک نویسه بسیط، بیانگر توالی چند واج است. این واج‌ها می‌توانند یکسان یا متفاوت از هم باشند. برای مثال نویسه $\langle x \rangle$ در نظام نوشتاری انگلیسی بیانگر توالی واجی /ks/ می‌باشد و یا همخوان مشدد در نظام نوشتاری زبان فارسی را می‌توان نمونه‌ای برای این احتمال منظور کرد.

احتمال شماره ۶: یک نویسه غیر بسیط یا توالی چند نویسه، مبین توالی چند واج باشد. این حالت خاص تنها در صورتی لحاظ می‌شود که در موارد مشابه دیگر هم، چنین تناظری بتوان در نظر گرفت. اما در صورتی که این حالت تنها در یک مورد اتفاق افتاده باشد، برای آن احتمال ششم را نمی‌توان منظور کرد و بایستی به نحوی این توالی را با یکی از پنج حالت قبلی انطباق داد. به عنوان نمونه، توالی نویسه‌های $\langle xwa \rangle$ در اکثر موارد در تناظر با توالی واج‌های /xâ/ واقع می‌شود، که در این صورت احتمال شماره شش برای این مورد لحاظ می‌گردد.

۴. تجزیه و تحلیل داده‌ها

در این بخش پس از تعیین رابطه بین هر نویسه با واج متناظر آن در واژه‌های فارسی و انگلیسی، نوع رابطه آنها بر اساس احتمالات شش‌گانه تعیین می‌گردد. آن‌گاه نتایج به دست آمده در قالب نمودارهای توزیع فراوانی و جداول به دست داده می‌شوند.

۴-۱. بررسی میزان انطباق صورت مکتوب و ملفوظ نوشتار در داده‌های زبان فارسی

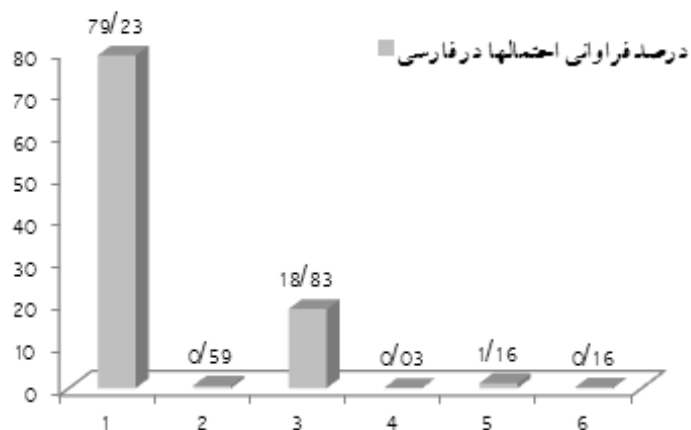
در این بخش، ابتدا با مقایسه هر نویسه و واج متناظرش که در ردیف پایین آن درج شده است، نوع رابطه آنها و در نتیجه شماره احتمال در نظر گرفته شده تعیین می‌گردد. در گام نخست، بسامد هر یک از احتمال‌های شش‌گانه معین شد که نتایج آن در بخش ۴-۱-۱ آمده است. سپس مواردی که احتمال شماره یک به آن‌ها تخصیص داده شده بود مورد تجزیه و تحلیل قرار گرفت و میزان رخداد خصیصه‌های سه‌گانه در بخش ۴-۱-۲ ارائه شده است.

۴-۱-۱. فراوانی احتمالات شش‌گانه در داده‌های زبان فارسی

با بررسی بسامد هر یک از احتمال‌های شش‌گانه، نتایج زیر حاصل شد که به صورت جدول و نمودار توزیع فراوانی نیز نشان داده شده است.

احتمالها	فراوانی احتمالها	درصد فراوانی
احتمال ۱	۱۷۳۲۵	۷۹/۲۳
احتمال ۲	۱۲۸	۰/۵۹
احتمال ۳	۴۱۱۸	۱۸/۸۳
احتمال ۴	۷	۰/۰۳
احتمال ۵	۲۵۴	۱/۱۶
احتمال ۶	۳۵	۰/۱۶
	۲۱۸۶۷	

جدول ۳: فراوانی احتمالها در داده‌های فارسی



نمودار ۱: درصد فراوانی احتمالهای شش‌گانه در داده‌های فارسی

همان‌طور که از نمودار پیداست، در نظام نوشتاری زبان فارسی احتمال شماره یک از درصد قابل ملاحظه‌ای برخوردار است، بدین صورت که حدود ۸۰ درصد از نویسه‌های مورد بررسی، نویسه‌های بسیطی هستند که متناظر با واجی واحد قرار می‌گیرند. پس از آن، احتمال شماره سه قرار می‌گیرد، یعنی واج‌هایی از زبان که در نظام نوشتاری توسط هیچ نویسه‌ای ظاهر نمی‌شوند که اکثر قریب به اتفاق آن‌ها واج‌های واکه‌ای

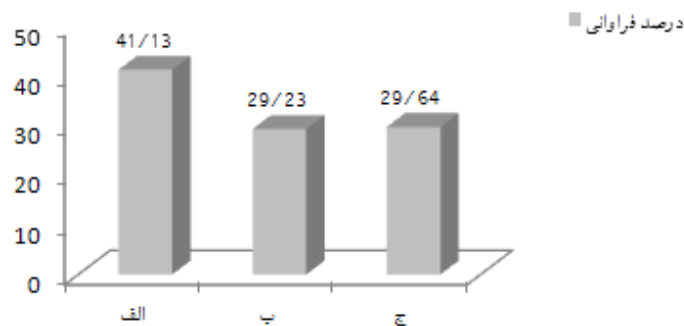
هستند. بقیه احتمال‌ها که بسامدشان بسیار اندک است، به ترتیب عبارتند از احتمال شماره پنج (نویسه بسیطی که معادل توالی چند واج قرار گرفته است)، احتمال شماره دو (نویسه غیربسیطی که در تناظر با واجی واحد واقع شده است) و احتمال شماره شش (توالی چند نویسه که معادل توالی چند واج قرار گرفته با لحاظ شرط تکرار) و کمترین میزان به احتمال شماره چهار (نویسه‌ای که معادل واجی در زبان ندارد) تعلق می‌گیرد.

۴-۱-۲. فراوانی خصیصه‌های احتمال شماره یک در داده‌های زبان فارسی

پس از تعیین فراوانی احتمال‌های شش‌گانه، در ادامه به خصیصه‌های احتمال شماره یک می‌پردازیم. در این مرحله، تنها موارد شامل احتمال شماره یک مورد تجزیه و تحلیل قرار می‌گیرند تا نوع خصیصه آنها مشخص گردد. ضمن بررسی دقیق داده‌ها مشخص گردید که تعداد ۱۴ نویسه از مجموع نویسه‌های نظام نوشتاری فارسی که همگی معادل واجی همخوانی می‌باشند، دارای خصیصه «الف» می‌باشند. ۴ نویسه هم شامل خصیصه «ب» هستند و تعداد ۶ تا از همخوان‌های زبان فارسی و ۴ تا از واژه‌های این زبان در گروه خصیصه «ج» جای می‌گیرند. به این ترتیب، با تعیین میزان فراوانی هر خصیصه می‌توان در جدول زیر به مقایسه آنها پرداخته و نمودار فراوانی آنها را نیز بدین صورت نشان داد.

احتمال‌ها	فراوانی احتمال‌ها	درصد فراوانی
خصیصه «الف»	۸۶۷۲	۴۱/۱۳
خصیصه «ب»	۶۱۶۲	۲۹/۲۳
خصیصه «ج»	۶۲۴۸	۲۹/۶۴
	۲۱۰۸۲	

جدول ۴: فراوانی خصیصه‌ها در داده‌های فارسی



نمودار ۲: درصد فراوانی خصیصه‌ها در داده‌های فارسی

چنانکه مشاهده می‌شود، در میان نویسه‌های بسیطی که در تناظر با واجی واحد قرار می‌گیرند، به طور تقریبی ۴۰ درصد از آن‌ها دارای رابطه یک به یک با هم هستند (خصیصه «الف») و مابقی دارای خصیصه «ب» و «ج» می‌باشند؛ بدین مفهوم که خصیصه هم‌نویسی (بروز یک نویسه واحد به صورت واج‌های مختلف) و خصیصه هم‌آوایی (نمایش یک واج واحد با نویسه‌های گوناگون) در حدود ۶۰ درصد موارد رخ می‌دهد.

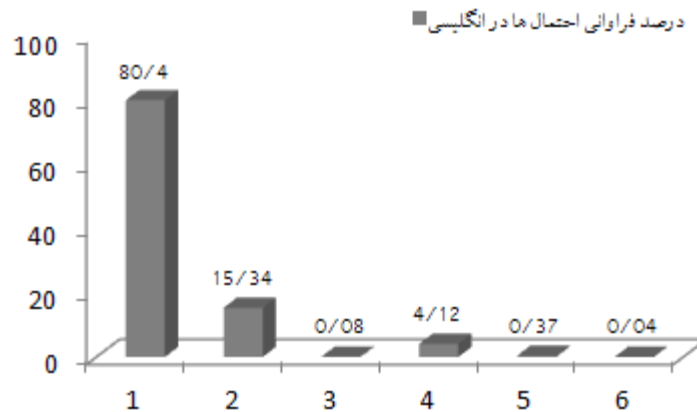
۴-۲. بررسی میزان انطباق صورت مکتوب و ملفوظ نوشتار در داده‌های زبان انگلیسی
در این بخش، تمامی مراحل که برای داده‌های فارسی انجام پذیرفت، برای داده‌های انگلیسی هم اعمال شد. بدین صورت که با مقایسه هر نویسه و واج معادلش، نوع رابطه و شماره احتمال آن مشخص گردید و سپس موارد شامل احتمال شماره یک به دقت بررسی شد تا نوع خصیصه‌های آن‌ها نیز تعیین گردد.

۴-۲-۱. فراوانی احتمالات شش‌گانه در داده‌های زبان انگلیسی

فراوانی هر یک از احتمال‌های شش‌گانه در میان داده‌های زبان انگلیسی به صورت زیر است که نتایج آن در جدول ارائه شده و نمودار درصد فراوانی آن هم به صورت زیر می‌باشد.

احتمال‌ها	فراوانی احتمال‌ها	درصد فراوانی
احتمال ۱	۱۳۲۷۷	۸۰/۰۴
احتمال ۲	۲۵۴۵	۱۵/۳
احتمال ۳	۱۴	۰/۰۸
احتمال ۴	۶۸۴	۴/۱۲
احتمال ۵	۶۱	۰/۳۷
احتمال ۶	۶	۰/۰۴
	۱۶۵۸۷	

جدول ۵: فراوانی احتمال‌های شش‌گانه در زبان انگلیسی



نمودار ۳: درصد فراوانی احتمال‌ها در داده‌های انگلیسی

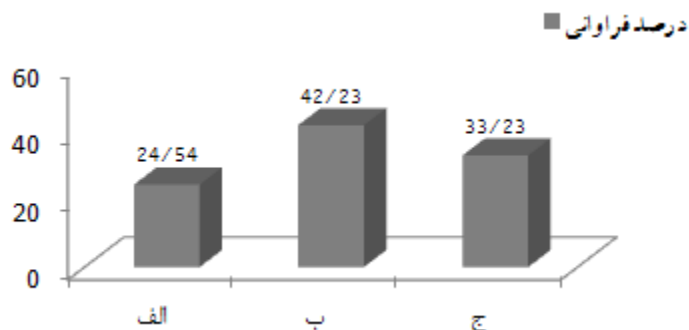
همانطور که از آمار جدول برمی‌آید، در میان داده‌های انگلیسی، بیشترین احتمال به احتمال شماره یک اختصاص می‌یابد؛ بدین مضمون که در بیش از ۸۰ درصد موارد مورد بررسی، یک نویسه بسیط در تناظر با واجی واحد قرار گرفت و پس از آن احتمال شماره دو است که حدود ۱۵ درصد موارد را شامل می‌شود. بقیه احتمال‌ها از میزان فراوانی بسیار کمی در مقایسه با این دو احتمال برخوردار می‌باشند و به ترتیب به صورت احتمال‌های شماره ۴، ۵ و ۳ هستند و احتمال شماره ۶ هم دارای کمترین میزان فراوانی است.

۴-۲-۲. فراوانی خصیصه‌های احتمال شماره یک در داده‌های زبان انگلیسی

در این بخش، موارد شامل احتمال شماره یک مورد تجزیه و تحلیل قرار گرفت تا نوع خصیصه هر یک تعیین شود. در این مرحله مشخص گردید که تعداد ۸ نویسه که هر یک معادل واجی همخوانی واقع می‌شوند، دارای خصیصه «الف» می‌باشند و ۱۱ نویسه از مجموع نویسه‌های نظام نوشتاری انگلیسی در گروه خصیصه «ب» جای می‌گیرند و ۱۷ واج زبان انگلیسی که اغلب واج واکه‌ای می‌باشند، شامل خصیصه «ج» می‌باشند. در ادامه، میزان فراوانی هر یک از خصیصه‌های سه‌گانه مدنظر مشخص گردید که نتایج حاصل از آن در قالب جدول و به صورت نمودار درصد فراوانی ارائه شده است.

درصد فراوانی	فراوانی	خصیصه‌ها
۲۴/۵۴	۴۶۲۷	خصیصه «الف»
۴۲/۲۳	۷۹۶۳	خصیصه «ب»
۳۳/۲۳	۶۲۶۵	خصیصه «ج»
	۱۸۸۵۵	

جدول ۶: فراوانی خصیصه‌ها در داده‌های انگلیسی



نمودار ۴: درصد فراوانی خصیصه‌ها در داده‌های انگلیسی

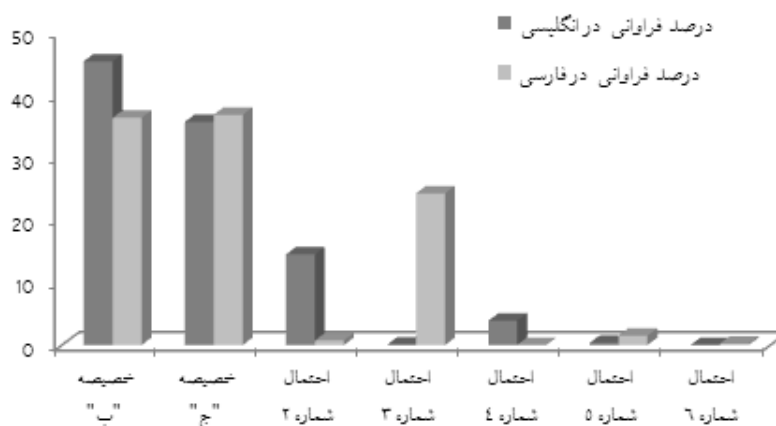
همان‌طور که مشاهده می‌شود، خصیصه «ب» دارای بیشترین میزان فراوانی است و پس از آن خصیصه-های «ج» و «الف» قرار می‌گیرند. بدین مفهوم که در میان داده‌هایی که یک نویسه بسیط معادل واجی واحد واقع می‌شود، در حدود ۴۰ درصد موارد، خصیصه هم‌نویسی وجود دارد و رابطه یک به یک میان نویسه و واج -خصیصه «الف»- در حدود ۲۵ درصد مواقع اتفاق می‌افتد.

۳-۴. مقایسه میزان انطباق صورت مکتوب و ملفوظ نوشتار در واژه‌های زبان فارسی و انگلیسی در این بخش با مطالعه و بررسی اطلاعات آماری به مقایسه میزان انطباق صورت مکتوب و ملفوظ نوشتار در زبان فارسی و انگلیسی پرداخته می‌شود. بدین منظور، ابتدا موارد عدم انطباق کامل نویسه و واج تعیین می‌شوند که همانا عبارتند از: احتمال‌های شماره ۲، ۳، ۴، ۵، ۶ و خصیصه‌های «ب» و «ج» از احتمال شماره

یک. تنها مورد تطابق کامل میان نویسه و واج نیز به خصیصه «الف» از احتمال شماره یک اختصاص می‌یابد. در زیر بخش‌های ۱-۳-۴ و ۲-۳-۴ به مقایسه این موارد در زبان‌های فارسی و انگلیسی پرداخته می‌شود. ۱-۳-۴. مقایسه میزان فراوانی موارد برهم زنده انطباق صورت مکتوب و ملفوظ نوشتار فارسی و انگلیسی

مورد برهم زنده انطباق صورت مکتوب و ملفوظ نوشتار	درصد فراوانی در فارسی	درصد فراوانی در انگلیسی
خصیصه «ب»	۳۶/۳۵	۴۵/۴
خصیصه «ج»	۳۶/۸۶	۳۵/۷۲
احتمال شماره ۲	۰/۷۵	۱۴/۵۱
احتمال شماره ۳	۲۴/۲۹	۰/۰۸
احتمال شماره ۴	۰/۰۴	۳/۹
احتمال شماره ۵	۱/۵	۰/۳۵
احتمال شماره ۶	۰/۲۱	۰/۰۳

جدول ۷: درصد فراوانی موارد برهم زنده انطباق صورت مکتوب و ملفوظ نوشتار فارسی و انگلیسی



نمودار ۵: مقایسه درصد فراوانی موارد برهم زنده انطباق صورت مکتوب و ملفوظ نوشتار فارسی و انگلیسی

با توجه به داده‌های آماری و مقایسه آن‌ها در دو زبان فارسی و انگلیسی، نتایج حاصله به این شرح می‌باشد:

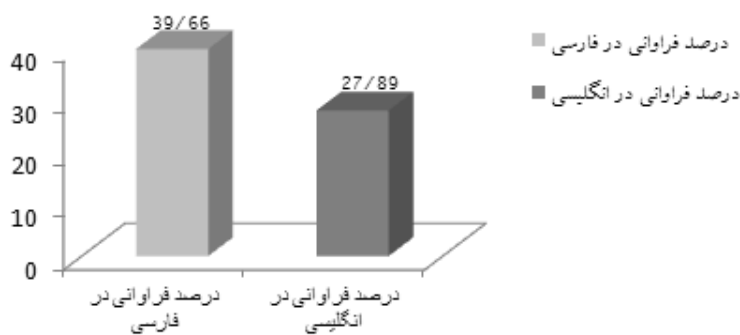
- الف- درصد فراوانی خصیصه «ب» در زبان انگلیسی به میزان $9/05$ درصد بیش از زبان فارسی می‌باشد.
- ب- درصد فراوانی خصیصه «ج» در زبان فارسی $1/14$ درصد بیشتر از زبان انگلیسی است.
- ج- درصد فراوانی احتمال شماره دو در زبان انگلیسی به میزان $13/76$ درصد در زبان انگلیسی بیش از زبان فارسی می‌باشد.
- د- درصد فراوانی احتمال شماره سه در زبان فارسی به میزان قابل توجهی و برابر $24/21$ درصد از زبان انگلیسی بیشتر است.
- ه- درصد فراوانی احتمال شماره چهار، $3/86$ درصد در زبان انگلیسی بیش از فارسی می‌باشد.
- و- احتمال شماره پنج در هر دو زبان به میزان کمی یافت می‌شود و در زبان فارسی به میزان $1/15$ درصد بیش از زبان انگلیسی است.
- ی- درصد فراوانی احتمال شماره شش نیز در زبان فارسی $0/18$ درصد بیشتر از زبان انگلیسی می‌باشد.
- بنابراین، به طور کلی می‌توان نتیجه گرفت که درصد فراوانی خصیصه «ج» و احتمال‌های شماره ۳، ۵ و ۶ در زبان فارسی بیش از زبان انگلیسی می‌باشد. و در مورد خصیصه «ب» و احتمال‌های شماره ۲ و ۴ زبان انگلیسی از درصد فراوانی بیشتری نسبت به زبان فارسی برخوردار است.

۴-۳-۲. مقایسه میزان فراوانی مورد انطباق کامل میان صورت مکتوب و ملفوظ نوشتار فارسی و انگلیسی

همان‌طور که قبلاً ذکر شد، از میان تمام روابط احتمالی که میان نویسه و واج در نظر گرفته شده است، تنها خصیصه «الف» نشانگر رابطه یک به یک میان نویسه و واج و در نتیجه، انطباق کامل صورت ملفوظ و مکتوب می‌باشد. میزان فراوانی این خصیصه در میان داده‌های مورد بررسی زبان فارسی برابر 8672 و در زبان انگلیسی به میزان 4627 مورد می‌باشد. حال، چنانچه این میزان فراوانی‌ها نسبت به فراوانی کل احتمال‌های موجود در پژوهش سنجیده شوند و درصد آن‌ها تعیین گردد، مشخص می‌شود که در میان داده‌های تحت بررسی، از کل احتمال‌های در نظر گرفته شده در زبان فارسی حدود 40 درصد و در زبان انگلیسی حدود 28 درصد رابطه یک به یک میان نویسه و واج و به عبارتی تطابق کامل میان صورت مکتوب و ملفوظ نوشتار وجود دارد.

درصد فراوانی خصیصه «الف»	
در زبان فارسی	۳۹/۶۶
در زبان انگلیسی	۲۷/۸۹

جدول ۸: درصد فراوانی خصیصه «الف» در انگلیسی داده‌های فارسی و انگلیسی



نمودار ۶: مقایسه درصد فراوانی خصیصه «الف» در داده‌های فارسی و انگلیسی

۵. نتیجه‌گیری

در این بخش، با تجزیه و تحلیل داده‌ها، سعی می‌شود نتایج در قالب یافته‌های پژوهش بیان گردند. با توجه به روابط احتمالی در نظر گرفته شده میان نویسه و واج، مشخص می‌شود که تنها خصیصه «الف» در احتمال شماره یک، رابطه یک به یک بین نویسه و واج را نشان می‌دهد و در این حالت، صورت ملفوظ و مکتوب نوشتار تطابق کامل دارند و خصیصه‌های «ب» و «ج» در احتمال شماره یک و تمامی احتمال‌های دو تا شش مواردی هستند که موجب عدم تناظر میان نویسه و واج و در نتیجه عدم انطباق میان صورت مکتوب و ملفوظ نوشتار در زبان فارسی و انگلیسی می‌شوند. همچنین، چنین می‌نماید که عدم تطابق میان نویسه و واج معادل آن در هر دو زبان تحت بررسی، در میان واژه‌ها بیشتر از همخوان‌ها یافت می‌شود و این رابطه در هر نظام نوشتاری خود را به نحوی می‌نماید؛ اکثراً در فارسی به صورت احتمال شماره ۳ و در انگلیسی به شکل خصیصه «ج» ظاهر می‌شود و نتیجه دیگر این که درصد فراوانی خصیصه «الف» در زبان فارسی ۱/۴۲ برابر بیش از زبان انگلیسی می‌باشد؛ بدین مفهوم که میزان انطباق کامل صورت ملفوظ و مکتوب نوشتار در زبان فارسی حدود یک و نیم برابر بیشتر از زبان انگلیسی است. بنابراین، طی پژوهش حاضر و با تجزیه و تحلیل

داده‌ها مشخص شد که مواردی که به عنوان کمبودها و معایب نظام نوشتاری فارسی عنوان می‌شود، کم و بیش در سایر نظام‌های نوشتاری از جمله نظام نوشتاری انگلیسی نیز یافت می‌شود. به نحوی که شاید بتوان ادعا کرد عدم انطباق صورت مکتوب و ملفوظ نوشتار نمی‌تواند به سبب نوع رده‌شناختی خط مورد استفاده زبان مورد نظر باشد، بلکه می‌توان گفت جزو ماهیت نظام‌های نوشتاری می‌باشد و به نوعی می‌توان آن را در زمره جهانی‌های زبان محسوب کرد.

منابع

- آل احمد، جلال، (۱۳۸۷)، *مدیر مدرسه*، تهران: آدینه سبز.
- ابوالقاسمی، محسن، (۱۳۸۱)، *تاریخ زبان فارسی*، تهران: سمت.
- انوری، حسن، (۱۳۸۱)، *فرهنگ بزرگ سخن*، تهران: سخن.
- باطنی، محمد رضا، (۱۳۸۰)، *پیرامون زبان و زبان‌شناسی*، تهران: فرهنگ معاصر.
- باطنی، محمد رضا، (۱۳۸۵)، *زبان و تفکر*، تهران: آگاه.
- باقری، مه‌ری، (۱۳۸۱)، *مقدمات زبان‌شناسی*، تهران: نشر قطره.
- بهمنیار، احمد، (۱۳۷۷)، *املائی فارسی*، تهران: دانشگاه تهران.
- حق شناس، علی محمد، (۱۳۸۶)، *آواشناسی (فونتیکی)*، تهران: آگاه.
- دستور خط فارسی؛ مصوب فرهنگستان زبان و ادب فارسی (۱۳۸۵)، تهران: فرهنگستان زبان و ادب فارسی (نشر آثار).
- ذکاء، یحیی، (۱۳۳۹)، *در پیرامون تغییر خط فارسی*، تهران: بی‌جا.
- صفوی، کورش، (۱۳۸۶)، *آشنایی با نظام‌های نوشتاری*، تهران: پژوهاک کیوان.
- عبدالتاج‌دینی، صادق، (۱۳۸۴)، *مقایسه زبان گفتار و زبان نوشتار فارسی معاصر در چارچوب واج-شناسی*، پایان نامه کارشناسی ارشد، دانشگاه اصفهان.
- کلباسی، ایران، (۱۳۸۰)، *فارسی گفتاری و نوشتاری، فصلنامه فرهنگ*، بهار و تابستان، سال ۱۴، ش، ۲۷-۳۸: ۴۹-۶۸، انتشارات پژوهشکده علوم انسانی و مطالعات فرهنگی.
- ناتل خانلری، پرویز، (۱۳۴۷)، *زبان‌شناسی و زبان فارسی*، تهران: انتشارات بنیاد فرهنگ ایران.
- Bernstein, J. and L. Nessler. 1981. Performance comparison of component algorithms for the phonemicization of orthography. In *Proceeding of the*

19 th Annual Meeting, Stanford University. Association for Computational Linguistics.

- Coulmas, F. 1989. *The Writing Systems of the World*. Oxford: Blackwell.
- Coulmas, Florian. 2003. *Writing Systems: An Introduction to their Linguistic Analysis*. New York: Cambridge University Press.
- Fischer, S. R.. 2001. *A History of Writing*. London: Reaktion Books Ltd.
- Gaur, A. 1985. *A History of Writing*. New York: Charles Scribners Sons.
- Gelb, I. J. 1963. *A Study of Writing*. Chicago: University of Chicago Press.
- Gnanadesikan, Amalia E. 2009. *The Writing Revolution: Cuneiform to the Internet*. Wiley- Blackwell.
- Gontijo, P. F. D. et al. 2003. Grapheme- Phoneme Probabilities in English. *Behavior Research Methods, Instruments and Computers*. 35(1), 136-157.
- Crowther, J. 1996. *Oxford advanced learner's dictionary*. Oxford university press.
- Rogers, H. 2005. *Writing Systems, A Linguistic Approach*. Oxford: Blackwell.
- Salinger, J.D. 1958. *Catcher in the rye*. London: Hamish Hamilton.
- Sampson, Geoffrey. 1985. *Writing Systems: A Linguistic Introduction*. London: Hutchinson.
- Saussure, Ferdinand de. 1959. *Course in General Linguistics*. Traslated by Wade Baskin. New York: The Philosophical Library.
- Sproat, Richards. 2000. *A Computational Theory of Writing Systems*. Cambridge: Cambridge University Press.
- Venezky, Richard L. 1999. *The American way of spelling*. New York: Guildford Press.
- Venezky, Richard L. 1970. *The Structure of English Orthography*. The Hague: Mouton.
- Vitale, A. J. 1991. An algorithm for high accuracy name pronunciation by parametric speech synthesizer. *Computational Linguistics* 17(3).