

آسیب‌شناسی نگارشی زیرنویس‌ها در شبکه‌های تلویزیونی خبر و آی‌فیلم: پژوهشی پیکره‌بنیاد

شهرام مدرس خیابانی^۱

تاریخ پذیرش: ۹۶/۰۳/۰۹

تاریخ دریافت: ۹۵/۱۰/۱۷

چکیده

در این پژوهش سعی بر آن بوده است، نظام نوشتاری موجود در زیرنویس‌های شبکه معیار خبر و زیرنویس‌های موجود در شبکه آی‌فیلم از منظر پیکره‌بنیاد بررسی شود تا ضمن دریافت نوع نگارش غالب در این زیرنویس‌ها، اشکالات املائی و فنی موجود در آنها، بر اساس دستورالعمل نگارش مصوب فرهنگستان زبان و ادب فارسی، بررسی و مقایسه شود. برای دستیابی به این هدف ابتدا پیکره‌ای شامل ۱۰۰۰ جمله از زیرنویس‌های خبری شبکه خبر و زیرنویس پیام‌های مردمی از شبکه آی‌فیلم در بازه زمانی مهرماه ۱۳۹۳ تا فروردین ماه ۱۳۹۴ آماده شد و سپس به کمک ابزار رایانه‌ای، فراوانی واژه‌ها و عبارت‌های موجود در زیرنویس‌ها استخراج شد. با روش تحلیل محتوا اشکالات املائی / سبکی، خطای حذف نیم‌فاصله، خطاهای حذف فاصله، خطای درج فاصله، سرهم‌نویسی و جدانویسی بررسی شد. یافته‌ها نشان داد، علاوه بر وجود انواع خطای نگارشی در هر دو شبکه، فراوانی خطاها در زیرنویس‌های شبکه آی‌فیلم بسیار زیاد است. بررسی میزان خطاهای نگارشی و مقایسه این نوع خطاها در هر یک از دو شبکه نشان می‌دهد که میان انواع خطاها به لحاظ نوع و فراوانی در بیشتر این جمله‌ها اختلاف معنی‌داری وجود دارد.

واژه‌های کلیدی

آسیب‌شناسی نظام نوشتاری، شبکه خبر، شبکه آی‌فیلم، پیکره زبانی، تلویزیون

۱. مقدمه

خط فارسی یکی از ابزارهایی است که نقشی غیرقابل انکار در اشاعه فرهنگ ایرانی و زبان و ادبیات فارسی دارد؛ اما نکاتی مانند نبود نویسه برای برخی واژه‌ها، نمایش یک آوا با چند نویسه، تنوع املائی و مشکلات سرهم‌نویسی یا جدانویسی، گاهی کاربران این نظام نوشتاری را با مشکل روبه‌رو می‌کند.

تلاش برای توسعه فرهنگی و حفظ زبان و ادبیات فارسی از یک‌سو و گسترش فضای مجازی و کاربرد روزافزون خط لاتین از سوی دیگر، اهمیت بررسی و آسیب‌شناسی خط فارسی را بیش‌ازپیش نشان می‌دهد. استفاده نکردن از الگویی معیار، استفاده از شکل‌های نامتعارف نوشتاری و همچنین کاربرد صورت‌های محاوره‌ای در نگارش فارسی‌زبانان به‌ویژه در شبکه‌های اجتماعی و زیرنویس‌های^۱ شبکه‌های مردمی از یک‌سو و همچنین بی‌توجهی به برخی ویژگی‌های معیار نگارشی در شبکه‌ای معیار چون شبکه خیر، نگارنده را بر آن داشت تا به بررسی مشکلات نگارشی در این شبکه‌ها بپردازد. از سویی دیگر، همان‌گونه که داوری اردکانی (۱۳۸۵) در یک نظرسنجی نشان داده است، سازمان صداوسیما جمهوری اسلامی ایران یکی از مهم‌ترین سازمان‌های مؤثر در حفظ و گسترش زبان و خط فارسی است که می‌تواند با تأکید بر یک نظام نوشتاری معیار، بسیاری از مشکلات املائی کاربران را برطرف کند.

بر این اساس در این پژوهش سعی بر آن است تا به بررسی نظام نوشتاری موجود در زیرنویس‌های حاوی خبر و پیام‌های مردمی پرداخته شود تا ضمن بررسی نوع نگارش غالب در این زیرنویس‌ها به بررسی اشکالات املائی و فنی موجود در آنها پرداخته و سازوکارهایی به‌منظور کاهش این اشکالات ارائه شود. زیرنویس‌های خبری به علت کاربرد زبان رسمی و پیام‌های مردمی به سبب کاربرد زبان غیررسمی و بررسی تمایز اشکالات بین این دو گونه در نوشتار انتخاب می‌شوند.

داده‌های پژوهش از پیکره زبانی^۲ استخراج شده است و مبنای تحلیل داده‌ها در این پژوهش دستورالعمل نگارش مصوب فرهنگستان زبان و ادب فارسی با عنوان «دستور خط فارسی» است.

۱. subtitle

۲. corpus

در ادامه پس از ارائه پیشینه‌ای از پژوهش‌های صورت گرفته در این زمینه و توصیف مبنای پژوهش، به نحوه گردآوری داده‌ها و تحلیل آنها پرداخته خواهد شد.

۲. پیشینه پژوهش

پژوهش پیش‌رو به بررسی ویژگی‌های املائی و فنی غالب در زیرنویس‌های موجود در برنامه‌های سیما اختصاص دارد. بر این اساس در این بخش ضمن اشاره به پژوهش‌های انجام شده درباره خط فارسی به پژوهش‌های انجام شده درباره زبان در صداوسیما خواهیم پرداخت و در نهایت پیشینه‌ای مختصر از پژوهش‌های مربوط به زیرنویس‌ها و به‌ویژه زیرنویس‌ها در سیما جمهوری اسلامی ایران، ارائه خواهد شد.

۲-۱. پژوهش‌های صورت گرفته در زمینه خط فارسی

درباره خط فارسی و مشکلات آن از دیرباز پژوهش‌های متعددی صورت گرفته است. در بیشتر این پژوهش‌ها به مشکلات خط فارسی پرداخته شده و سازوکارهایی برای رفع مشکلات آن ارائه شده است. در این میان، آثار اعتصام‌زاده (۱۳۰۴)، بهمن‌یار (۱۳۲۲)، خیام‌پور (۱۳۳۸)، حق‌شناس (۱۳۷۴)، هندی (۱۳۸۱)، کبیری (۱۳۸۶)، ذوالفقاری (۱۳۸۷)، راسخ‌نورد (۱۳۸۹) و فرهمندنژاد (۱۳۹۰) بر اصلاح در خط فارسی تمرکز کرده‌اند. دولت‌آبادی (۱۳۳۸) و نفیسی (۱۳۴۲) ضمن اشاره به این مشکلات به تغییر خط فارسی توجه کرده‌اند. پژوهشگرانی چون غلامحسین‌زاده (۱۳۷۹)، ضیا (۱۳۸۱) و اسلامی (۱۳۸۱)، به نقش مهم فرهنگستان زبان و ادب فارسی در برطرف کردن مشکلات خط و یکسان‌سازی نظام نگارشی تأکید کرده‌اند و در این خصوص سازوکارهای اساسی ارائه کرده‌اند. افرادی چون عرفانی (۱۳۸۷) معتقدند که تغییر خط منشأ پراکندگی، دگردیسی و نابودی بخش اعظم انباشت اندیشه و ثروت فرهنگی می‌شود. نصیری‌پور (۱۳۹۱) در تحلیلی پیکره‌بنیاد که پژوهش پیش‌رو به لحاظ روش‌شناختی شباهتی بسیار با آن دارد به گرایش گویشوران به جدانویسی یا سرهم‌نویسی واژه‌ها پرداخته و آن را با شیوه‌نامه مصوب فرهنگستان مقایسه می‌کند و نتیجه می‌گیرد که گرچه اغلب گویشوران فارسی‌زبان از الگوی فرهنگستان تبعیت می‌کنند مواردی وجود دارد که با شیوه‌نامه فرهنگستان همخوانی ندارد.

۲-۲. پژوهش‌های صورت گرفته درباره زبان در صداوسیما

پژوهش‌های صورت گرفته اوجی (۱۳۷۴)، ذاکری (۱۳۷۵)، شازاد (۱۳۷۵)، میرزایی (۱۳۷۸)، دمالی امیری (۱۳۸۰)، مهریار (۱۳۸۲)، بهمنی (۱۳۸۴)، مدرس خیابانی (۱۳۸۹) و مدرس خیابانی (۱۳۹۱) از جمله نمونه‌های مطرح درباره زبان در صداوسیما است. اوجی (۱۳۷۴) در پژوهش خود به بررسی واژه‌های قرصی اروپایی در خبر رسانه‌های گروهی ایران می‌پردازد. ذاکری (۱۳۷۵) به بررسی مسائل مربوط به اجرای نوشته‌های رادیویی و مشکلات گویندگان در فهمیدن و فهماندن مطالب می‌پردازد. شازاد (۱۳۷۵) به تفاوت زبانی زنان و مردان در شروع و پایان مکالمات حضوری در سریال‌های تلویزیونی پرداخته و نتیجه گرفته است که زبان مورد استفاده زنان و مردان هنگام آغاز و پایان مکالمه متفاوت است. میرزایی (۱۳۷۸) به مواردی در فیلم‌های دوبله اشاره می‌کند که با زبان معمول فارسی‌زبانان سازگار و هماهنگ نیست. دمالی امیری (۱۳۸۰) در پژوهش خود با عنوان «تبلور ایدئولوژی و قدرت در کلام (تحلیل انتقادی اخبار رسانه‌ای)» به این موضوع می‌پردازد که هیچ متنی خالی از بار ایدئولوژیک نیست و هر متن به منبع قدرت با اقتدار و نه لزوماً سیاسی مرتبط است. مهریار (۱۳۸۲) در پژوهش خود به میزان کاربرد واژگان مصوب فرهنگستان زبان در برنامه‌های سیمای جمهوری اسلامی ایران می‌پردازد. به اعتقاد وی بیشترین میزان کاربرد واژه‌های مصوب در برنامه‌های خبری به چشم می‌خورد. بهمنی (۱۳۸۴) در پژوهش خود به تأثیر برنامه‌های سیما بر رشد گفتار در کودکان و نوجوانان می‌پردازد. مدرس خیابانی (۱۳۸۹) زبان گزارشگران فوتبال سیما را تجزیه و تحلیل کرده است و ضمن ارائه نمونه‌هایی از هنجارگریزی‌های زبانی در این نوع گزارش‌ها، راهکارهایی به منظور کاهش این موارد ارائه کرده است. وی (۱۳۹۱) نیز با بهره‌گیری از ابزار پیکره‌بنیاد به بررسی میزان کاربرد دشواژه‌ها و رعایت نزاکت زبانی در برنامه‌های طنز سیما می‌پردازد و نتیجه می‌گیرد که با وجود درصد نسبتاً کم دشواژه‌ها در این برنامه‌ها، کاربرد چنین واژه‌ها و عبارت‌هایی از یک‌سو و کاربرد زیاد درخواست‌های مستقیم به جای درخواست‌های مؤدبانه، بدون آگاه کردن مخاطب از کاربرد آنها در این برنامه‌ها، نقش صداوسیما را در تربیت فرهنگی و اخلاقی جامعه و به‌ویژه نسل جوان، کم‌رنگ خواهد کرد.

۳-۲. پژوهش‌های انجام‌شده در زمینه زیرنویس و خط در صداوسیما

درباره زیرنویس می‌توان به برخی آثار پژوهشگران غیرایرانی و ایرانی اشاره کرد. لیند^۱ و کای^۲ (۱۹۹۹) در مقاله خود با عنوان «نشانه‌شناسی زیرنویس» مقدمه‌ای بسیار مفید برای آشنایی با ساختار و روش ترجمه مکتوب^۳ فیلم‌های سمعی - بصری ارائه می‌دهند. کوردلا^۴ (۲۰۰۶) در مقاله خود با عنوان «تحلیل گفتمان زیرنویس‌های فیلم‌های مستند: مطالعه موردی بچه‌های روسیه» به تجزیه و تحلیل زیرنویس‌های فیلم مستند «بچه‌های روسیه» می‌پردازد. به عقیده وی در فرایند تبدیل وسیله ارتباطی از گفتاری به نوشتاری باید به محدودیت‌های مکانی، زمانی و انطباق ساختارهای فرهنگی زبان مبدأ با زبان مقصد توجه شود.

برخی پژوهشگران ایرانی نیز به تحلیل زیرنویس‌ها پرداخته‌اند. جبارزاده و شهبها (۱۳۸۹) هدف از نگارش مقاله خود با عنوان «چالش‌های ترجمه فیلم» را بررسی نحوه عملکرد مترجمان در برخورد با موارد چالش‌برانگیز در فرایند ترجمه فیلم به‌خصوص زیرنویس می‌دانند. این دو پژوهشگر یکی از مزایای مهم زیرنویس را برطرف کردن ابهام موجود در آن با افزودن جمله یا حتی یک کلمه از جانب مترجم می‌دانند.

اما در زمینه آسیب‌شناسی خط فارسی در تلویزیون که به‌نوعی همسو با پژوهش حاضر است، می‌توان به آثار کشاورز و ستاری (۱۳۹۰) و سمائی (۱۳۸۸) اشاره کرد. کشاورز و ستاری تنوع املاء و عدم رعایت نیم‌فاصله را از مهم‌ترین مشکلات زیرنویس‌های تلویزیون می‌دانند. سمائی نیز در تحلیلی جامع ضمن توصیف مشکلات زبانی و فنی زیرنویس‌های شبکه خبر، سازوکارهایی را به‌منظور کاهش آن ارائه می‌دهد.

آنچه مسلم است این است که گرچه پژوهش‌هایی متعدد درباره زبان و خط فارسی در برنامه‌های سیما صورت گرفته است، در هیچ‌کدام تحلیل زیرنویس‌ها مبتنی بر رویکردی پیکره‌بنیاد^۵ انجام نشده است.

1. Z. D. Linde
2. N. Kay
3. translation
4. M. Cordella
5. corpus-based

۳. ادبیات نظری پژوهش

همان‌گونه که در بخش مقدمه پژوهش اشاره شد، در پژوهش به‌منظور دستیابی به ساختار املائی غالب در زیرنویس‌ها از دستاوردهای زبان‌شناسی پیکره‌ای^۱ استفاده می‌شود و ملاک ارزیابی نحوه نگارش زیرنویس‌ها، اصول مطرح در کتاب «دستور خط فارسی» مصوب فرهنگستان زبان و ادب فارسی (چاپ نهم، ۱۳۸۹) است. بنابراین ابتدا نکاتی درباره پیکره و زبان‌شناسی پیکره‌ای مطرح می‌شود و سپس به برخی اصول نگارشی مرتبط با پژوهش در «دستور خط فارسی» اشاره شده و درنهایت بر اساس الگوی مطرح‌شده، خطاهای نگارشی طبقه‌بندی خواهد شد.

۳-۱. پیکره و زبان‌شناسی پیکره‌ای

در زبان‌شناسی، به مجموعه‌ای از متون، پیکره گفته می‌شود؛ متونی که در پایگاه‌های الکترونیکی ذخیره شده است. پیکره‌ها معمولاً مجموعه بزرگی از متون قابل خوانش توسط ماشین را دربر دارند که حجم آن به هزاران یا میلیون‌ها واژه می‌رسد. پیکره از آرشیو متمایز است. در پیکره‌ها متن‌ها معمولاً به گونه‌ای انتخاب می‌شود که بیانگر گونه یا سبک زبانی خاص باشد و بدین‌ترتیب پیکره‌ها می‌توانند مرجعی معیار تلقی شوند. پیکره‌ها اغلب با اطلاعات اضافی مانند «برچسب‌های مقوله کلام^۲» یا حتی برخی ویژگی‌های نوایی گفتار برچسب‌دهی می‌شود (بیکر^۳ و دیگران، ۲۰۰۶: ۴۸-۴۹). پیکره‌ها انواع مختلفی دارند که در این میان می‌توان به پیکره‌های عمومی^۴، پیکره‌های تخصصی^۵، مرجع^۶، چندزبانه^۷، موازی^۸، زبان‌آموز^۹، در زمانی^{۱۰} یا تاریخی و فرابین یا ناظر^{۱۱} اشاره کرد (بیکر و دیگران، ۲۰۰۶: ۴۸-۴۹). پیکره‌هایی که تنها از متون زبانی تشکیل شده و تحلیل‌گر را در یافتن برخی سؤالات درباره واژه‌ها، دستور یا ساختار کلامی یک

۱. corpus linguistics

۲. part-of-speech tags

۳. P. Baker

۴. general corpora

۵. specialized

۶. reference

۷. multilingual

۸. parallel

۹. learner

۱۰. diachronic

۱۱. monitor

زبان کمک می‌کند، پیکره عمومی نامیده می‌شود (کندی، ۱۹۹۸: ۲۰). البته این نوع پیکره‌ها معمولاً شکل متعادل^۱ دارد، به این مفهوم که سبک‌های زبانی متفاوت و همچنین صورت گفتاری و نوشتاری را در برمی‌گیرد. شکل دیگری از پیکره‌ها وجود دارد که پیکره‌های تخصصی نامیده می‌شود. این پیکره‌ها در مطالعه موارد زبانی خاص همچون شکل‌گیری زبان کودک یا حتی زبان تخصصی مثلاً زبان مربوط به حوزه اکتشاف نفت، استفاده می‌شود (کندی، ۱۹۹۸: ۱۹-۲۰). از سوی دیگر برخی پیکره‌ها به‌منظور اهدافی خاص همچون تدوین فرهنگ لغت یا بررسی تفاوت‌های زبانی موجود در بافت‌های مختلف جغرافیایی، اجتماعی، تاریخی یا شغلی طراحی شده است (کندی، ۱۹۹۸: ۱۹-۲۰).

عاصی (۱۳۸۲) به نقل از توبرت^۲ (۱۹۹۹) در طبقه‌بندی انواع پیکره به این نکته اشاره می‌کند که پیکره‌ها را می‌توان از نظر اندازه و گستره به دسته‌های محدود، متوسط و عظیم تقسیم کرد. پیکره‌ها به دلایل متفاوت استفاده می‌شود و این مسئله به نوبه خود در طرح، اندازه و ماهیت پیکره‌ها اثر می‌گذارد. برخی پیکره‌ها به‌منظور توصیف ویژگی‌های عام زبان طراحی شده است و به کمک آنها می‌توان سطوح مختلف زبان همچون واژگان، دستور و الگوهای کلامی و کاربردشناختی را بررسی کرد (کندی، ۱۹۹۸: ۳). با توجه به اینکه 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 (عاصی، ۱۳۸۲: ۴). استابز^۳ (۲۰۰۰: ۸) به مقایسه پیکره‌ها بر اساس حجم آنها می‌پردازد و نتیجه می‌گیرد که تکیه بر یک نوع پیکره واحد امری پرمخاطره است و ترکیبی از پیکره‌های عمومی و تخصصی می‌تواند نتایج سودمندتر در بر داشته باشد. از سوی دیگر وی به این نکته اشاره می‌کند که حتی پیکره‌های کوچک می‌توانند در نشان دادن برخی الگوهای زبانی مؤثر باشند، اما تنوع واژگانی و ویژگی‌های کلامی عبارت‌های کم‌بسامد، توجه به پیکره‌های بزرگ زبانی را ضروری می‌سازد.

در امتداد استفاده از پیکره زبانی، شاخه‌ای در زبان‌شناسی موسوم به زبان‌شناسی پیکره‌ای پدید آمده است. از دیدگاه مکنری^۴ و هاردی^۵ (۲۰۱۲) زبان‌شناسی پیکره‌ای حوزه‌ای است که بر

۱. balanced

۲. W. Teubert

۳. M. Stubbs

۴. T. MCenery

۵. A. Hardie

روال‌ها یا روش‌هایی برای مطالعهٔ زبان تمرکز دارد. زبان‌شناسی پیکره‌ای کمک می‌کند تا نظریه‌ای دربارهٔ زبان ارائه شود یا طیفی از نظریه‌هایی دربارهٔ زبان بازبینی یا توصیف شود. امتیاز زبان‌شناسی پیکره‌ای و پژوهش‌های پیکره‌بنیاد این است که ما را در بررسی دقیق‌تر مواردی که تنها بر اساس درون‌نگری یا به‌عبارت‌دیگر شَمّ زبانی توجیه شده‌اند، یاری می‌کند. به کمک مطالعهٔ پیکره‌بنیادِ زبان طبیعی، می‌توان به اطلاعاتی دست یافت که در مباحثی چون یادگیری زبان، آموزش زبان، پردازش زبان طبیعی^۱، درک گفتار^۲ و ترجمه، نقشی اساسی ایفا می‌کند (کندی^۳، ۱۹۹۸: ۸).

لیچ^۴، ویژگی‌های مهم این رشته را چنین برمی‌شمارد:

۱. تمرکز بر کنش زبانی و نه توانش زبانی؛
 ۲. تمرکز بر توصیف زبانی و نه بر همگانی‌های زبان؛
 ۳. تمرکز بر الگوهای کمی زبانی همانند الگوهای کیفی آن؛
 ۴. تمرکز بر دیدگاه‌های تجربی (و نه عقلانی) در بررسی‌های علمی زبان؛
- همان‌گونه که مشاهده می‌شود، این ویژگی‌ها مجموعه‌ای را به وجود می‌آورد که توجه بیشتری به جنبه‌های رفتاری زبان و بروز طبیعی گفتار و نوشتار دارد و عملاً در مقابل دیدگاه‌های چامسکی و پیروان وی قرار می‌گیرد (عاصی، ۱۳۸۲: ۴).
- به عقیدهٔ بایبر^۵ (۱۹۹۹) نگرش پیکره‌بنیاد را می‌توان نگرشی با چهار جنبه برشمرد:
۱. با معرفی الگوهای متون طبیعی، سعی در نیل به اهداف تجربی خود دارد؛
 ۲. به مطالعهٔ متون طبیعی در پیکره‌های معتبر می‌پردازد؛
 ۳. بررسی پیکره‌ها توسط فن‌آوری ارتباطی رایانه‌ای صورت می‌گیرد؛
 ۴. با ترکیب روش‌های کمی و کیفی، تجزیه و تحلیل انجام می‌شود.

۲-۳. اصول نگارش در دستور خط فارسی

^۱. Natural Language Processing (NLP)

^۲. speech recognition

^۳. G. Kennedy

^۴. G. Leech

^۵. D. Biber

همان‌گونه که در بخش مقدمه نیز اشاره شد: در این پژوهش ملاک ارزیابی نحوه نگارش زیرنویس‌ها، کتاب «دستور خط فارسی» مصوب فرهنگستان زبان و ادب فارسی (چاپ نهم، ۱۳۸۹) است. در کتاب «دستور خط فارسی» ضمن اشاره به حروف زبان فارسی و نحوه تلفظ آنها، حدود ۴۷ قاعده درباره سرهم‌نویسی و جدانویسی واژه‌ها و همچنین املاهای صحیح برخی کلمات نیز اشاره شده است. برخی از این قواعد که به پژوهش حاضر ارتباط دارد به شرح زیر است (دستور خط فارسی: ۲۲-۴۳):

- «این» و «آن» جدا از جزء و کلمه پس از خود نوشته می‌شود به‌استثنای آنچه، آنکه، اینکه، اینها، آنها، اینجا، آنجا، وانگهی.
- «همین» و «همان»، همواره جدا از کلمه پس از خود نوشته می‌شود.
- «هیچ»، همواره جدا از کلمه پس از خود نوشته می‌شود.
- «چه»، جدا از کلمه پس از خود نوشته می‌شود، مگر در چرا، چگونه، چقدر، چطور، چسان.
- «چه» همواره به کلمه پیش از خود می‌چسبد.
- «بی»، همیشه جدا از کلمه پس از خود نوشته می‌شود، مگر آنکه کلمه بسیط‌گونه باشد، یعنی معنای آن دقیقاً مرکب از معنای اجزای آن نباشد؛ مانند بیهوده، بیخود، بیراه، بیچاره، بینوا، بیجا.
- «می» و «همی»، همواره جدا از کلمه پس از خود نوشته می‌شود.
- «هم»، همواره جدا از کلمه پس از خود نوشته می‌شود، مگر کلمه بسیط‌گونه باشد؛ مانند همشهری، همشیره. - «هم»، همواره جدا از کلمه پس از خود نوشته می‌شود، مگر جزء دوم تک‌هجایی باشد؛ مانند همکار یا همراه.
- «تر» و «ترین»، همواره جدا از کلمه پیش از خود نوشته می‌شود، مگر در بهتر، مهتر، کهتر، بیشتر، کمتر.
- «ها» (نشانه جمع) در ترکیب با کلمات به هر دو صورت (پیوسته و جدا) صحیح است اما هرگاه «ها» بعد از واژه‌های بیگانه نامأنوس به کار رود جدانویسی الزامی است؛ مانند پزیتیویست‌ها، فرمالیست‌ها.
- هنگامی که بخواهیم اصل کلمه را برای برجسته‌سازی مشخص کنیم جدانویسی «ها» الزامی است؛ مانند کتاب‌ها، باغ‌ها، متمدن‌ها، ایرانی‌ها.
- هرگاه کلمه پردندانه (بیش از سه دندانه) شود یا به «ط» و «ظ» ختم شود جدانویسی الزامی است؛ همچون پیش‌بینی‌ها، حساسیت‌ها، استنباط‌ها، تلفظ‌ها.

- هرگاه جمع اسامی خاص مدنظر باشد جدانویسی الزامی است؛ مانند سعدی‌ها، فردوسی‌ها، مولوی‌ها، هدایت‌ها.

- اگر واژه به «های غیرملفوظ» ختم شود جدانویسی الزامی است؛ مانند میوه‌ها، خانه‌ها.

- هرگاه واژه به «های ملفوظی» ختم شود که حرف قبل از آن حرف متصل باشد، جدانویسی الزامی است؛ برای نمونه سفیه‌ها، فقیه‌ها، پیه‌ها، به‌ها.

- کلماتی که از ترکیب با پسوند ساخته می‌شود همیشه پیوسته نوشته می‌شود، مگر هنگامی که حرف پایانی جزء اول با حرف آغازی جزء دوم یکسان باشد؛ مانند نظام‌مند، آب‌بان.

- کلماتی که از ترکیب با پسوند ساخته می‌شود همیشه پیوسته نوشته می‌شود، مگر هنگامی که جزء اول آن عدد باشد؛ مانند پنج‌گانه، ده‌گانه، پانزده‌گانه.

- پسوند «-وار» از حیث جدا یا پیوسته‌نویسی تابع قاعده‌ای نیست، در بعضی کلمه‌ها جدا و در بعضی دیگر پیوسته نوشته می‌شود:

طوطی‌وار، فردوسی‌وار، طاووس‌وار، پری‌وار، بزرگوار، سوگوار، خانوار.

آنچه مسلم است این است که حتی «دستور خط فارسی» نیز دارای مشکلاتی است که در اینجا به نکاتی اشاره می‌شود:

الف. در مواردی گفته شده است که هم سرهم‌نویسی و هم جدانویسی مجاز است؛ مانند بحث «-ها».

ب. در دستور خط فارسی صرفاً بحث واحدهای نوشتاری مطرح است و تفاوت‌های نقشی یا معنایی لحاظ نشده است. بدیهی است میان ضمیر «چه» در «آنچه» و پسوند اشتقاقی «-چه» در «کتابچه» به لحاظ صرفی، تفاوت اساسی وجود دارد.

پ. به نظر می‌رسد مؤلفان «دستور خط فارسی» در بسیاری از موارد تمایزی میان واژه‌های مشتق و مرکب قائل نبوده‌اند و همه واحدهای زبانی را که بعد از واحد اول آمده است را پسوند در نظر گرفته‌اند. به همین علت است که هم تکواژ دستوری مانند «-وار» (که پسوند است) و هم تکواژهای واژگانی مانند «خوار» یا «ریا» (که پسوند نیستند) را «پسوند» تلقی کرده‌اند.

ت. به برخی موارد مانند جدانویسی اعداد یا جدانویسی آواهای هم‌مخرج بیش از یک‌بار اشاره شده است.

ث. با توجه به نکات ذکر شده در بالا تصحیح «دستور خط فارسی» ضروری به نظر می‌رسد. در این میان می‌توان به موضوع فاصله مجازی یا نیم‌فاصله اشاره کرد که امروزه نقشی مهم در

نگارش رایانه‌ای خط فارسی دارد؛ موضوعی که در بحث فاصله‌گذاری (دستور خط فارسی، ۱۳۸۹: ۱۰)، به اختصار، به آن اشاره شده است؛ بنابراین انتظار می‌رود که در ویراست جدید «دستور خط فارسی» به این موضوع‌ها بیشتر پرداخته شود.

۳-۳. طبقه‌بندی خطاهای نگارشی

در این پژوهش با الگوبری از دستور خط فارسی و اساساً شکل صحیح نوشتاری واژه‌ها، خطاها در شش گروه اشکال املائی/سبکی، درج فاصله، حذف فاصله، حذف نیم‌فاصله، سرهم‌نویسی و جدانویسی بررسی شد که در اینجا به توصیف هر یک از انواع این خطاها خواهیم پرداخت:

الف. اشکال املائی/سبکی: منظور از اشکال املائی این است که واژه موردنظر به لحاظ املائی نادرست باشد. برای نمونه نوشتن «متچکرم» یا «مچکرم» به جای «متشکرم» اشکال املائی محسوب می‌شود، اما منظور از اشکال سبکی این است که گرچه واژه موردنظر به لحاظ نوشتاری غلط نیست به سبک شبکه موردنظر که در اینجا بیشتر شبکه رسمی خبر مدنظر است، نامرتب است. برای نمونه کاربرد «جیگرم» به جای «جگرم» در شبکه خبر اشکال سبکی محسوب شده است.

به‌منظور حفظ تمام موارد ذکر شده در دستورالعمل فرهنگستان، حتی مواردی که علامت تنوین یا حتی همزه حذف شده است نیز به‌عنوان اشکال املائی در نظر گرفته شده است. ب. درج فاصله: این نوع خطا به مواردی مربوط می‌شود که نیاز به درج فاصله نبوده ولی نویسنده فاصله گذاشته است؛ برای نمونه پیش از علائم سجاوندی مانند «:» یا «،» نباید فاصله‌ای درج شود و در صورت درج فاصله، خطای درج فاصله لحاظ شده است.

پ. حذف فاصله: این مورد عکس مورد ب است. در این حالت به‌جای درج فاصله، فاصله برداشته شده است؛ نمونه «باعرض» به‌جای «با عرض»، از این دست است.

ت. حذف نیم‌فاصله: بر اساس دستورالعمل فرهنگستان برخی پسوندها و پیشوندها به کمک نیم‌فاصله بعد یا قبل از واژه پایه نوشته می‌شوند. این مسئله در خصوص واژه‌های مرکبی که اجزای آنها باید جدا نوشته شود نیز صدق می‌کند. به کمک نیم‌فاصله یا فاصله مجازی، هم جدا بودن واژه مشخص می‌شود و هم نشان داده می‌شود که بخش موردنظر، بخشی از خود واژه است. پیشوند «می-» یا پسوند «-ها» از جمله مواردی هستند که به دلیل درج فاصله، به‌عنوان

واژه‌ای مستقل در پیکره مشاهده شده‌اند و می‌توانند نمونه‌هایی از خطای حذف نیم‌فاصله محسوب شوند.

ث. سرهم‌نویسی: این نوع خطا به مواردی اطلاق شده است که باید جدا نوشته می‌شدند ولی سرهم نوشته شده‌اند. برای نمونه پیشوند تصریفی «می-» باید جدا نوشته شود که در نمونه «میگیرد» سرهم نوشته شده است.

ج. جدانویسی: این مورد به مواردی مربوط می‌شود که باید سرهم نوشته شوند ولی جدا نوشته شده‌اند. برای نمونه اگر واژه «همشهری» به صورت «هم‌شهری» نوشته شود، خطای جدانویسی رخ داده است. نکته قابل توجه این است که در هیچ‌یک از دو پیکره بررسی شده حتی یک مورد خطای جدانویسی مشاهده نشده است.

نکته بسیار مهم این است که ممکن است در یک نمونه شاهد وقوع بیش از یک خطا باشیم. برای نمونه درحالی‌که در صورت نوشتاری «میگیره»، خطای سرهم‌نویسی رخ داده است، کاربرد آن در پیکره خبر، اشکال سبکی نیز محسوب می‌شود و در نتیجه در بخش اشکالات املائی/ سبکی نیز به آن اشاره شده است.

در بخش بعدی به ترتیب به تحلیل خطاهای پیکره خبر، پیکره آی‌فیلم و مقایسه خطاهای دو پیکره خواهیم پرداخت.

۴. روش پژوهش

روش این پژوهش رویکردی پیکره‌بنیاد است و مبنای تحلیل داده‌ها بر اساس مقایسه نظام نگارشی زیرنویس‌ها با شیوه‌نامه «دستور خط فارسی» مصوب فرهنگستان زبان و ادب فارسی (چاپ نهم، ۱۳۸۹) است. بر این اساس ابتدا لازم است پیکره‌ای از زیرنویس‌ها در دوره زمانی مهرماه ۱۳۹۳ تا فروردین‌ماه سال ۱۳۹۴، شامل ۱۰۰۰ جمله از زیرنویس‌های خبری شبکه خبر و زیرنویس پیام‌های مردمی از شبکه آی‌فیلم را آماده کرده و سپس به کمک ابزارهای رایانه‌ای، فراوانی واژه‌ها و عبارت‌های موجود در زیرنویس‌ها را استخراج کنیم و در نهایت به بررسی مشکلات موجود در این تولیدات نوشتاری بپردازیم. ضمن اینکه نحوه کاربرد علائم سجاوندی به‌کاررفته در چنین زیرنویس‌هایی را نیز بررسی خواهیم کرد.

دلیل انتخاب شبکه خبر گرایش زیاد بینندگان به آگاهی از اخبار روز است و از طرفی دیگر شبکه آی‌فیلم از معدود شبکه‌هایی است که نظر مردم را در قالب زیرنویس در برنامه‌های در

حال پخش نمایش می‌دهد. از سوی دیگر علاوه بر دست‌اندرکاران سیما که تولیدکننده زیرنویس‌های خبری هستند، ممکن است پیام‌های مردمی به همان شکل اولیه در قالب زیرنویس درآیند و بنابراین در این پژوهش علاوه بر بررسی اشکالات املایی و فنی زیرنویس‌های تولیدشده توسط مسئولان سیما، اشکالات ناشی از ارائه مستقیم پیام‌های مردمی را نیز نمایان می‌سازد.

از هر یک از دو شبکه ۱۰۰۰ جمله به شکلی که در جدول‌های ۱ و ۲ نمایش داده شده است، حروف‌چینی شد. سپس زیرنویس‌ها به صورت فایل متنی^۱ درمی‌آید تا امکان اعمال پردازش‌های رایانه‌ای میسر شود. به کمک نرم‌افزار ان‌اس‌پی^۲ واژه‌ها یا علائم و فراوانی آنها به صورت یک‌نگاشتی^۳ تا چندنگاشتی^۴ یعنی به شکل تک‌تک و توالی دو تا چند واژه یا علامت، ارائه می‌شود تا در نهایت پس از بررسی املایی و فنی آنها، نوع و میزان اشکالات املایی و فنی زیرنویس‌ها هم به صورت کلی و هم به صورت مجزا، مشخص شود. بدیهی است با مشخص بودن فراوانی این اشکالات، ساختار املایی غالب در زیرنویس‌ها مشخص خواهد شد.

جدول ۱. سه جمله نخست از پیکره شبکه خبر

ردیف	جمله
۱	یک کشته و ۱۰ زخمی بر اثر انفجار در کوئته پاکستان
۲	وقتی تو تلویزیون تصاویر کودکان کشته شده را می‌بینم، جیگرم آتش میگیره. من از هموطنانم می‌خوام فردا به ندای رهبرمون لیبیک بگیریم و به حمایت از فلسطین صدامون را به گوش جهانیان برسانیم
۳	وزیرکار فلسطین از تخریب ۲۳۳۰ واحد مسکونی در مناطق مختلف غزه بر اثر حملات رژیم صهیونیستی خبر داد

۱. txt

۲. N-gram Statistical Package (NSP)

۳. Unigram

۴. N-gram

جدول ۲. سه جمله نخست از پیکره آی فیلم

ردیف	جمله
۱	باسلام خسته نباشی خدمت ای فیلمی‌های عزیزممنون از فیلم‌هایی که برامون پخش می‌کنید خانواده‌ها را دور هم جمع می‌کنید من نخستین بار هست که پیام میدم امیدوارم پیامم پخش بشه. نادر شایان و محمدرضا وشعیب یادگاری از ایرانشهر
۲	زهرامردی از کاشان: هروقت که حوصله ام سر میره، پای برنامه های بسیارخوب وجذاب آی فیلم می شینم، نمی دونم چطورازتون تشکرکنم، فقط می تونم بگم: دستتون دردنکنه، خدقوت
۳	خورشید،محمد است وصادق ماه است خورشید همیشه با قمر همراه است یعنی که ولادت امام صادق در روز ولادت رسول الله است میلاد پیامبر مهر و رحمت و فرزند گرامیشان امام صادق (ع)به همه مسلمانان عزیزوخصوصا به مقام معظم رهبری تبریک عرض می کنم.کاظم تسبیه از شهرستان صومعه سرا

در اینجا بخش کوچکی از ابتدای هر دو متن ارائه شده است. متن نخست از شبکه خبر و متن دوم از شبکه آی فیلم است:

... یک کشته و ۱۰ زخمی بر اثر انفجار در کویته پاکستان وقتی تو تلویزیون تصاویر کودکان کشته شده را می‌بینم، جیگرم آتش میگیره. من از هموطنانم می‌خوام فردا به ندای رهبرمون لبیک بگیم و به حمایت از فلسطین صدامون را به گوش جهانیان برسانیم وزیرکار فلسطین از تخریب ۲۳۳۰ واحد مسکونی در مناطق مختلف غزه بر اثر حملات رژیم صهیونیستی خبر داد وزیران امور خارجه امریکا و روسیه در گفت وگویی تلفنی بر اهمیت آتش‌بس فوری در شرق اوکراین تاکید کردند وزیر امور خارجه روسیه : شبه جزیره کریمه ...

... باسلام خسته نباشی خدمت ای فیلمی‌های عزیزممنون از فیلم‌هایی که برامون پخش می‌کنید خانواده‌ها را دور هم جمع می‌کنید من اولین بار هست که پیام میدم امیدوارم پیامم پخش بشه. نادر شایان ومحمدرضا وشعیب یادگاری از ایرانشهر زهرامردی از کاشان:هروقت که حوصله‌ام سر میره،پای برنامه‌های بسیارخوب وجذاب آی فیلم می شینم،نمی دونم چطورازتون تشکرکنم،فقط می تونم بگم:دستتون دردنکنه،خدقوت خورشید، محمد است و صادق ماه است خورشید همیشه ...

همان‌گونه که در جدول ۳ مشاهده می‌شود، پیکره شبکه خبر شامل ۲۹۵۴ نوع^۱ و ۱۴۱۹۴ نماد^۲ است و این در حالی است که پیکره آی‌فیلم شامل ۵۷۴۹ نوع و ۱۹۳۲۶ نماد است. منظور از نوع واژه، عدد یا هر نوع علامت سجاوندی است و منظور از نماد وقوع مکرر یک نوع است. برای نمونه حرف اضافه «در» در شبکه خبر ۶۵۹ بار به کار رفته است که در چنین شرایطی می‌گوییم نوع «در» دارای ۶۵۹ نماد است.

جدول ۳. تعداد نوع و نماد در هر دو پیکره

مورد/شبکه	خبر	آی‌فیلم
نوع	۲۹۵۴	۵۷۴۹
نماد	۱۴۱۹۴	۱۹۳۲۶

با وجود برابر بودن تعداد جمله‌های منتخب هر یک از دو شبکه، مشاهده می‌شود که تعداد نوع و نماد در شبکه آی‌فیلم به مراتب بیشتر از شبکه خبر است که دلیل آن می‌تواند تنوع واژه‌های به کار رفته توسط کاربران بی‌شمار شبکه آی‌فیلم باشد. این در حالی است که در شبکه خبر به دلیل تخصصی بودن امر خبر و همچنین تعداد مشخص نویسندگان آن از دایره واژگانی محدودتری استفاده می‌شود.

در جدول‌های ۴ و ۵ به ترتیب واژه‌ها و عبارتهای استخراج‌شده از دو پیکره خبر و آی‌فیلم به همراه فراوانی هر یک نمایش داده شده است.

جدول ۴. نمونه‌هایی از واژه‌ها و عبارتهای پربسامد پیکره خبر و فراوانی آنها

فراوانی		نوع											
۶۵۹	در												
۶۲	در	را											
۲۷	خارج	امور	وزیر										
۲۲	خامنه‌ای	الله	آیت	حضرت									
۸	خامنه‌ای	الله	آیت	حضرت	است								
۶	های	سرزمین	مقصد	به	مختلف	کشورهای							
۶	اشغالی	های	سرزمین	مقصد	به	مختلف	کشورهای						
۶	متوقف	اشغالی	های	سرزمین	مقصد	به	مختلف	کشورهای					
۶	شده	متوقف	اشغالی	های	سرزمین	مقصد	به	مختلف	کشورهای				
۶	است	شده	متوقف	اشغالی	های	سرزمین	مقصد	به	مختلف	کشورهای			

1. type
2. token

باید به این نکته اشاره کرد که به دلیل اهمیت درج یا حذف فاصله و نیم‌فاصله در موارد مختلف، نرم‌افزار به‌گونه‌ای طراحی شد که بتواند تمامی این موارد را استخراج کند. برای نمونه، همان‌گونه که در جدول ۴ مشاهده می‌شود، واژه «سرزمین‌های» به دلیل عدم درج نیم‌فاصله به جای یک واژه یا به عبارت فنی‌تر یک‌نگاشتی، دونگاشتی یا دو واژه پیاپی یعنی «سرزمین + های» در نظر گرفته شده است.

عکس این نمونه در جدول ۵ مشاهده می‌شود؛ دونگاشتی «در چشم» به دلیل درج نکردن فاصله به صورت تک‌نگاشتی «درچشم» درج شده است.

جدول ۵. نمونه‌هایی از واژه‌ها و عبارتهای پربسامد پیکره‌آی فیلم و فراوانی آنها

فراوانی	نوع									
۵۵۱	از									
۱۶۱	کنید	پخش								
۴۹	کنید	پخش	را							
۱۳	نباشید	خسته	و	سلام						
۶	خسته	و	سلام	عرض	با					
۵	ازنهاوند	دارایی	حسین	کنید	پخش	دوباره				
۵	حسین	کنید	پخش	دوباره	را	دوباره	درچشم			
۵	دارایی	حسین	کنید	پخش	دوباره	را	باد	درچشم		
۵	ازنهاوند	دارایی	حسین	کنید	پخش	دوباره	را	باد	درچشم	
۴	ازنهاوند	دارایی	حسین	کنید	پخش	دوباره	را	باد	درچشم	فیلم

اهمیت استخراج چندنگاشتی‌ها این است که در برخی موارد مشخص نیست که واژه مدنظر چه بوده است و در نتیجه با مراجعه به توالی‌های چندواژه‌ای منظور نویسنده مشخص خواهد شد. پس از استخراج داده‌ها، امکان تحلیل خطاهای نوشتاری در دو شبکه فراهم می‌شود. در بخش بعدی این فصل به بررسی خطاها در هر یک از دو شبکه خواهیم پرداخت و در نهایت خطاهای دو شبکه را مقایسه خواهیم کرد.

۵. یافته‌های پژوهش

تحلیل داده‌ها در سه مرحله صورت گرفته است. ابتدا خطاهای هر پیکره به‌صورت مجزا بررسی می‌شود و سپس خطاهای دو پیکره با یکدیگر مقایسه می‌شود.

۵-۱. تحلیل خطاهای پیکره خبر

در این بخش ابتدا به تحلیل خطاهای ذکر شده در پیکره خبر خواهیم پرداخت و سپس فراوانی خطاها در این پیکره را مقایسه خواهیم کرد.

۵-۱-۱. اشکال املائی/سبکی در پیکره خبر

در کل ۱۷ نوع اشکال املائی/سبکی در پیکره خبر مشاهده شده است که با احتساب فراوانی ۱۱ برای صورت «تأکید»، شاهد ۲۷ نماد هستیم. با توجه به حجم داده‌های بررسی شده، فراوانی این نوع خطاها بسیار کم است که دلیل آن تخصص نویسندگان این متون است؛ برای نمونه به جای «پدر» از «پدره» یا به جای «تتانیاهو» از «تتانیاهور» استفاده شده است. از ۱۷ ایراد مشاهده شده، تنها سه نمونه خطای املائی است و در سایر موارد یا سبک نوشتاری معیار خبری رعایت نشده و از نوشتار غیررسمی استفاده شده است یا از علائم سجاوندی به‌درستی استفاده نشده است. در سایر موارد نیز علائم همزه و تنوین لحاظ نشده است؛ مانند «تاکید» به جای «تأکید».

۵-۱-۲. خطای درج فاصله در پیکره خبر

تنها سه نوع خطای درج فاصله در پیکره خبر مشاهده می‌شود که آن‌هم به دلیل درج فاصله پیش از علائم دونقطه، ویرگول و نقطه‌ویرگول است. کاربرد زیاد این خطا فراوانی کل این نوع خطا را که صرفاً در رعایت نکردن فاصله علائم سجاوندی به چشم می‌خورد را به ۲۷۲ مورد رسانده است که از حیث فراوانی نمادها تأمل‌برانگیز است و با توجه به وجود نمونه‌های بدون فاصله در پیکره خبر، نشان می‌دهد که هنگام درج علائم سجاوندی در زیرنویس‌های شبکه خبر، درج فاصله، شکل معیار تلقی می‌شود.

۵-۱-۳. خطای حذف فاصله در پیکره خبر

در این نوع خطا میان دو واژه مستقل هیچ نوع فاصله‌ای درج نشده است؛ برای نمونه به جای «وزیر کشور»، «وزیرکشور» درج شده است و به‌نوعی به لحاظ زبان‌شناختی، هر یک از این موارد یک واژه مرکب محسوب می‌شود. به‌غیراز نمونه «از جمله» که از نظر نگارنده حاضر می‌تواند یک واژه مرکب محسوب شود و به‌طور موردی قابل بحث است، در سایر موارد باید میان دو

واژه مستقل فاصله درج شود. با توجه به معیار بودن زبان شبکه خبر، فراوانی نسبتاً زیاد چنین خطایی یعنی ۳۳ نوع و ۳۷ نماد، قابل توجه است.

۴-۱-۵. خطای حذف نیم‌فاصله در پیکره خبر

در میان همه خطاهای بررسی شده در پیکره خبر، این نوع خطا از بیشترین فراوانی برخوردار است. برای نمونه در واژه «بین‌المللی» درج نکردن نیم‌فاصله باعث شده است که «المللی» یک واژه جدا تلقی شود. ۵۴۴ نماد از ۵۶ نوع این خطا نشان می‌دهد که اساساً به کاربرد نیم‌فاصله یا فاصله مجازی در واژه‌های مرکب و همچنین پس از برخی وندهای تصریفی مانند «می-» با ۹۹ نماد یا پیش از آنها مانند «های-» با ۱۹۰ نماد یا «ها-» با ۹۰ نماد در زیرنویس‌های شبکه خبر که قاعداً باید در آن از الگوی نوشتاری معیار تبعیت شود، توجهی نمی‌شود.

۵-۱-۵. خطای سرهم‌نویسی در پیکره خبر

تنها دو مورد خطای سرهم‌نویسی در پیکره خبر وجود دارد، آن‌هم «میگیره» به جای «می‌گیرد» و «صهیونیست‌های» به جای «صهیونیست‌های» که این خود مؤید رعایت اصول نگارش مصوب فرهنگستان است.

۶-۱-۵. خطای جدانویسی در پیکره خبر

در پیکره خبر حتی یک مورد هم خطای جدانویسی مشاهده نشد.

۷-۱-۵. مقایسه خطاهای پیکره خبر

پس از بررسی هر یک از خطاها در پیکره خبر به مقایسه فراوانی نوع و نماد خطاها در این پیکره خواهیم پرداخت. در جدول ۶ فراوانی نوع و نماد انواع خطا در پیکره خبر به تصویر کشیده شده است.

جدول ۶. فراوانی نوع و نماد انواع خطا در پیکره خبر

تعداد	واژه‌ها	کل اشکال املائی/سبکی	درج فاصله	حذف فاصله	حذف نیم‌فاصله	سرهم نویسی	جدانویسی	کل اشکالات
نوع	۲۹۵۴	۱۷	۳	۳۳	۵۹	۲	۰	۱۱۴
نماد	۱۴۱۹۴	۲۷	۲۷۲	۳۷	۵۴۴	۲	۰	۸۸۲

برای بررسی آماری داده‌ها، داده‌ها با فراوانی و درصد گزارش شدند. برای مقایسه اشتباهات رخ داده، از آزمون کای دو پیرسون^۱ استفاده شد. تحلیل داده‌ها با نرم‌افزار اسپ‌اس‌اس ۲۳ در سطح معنی‌داری ۰/۰۵ انجام شد. در جدول‌های ۷ و ۸ به ترتیب درصد فراوانی و نتیجه آزمون برحسب نوع و نماد خطاها نمایش داده شده است.

جدول ۷. تحلیل آماری انواع خطا در پیکره خبر (برحسب نوع)

نوع خطا	اشکال املائی/سبکی	درج فاصله	حذف فاصله	حذف نیم‌فاصله	سرهم‌نویسی	جدانویسی	نتایج آزمون کای دو
درصد	۱۴،۵٪	۲،۹٪	۳۹،۰٪	۵۲،۳٪	۱،۴٪	۰،۰٪	Pearson Chi-Square (5) = ۰۰۱.۰، P<۱۳.۱۰۵

بر اساس نتایج آزمون کای دو، برای این شبکه، برحسب نوع، میان درصد انواع مختلف خطا، اختلاف معنی‌داری وجود دارد. نتایج نشان می‌دهد که بیشترین درصد اشتباه برای خطای حذف نیم‌فاصله است و خطاهای حذف فاصله و اشکالات املائی/سبکی در رده بعدی قرار دارند. همچنین درج فاصله و سرهم‌نویسی در مراتب بعدی قرار داشتند و درنهایت هیچ نوع خطای جدانویسی وجود ندارد.

جدول ۸. تحلیل آماری انواع خطا در پیکره خبر (برحسب نماد)

نوع خطا	اشکال املائی/سبکی	درج فاصله	حذف فاصله	حذف نیم‌فاصله	سرهم‌نویسی	جدانویسی	نتایج آزمون کای دو
درصد	۳،۰٪	۳۰،۸٪	۴،۱٪	۶۱،۷٪	۰،۲٪	۰،۰٪	Pearson Chi-Square (5) = ۰۰۱.۰، P<۳۰۲.۱۱۹۹

بر اساس نتایج آزمون کای دو، برای این شبکه، برحسب نماد، میان درصد انواع مختلف خطا اختلاف معنی‌داری وجود دارد. همچنین نتایج نشان می‌دهد که بیشترین درصد اشتباه برای خطای حذف نیم‌فاصله بوده و خطاهایی چون درج فاصله، اشکال املائی/سبکی و سرهم‌نویسی در مراتب بعدی قرار گرفتند و درنهایت هیچ مورد اشتباه در جدانویسی به چشم نخورده است.

۱. Pearson Chi-Square

۲. SPSS 23

مشاهده می‌شود که وقتی نمادها بررسی می‌شود، فراوانی خطای درج فاصله بسیار زیاد می‌شود و این بیانگر این مسئله است که فراوانی این خطا در تعدادی مورد خاص دائماً تکرار شده است.

۲-۵. تحلیل خطاهای پیکره‌آی فیلم

در این بخش ابتدا به تحلیل خطاها در پیکره‌آی فیلم خواهیم پرداخت و سپس فراوانی خطاها در این پیکره را مقایسه خواهیم کرد.

۱-۲-۵. اشکال املائی/سبکی در پیکره‌آی فیلم

در پیکره‌آی فیلم ۳۹۱ نوع و ۸۴۲ نماد خطای املائی/سبکی وجود دارد. برخلاف پیکره‌آی خبر، تعداد زیادی از این نوع خطا به اشکالات املائی مربوط می‌شود. نمونه‌های نظیر «طرفطار»، «مچکرم»، «خواهشن»، «نزارین» و «سپاسگذارم» از جمله نمونه‌های قابل توجه است؛ هرچند اشکالاتی مانند استفاده نکردن از تنوین، سرکش و همزه و همچنین استفاده نادرست از علائم سجاوندی یا ویرگول لاتین به جای فارسی، استفاده از اعداد لاتین در متن فارسی نیز در میان این اشکالات به چشم می‌خورد.

در میان خطاهای مشاهده شده، کاربرد پرسامد صورت‌هایی مانند «مچکرم»، «خواهشن»، «نزارین» و «سپاسگذارم» از جمله مواردی است که در شبکه‌های اجتماعی به کرات مشاهده می‌شود و به نظر می‌رسد که علیرغم تأثیر بسزای اینترنت و شبکه‌های اجتماعی در بهبود زندگی اجتماعی، بی‌توجهی به نحوه نگارش، جامعه را با خطر تغییر نظام نوشتاری فارسی و ایجاد بی‌ثباتی در آن مواجه خواهد کرد.

۲-۲-۵. خطای درج فاصله در پیکره‌آی فیلم

در پیکره‌آی فیلم ۲۳ نوع و ۵۵ نماد خطای درج فاصله وجود دارد. نگارش «خواهش» به جای «خواهش» و «تقاضای» به جای «تقاضای» از جمله این نوع خطاست. برخلاف شبکه‌آی خبر، که این نوع خطا محدود به سه مورد می‌شود، تنوع این خطا در شبکه‌آی فیلم بسیار زیاد است. در این میان بازهم درج فاصله پس از علامت «:»، بیشترین فراوانی را دارد.

۳-۲-۵. خطای حذف فاصله در پیکره‌آی فیلم

۳۱۴۰ نماد از ۲۱۵۸ نوع خطای حذف فاصله بیانگر فراوانی بسیار زیاد این نوع خطاست. مشاهده این تعداد زیاد حکایت از ناآگاهی بسیاری از نگارندگان زیرنویس‌های شبکه‌آی فیلم از درج فاصله میان واژه‌ها دارد.

این در حالی است که ملاک شمارش خطا در پژوهش حاضر صرفاً نوع مورد بررسی بوده و اگر در یک نوع، چند بار این خطا تکرار شده باشد، تعداد این تکرار محاسبه نشده است. برای نمونه در عبارت «سرباز، مسافر زمان ۱ و ۲ رویدارین. مرسی» قاعدتاً، میان واژه‌ها و علائم سجاوندی باید ۸ بار فاصله درج می‌شده است ولی تنها یک مورد خطا محسوب شده است. دلیل این نوع محاسبه این است که اگر نگارنده زیرنویس نسبت به درج فاصله آگاه می‌بود حداقل در یک مورد، فاصله را درج می‌کرد؛ ولی اساساً به فاصله‌گذاری اعتقادی نداشته است و در نتیجه می‌توان گفت که یک‌بار خطا کرده است. بدیهی است در صورت محاسبه هر یک از این موارد حذف فاصله، تعداد نمادها بسیار بیشتر از تعداد فعلی می‌بود.

۴-۲-۵. خطای حذف نیم‌فاصله در پیکره‌آی فیلم

۱۰۳ نوع و ۶۷۳ نماد از این نوع خطا مشاهده شده است. ویژگی این نوع خطا این است که نگارنده نسبت به جدانویسی برخی عناصر واژگانی همچون وندهای «-ها»، یا «می-» آگاهی دارد ولی از نیم‌فاصله یا به دلیل ناآگاهی یا به دلیل مشکل درج آن استفاده نمی‌کند.

۵-۲-۵. خطای سرهم‌نویسی در پیکره‌آی فیلم

با بررسی این پیکره مشخص می‌شود که ۲۹۸ نوع خطای سرهم‌نویسی در این پیکره مشاهده می‌شود که با احتساب تکرار آنها تعداد نمادهای این نوع خطا به ۵۸۷ مورد می‌رسد. بیشترین میزان خطا در سرهم‌نویسی صورت‌های تصریفی «می-» و «-ها» مشاهده می‌شود، هرچند در مواردی مانند «هاشمپور» به جای «هاشم‌پور» یا «گیلانغرب» به جای «گیلان غرب» نیز این نوع خطا مشاهده می‌شود.

۶-۲-۵. خطای جدانویسی در پیکره‌آی فیلم

همان‌گونه که انتظار می‌رفت در پیکره‌آی فیلم نیز حتی یک مورد خطای جدانویسی مشاهده نشد.

۷-۲-۵. مقایسه خطاهای پیکره‌آی فیلم

پس از بررسی هر یک از انواع خطا در پیکره‌آی فیلم می‌توان به مقایسه فراوانی آنها پرداخت. در جدول ۹ فراوانی نوع و نماد هر یک از خطاها مشخص شده است.

جدول ۹. فراوانی نوع و نماد انواع خطا در پیکره‌آی فیلم

تعداد	کل واژه‌ها	اشکال املائی/سبکی	درج فاصله	حذف فاصله	حذف نیم‌فاصله	سرهم‌نویسی	جدانویسی	کل اشکالات
نوع	۵۷۴۹	۳۹۱	۲۳	۲۱۵۸	۱۰۴	۲۹۸	۰	۲۹۷۴
نماد	۱۹۳۳۶	۸۴۲	۵۵	۳۱۴۰	۶۷۳	۵۸۷	۰	۵۲۹۷

مشاهده می‌شود که فراوانی نوع خطا نزدیک به نیمی از تعداد کل واژه‌هاست و فراوانی نماد خطاها نزدیک به یک‌چهارم کل نمادهای پیکره است و این میزان زیاد خطا تأمل‌برانگیز است. به‌منظور انجام مقایسه دقیق‌تر انواع خطا از آزمون کای‌دو استفاده شده است. در جدول ۱۰ تحلیل آماری انواع خطا در پیکره‌آی فیلم برحسب نوع و در جدول ۱۱ برحسب نماد، قابل مشاهده است.

جدول ۱۰. تحلیل آماری انواع خطا در پیکره‌آی فیلم (برحسب نوع)

نوع خطا	اشکال املائی/ سبکی	درج فاصله	حذف فاصله	حذف نیم‌فاصله	سرهم‌نویسی	جدانویسی	نتایج آزمون کای دو
درصد	۱۳،۲٪	۰،۸ ٪	۷۲،۶٪	۳،۵٪	۱۰،۰٪	۰،۰٪	Pearson Chi-Square (5) = ۷۸،۵۳۹۸ ۰،۰۰۱P<

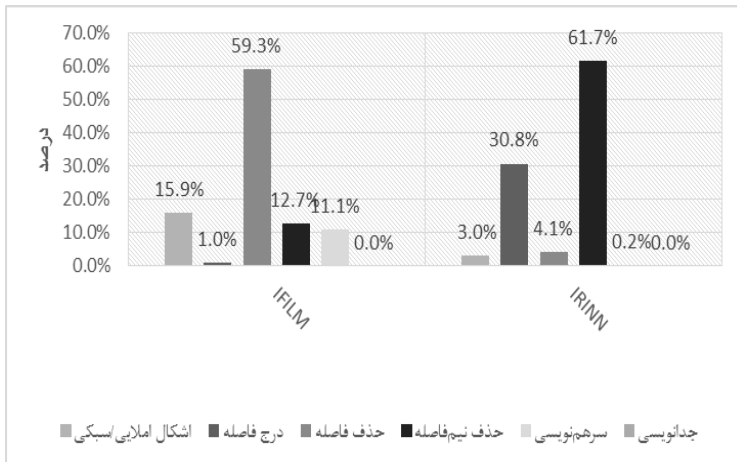
بر اساس نتایج آزمون کای دو، در پیکره‌آی فیلم، میان درصد انواع مختلف خطا اختلاف معنی‌داری وجود دارد. نتایج نشان می‌دهد که بیشترین درصد خطا در حذف فاصله وجود دارد و اشکال املائی/سبکی و سرهم‌نویسی در رده بعدی و همچنین حذف نیم‌فاصله و درج فاصله در مراتب بعدی قرار دارند و درنهایت هیچ مورد اشتباه در جدانویسی به چشم نمی‌خورد.

جدول ۱۱. تحلیل آماری انواع خطا در پیکره‌آی فیلم (برحسب نماد)

نوع خطا	اشکال املائی / سبکی	درج فاصله	حذف فاصله	حذف نیم‌فاصله	سرهم‌نویسی	جدانویسی	نتایج آزمون کای دو
درصد	۱۵,۹٪	۱,۰٪	۵۹,۳٪	۱۲,۷٪	۱۱,۱٪	۰,۰٪	Pearson Chi-Square (5) = ۰,۰۰۱, P<۶۹.۵۹۰۰

اما هنگام تحلیل آماری تعداد نماد خطا در شبکه‌آی فیلم، نیز مشخص می‌شود که میان درصد انواع مختلف خطا اختلاف معنی‌داری وجود دارد. پژوهش نشان می‌دهد که بیشترین درصد اشتباه برای مورد حذف فاصله است و اشکال املائی/سبکی، حذف نیم‌فاصله و سرهم‌نویسی در رده‌های بعدی قرار دارند؛ خطای درج فاصله در مرتبه بعدی قرار دارد و درنهایت هیچ مورد اشتباه در جدانویسی وجود ندارد.

تفاوت میان فراوانی نوع و نماد خطا در پیکره‌آی فیلم در رتبه‌بندی خطا مشاهده می‌شود. درحالی‌که خطای سرهم‌نویسی در میان انواع خطا برحسب نوع در جایگاه سوم قرار دارد، این خطا برحسب نماد در جایگاه چهارم قرار گرفته است و این در حالی است که خطای حذف نیم‌فاصله برحسب نوع در رتبه چهارم و برحسب نماد در رتبه سوم است.

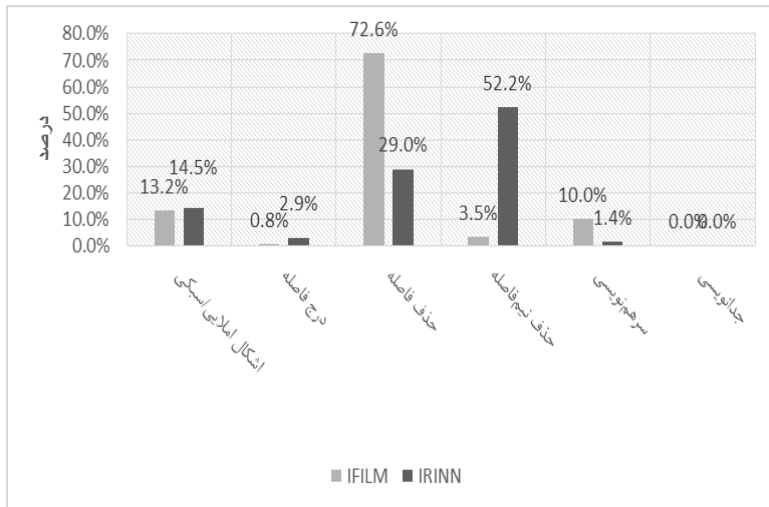


نمودار ۱. مقایسه درصد خطاها برحسب نماد با تفکیک دو پیکره‌آی فیلم

دلیل این امر کاربرد مکرر خطای حذف نیم‌فاصله برای برخی موارد خاص است که باعث افزایش فراوانی نمادهای این نوع شده است. این در حالی است که تفاوت زیادی میان درصد نوع و نماد در خطای سرهم‌نویسی مشاهده نمی‌شود. نمودار ۱ درصد خطاها برحسب نماد را در هر دو پیکره مورد بررسی به تصویر کشیده است. ملاحظه می‌شود که درصد خطای حذف فاصله در پیکره آی‌فیلم و خطای حذف نیم‌فاصله در پیکره خبر بسیار بیشتر از سایر خطاهاست.

۳-۵. مقایسه خطاهای دو پیکره خبر و آی‌فیلم

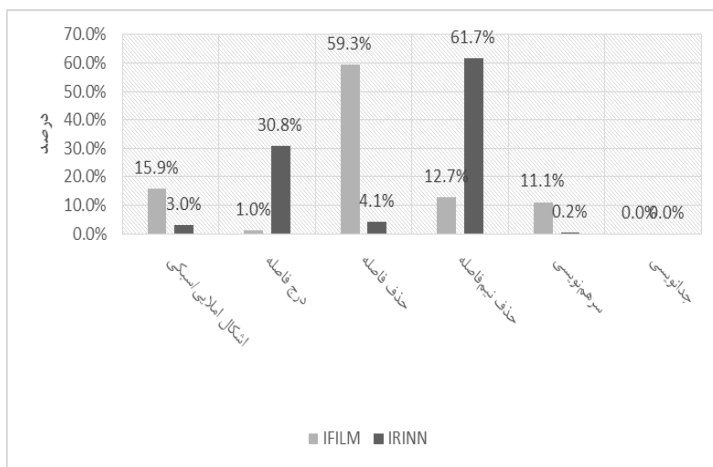
برای مقایسه دو شبکه، داده‌ها با فراوانی و درصد خلاصه، گزارش شدند. برای مقایسه دو شبکه در اشتباهات رخ داده در حالت‌های خطاهای مختلف از آزمون کای دو استفاده شد. برای تعدیل تعداد موارد بررسی شده در دو شبکه که متفاوت بودند، فراوانی کل تعریف شده در هر شبکه بر حسب میانگین مجموع تعداد موارد بررسی شده در دو شبکه تعدیل (نرمالایز) شد و به عنوان وزن تعریف شد. تحلیل داده‌ها با نرم‌افزار اسپ‌اس‌اس ۲۳ در سطح معنی‌داری ۰/۰۵ انجام شد. در نهایت مقایسه در دو سطح صورت گرفت؛ نخست در سطح نوع و سپس در سطح نماد. بر اساس نتایج آزمون کای دو، میان دو شبکه اختلاف معنی‌داری وجود دارد. همان‌گونه که انتظار می‌رفت، درصد کل اشتباهات در پیکره آی‌فیلم به‌طور معنی‌داری بالاتر است. در نمودار ۲ به مقایسه انواع خطا برحسب نوع در دو شبکه پرداخته‌ایم.



نمودار ۲. مقایسه درصد اشتباهات نوعی، بین دو شبکه خبر و آی‌فیلم

آن‌گونه که نمودار نشان می‌دهد اختلاف معنی‌داری در دو مورد اشکالات املائی و درج فاصله میان دو پیکره وجود ندارد، اما میزان خطای حذف فاصله و سرهم‌نویسی در پیکره آی‌فیلم به‌طور معنی‌داری بیشتر از پیکره خبر است و تنها خطای حذف نیم‌فاصله است که در پیکره خبر بیش از پیکره آی‌فیلم مشاهده می‌شود. از سوی دیگر به دلیل اینکه خطای جدانویسی در هیچ‌یک از دو پیکره مشاهده نمی‌شود، هیچ تفاوتی میان دو شبکه وجود ندارد.

بر اساس نتایج آزمون کای دو، میان دو شبکه اختلاف معنی‌داری وجود دارد. همان‌گونه که پیش‌بینی می‌شد درصد اشتباهات در شبکه آی‌فیلم به‌طور معنی‌داری بالاتر است. در نمودار ۳ تفاوت خطاها برحسب نماد به‌صورت کلی و تفکیک‌شده به تصویر کشیده شده است.



نمودار ۳. مقایسه درصد خطاها برحسب نماد، میان دو پیکره خبر و آی‌فیلم

نمودار نشان می‌دهد که به‌غیراز میزان خطای جدانویسی که در هر دو پیکره صفر است و تفاوتی وجود ندارد، در سایر موارد میان دو پیکره اختلاف معنی‌داری وجود دارد. درحالی‌که در پیکره خبر درصد نمادهای خطای درج فاصله و حذف نیم‌فاصله بسیار بیشتر از پیکره آی‌فیلم است. در مقابل درصد فراوانی خطاهای املائی/اسبکی، حذف فاصله و سرهم‌نویسی به شکل معنی‌داری در پیکره آی‌فیلم بیشتر است.

۶. بحث و نتیجه‌گیری

در این بخش به تبیین برخی مشکلات نوشتاری و ارائه پیشنهادهایی برای کاهش خطاهای نوشتاری خواهیم پرداخت.

- بررسی درصد انواع خطا در دو پیکره نشان می‌دهد که هیچ نوع خطای جدانویسی وجود ندارد. شاید بتوان دلیل نبود این خطا را گرایش کاربران خط فارسی به سرهم‌نویسی صورت‌های زبانی، و به‌ویژه صورت‌های مشتق، دانست.

- بررسی انواع مختلف خطا در شبکه‌خبر، در هر دو سطح نوع و نماد، نشان می‌دهد که بیشترین درصد اشتباه برای خطای حذف نیم‌فاصله است. مسلماً نویسندگان زیرنویس‌های خبری درخصوص جدا کردن عناصر زبانی و به‌ویژه عناصر تصریفی زبان آگاهی دارند، اما شاید بتوان جدید بودن نسبی فناوری فاصله مجازی یا همان نیم‌فاصله را دلیل اصلی فراوانی زیاد در خطای حذف نیم‌فاصله در نظر گرفت. از سوی دیگر هنگام بحث درباره فاصله‌گذاری در «دستور خط فارسی» (۱۳۸۹: ۱۰)، به‌اختصار، به موضوع استفاده از فاصله مجازی یا نیم‌فاصله، که امروزه نقشی مهم در نگارش رایانه‌ای خط فارسی دارد، اشاره شده است و این خود می‌تواند دلیلی دیگر بر زیاد بودن میزان خطای حذف نیم‌فاصله در شبکه رسمی خبر باشد. باوجود این انتظار می‌رود در شبکه رسمی همچون شبکه خبر به این مهم توجه شود.

- مقایسه نمادهای خطا در دو شبکه نشان می‌دهد که درصد فراوانی خطاهای املائی/سبکی، حذف فاصله و سرهم‌نویسی به شکل معنی‌داری در پیکره آی‌فیلم بیشتر است. دلیل این امر این است که برخلاف مخاطبان عمومی شبکه آی‌فیلم، نویسندگان زیرنویس‌های شبکه خبر از الگوی نگارشی معیار و یکسان استفاده می‌کنند و در نتیجه میزان فراوانی خطاها به‌ویژه خطاهای املائی/سبکی در شبکه خبر بسیار کمتر است.

- همان‌گونه که در بخش سوم پژوهش حاضر اشاره شد شیوه‌نامه «دستور خط فارسی» دارای اشکالاتی است و نیاز به تصحیح دارد. بازبینی و تصحیح دستور خط فارسی مصوب فرهنگستان به‌گونه‌ای که برای همه افشار باسواد جامعه قابل‌درک باشد، بسیار مهم است. نوع نگارش متن فرهنگستان به شکلی است که برخی موارد حتی برای زبان‌شناسان و پژوهشگران حوزه زبان و ادب فارسی نیز مبهم است.

- دستاوردهای پژوهش حاضر و به‌ویژه زیاد بودن خطاهای املائی در زیرنویس‌های مردمی از یک‌سو و نگاهی اجمالی به نحوه استفاده از خط فارسی در شبکه‌های اجتماعی رایانه‌ای،

هشدار جدی به مسئولان و متولیان حفظ و ترویج خط فارسی است که نادیده گرفتن آن می‌تواند باعث پراکندگی و بی‌ثباتی در نظام نگارش فارسی شده و درنهایت آینده زبان فارسی را با مشکل روبه‌رو سازد.

- با توجه به اینکه صداوسیما جمهوری اسلامی ایران مهم‌ترین نقش را در ارائه الگوی صحیح نگارش فارسی دارد، لازم است ضمن پیروی از اصول دستور خط فارسی فرهنگستان در تولید زیرنویس‌های خبری و تبلیغاتی، به تصحیح زیرنویس‌هایی که توسط بینندگان نوشته شده است، بپردازد. همان‌گونه که در پژوهش حاضر نشان داده شده است، حتی در زیرنویس‌های شبکه خبر که عالی‌ترین جلوه زبان فارسی معیار نوشتاری است، نیز شاهد خطاهای نوشتاری هستیم.

- با توجه به نقش پررنگ صداوسیما در حفظ و گسترش خط فارسی معیار و با توجه به ناآگاهی و گاهی بی‌توجهی کاربران خط فارسی نسبت به دستور خط فارسی، به‌ویژه در شبکه‌های اجتماعی، برگزاری نشست‌های تخصصی متعدد با حضور زبان‌شناسان، ادیبان و متخصصان حوزه‌های مختلف علمی، درباره خط فارسی معیار و بررسی مشکلات آن، امری ضروری محسوب می‌شود. آموزش دستور خط فارسی در مراکز آموزشی و پژوهشی، حتی از مقاطع ابتدایی، از طریق برگزاری کلاس‌های تخصصی، بازبینی کتاب‌های درسی در مدارس و دانشگاه‌ها و تصحیح آنها بر اساس دستور خط فارسی فرهنگستان، نظارت بر شیوه نگارش رسانه‌های نوشتاری کشور و تأکید بر پیروی آنها از اصول خط فارسی از جمله مواردی است که باید بیش‌ازپیش مورد توجه قرار گیرد.

منابع و مأخذ

- اسلامی، محرم (۱۳۸۱). «دشواری‌های پردازش رایانه‌ای خط فارسی»، *مجله نشر دانش*، شماره ۳، صص ۲۸-۳۲.
- اعتصام‌زاده، ابوالقاسم (۱۳۰۴). «لزوم اصلاح خط فارسی»، *مجله آینده*، شماره ۳.
- اوجی، سیدعلی (۱۳۷۴). *پژوهشی در واژه‌های قرصی اروپایی در خبر رسانه‌های گروهی ایران*، پایان‌نامه کارشناسی ارشد علوم ارتباطات اجتماعی، تهران: دانشگاه علامه طباطبائی.
- بهمن‌یار، احمد (۱۳۲۲). «املائی فارسی پیشنهاد به مقام فرهنگستان»، *نامه فرهنگستان قدیم*، شماره ۴، صص ۴۲-۶۶.
- بهمنی، مهدی (۱۳۸۴). *تأثیر تلویزیون بر رشد گفتار در کودکان و نوجوانان*، پایان‌نامه کارشناسی ارشد، مشهد: دانشگاه فردوسی.
- جبارزاده، فاطمه و محمد شهبها (۱۳۸۹). «چالش‌های ترجمه‌ی فیلم»، *کتاب ماه ادبیات*، شماره ۱۵۰، ۸۷-۹۴.
- حق‌شناس، علی‌محمد (۱۳۷۴). «همزه در زبان و خط فارسی»، *زبان و ادبیات فارسی دانشکده علوم انسانی دانشگاه تربیت معلم*، شماره ۱۱، صص ۳۶-۲۳.
- خیام‌پور، عبدالرسول (۱۳۳۸). «رسم خط فارسی»، *زبان و ادبیات فارسی دانشکده ادبیات و علوم انسانی تبریز*، شماره ۴۸، صص ۱۰۴-۸۹.
- داوری اردکانی، نگار (۱۳۸۵). «رسانه ملی متولی برنامه‌ریزی زبان»، *فصل‌نامه پژوهش و سنجش*، شماره ۴۷، صص ۴۴-۲۱.
- دمالی امیری، مهدی (۱۳۸۰). *تبلور ایدئولوژی و قدرت در کلام (تحلیل انتقادی اخبار رسانه‌ای)*، پایان‌نامه کارشناسی ارشد، تهران: پژوهشگاه علوم انسانی و مطالعات فرهنگی.
- دولت‌آبادی، عزیز (۱۳۳۸). «رسم خط فارسی»، *زبان و ادبیات دانشکده ادبیات و علوم انسانی تبریز*، شماره ۵۰، صص ۲۳۲-۲۲۹.
- ذاکری، محمدصالح (۱۳۷۱). *زبان و رسانه‌های گروهی: آشنایی با زبان رادیو*، پایان‌نامه کارشناسی ارشد، تهران: دانشگاه تهران.
- ذوالفقاری، حسن (۱۳۸۷). *راهنمای ویراستاری و درست‌نویسی*، تهران: نشر علم.
- راسخ‌نورد، اسماعیل (۱۳۸۹). *چرا املا می‌نویسیم*، دریافت از وبگاه <http://rasekh-navard.blogfa.com/post-12.aspx>، ۱۳۸۹/۰۸/۰۷.
- سمائی، فرشید (۱۳۸۸). *بررسی مسائل زبان‌شناختی زیرنویس‌های شبکه خبر*، تهران: مرکز تحقیقات صداوسیما جمهوری اسلامی ایران.
- شازاد، لیلا (۱۳۷۵). *نقش جنسیت در شروع و پایان مکالمات حضوری سریال‌های ایرانی تلویزیون*، پایان‌نامه کارشناسی ارشد، تهران: دانشگاه آزاد اسلامی واحد تهران مرکز.
- ضیاء، اسماعیل (۱۳۸۱). «شیوه خط فارسی: دیدگاهی دیگر»، *مجله چیستا*، شماره ۱۹۰، صص ۷۹۴-۷۹۷.
- عاصی، مصطفی (۱۳۸۲). «از پیکره زبانی تا زبان‌شناسی پیکره‌ای»، *مجموعه مقالات پنجمین کنفرانس زبان‌شناسی*، تهران: پژوهشگاه علوم انسانی و مطالعات فرهنگی.

آسیب‌شناسی نگارشی زیرنویس‌ها در شبکه‌های... ❖ ۵۹

عرفانی، فرهاد (۱۳۸۷). **اهمیت حفظ و گسترش زبان فارسی**، دریافت از وبگاه <http://persianlanguage.ir> ۱۳/۰۸/۱۳۸۶.

غلامحسین زاده، غلامحسین (۱۳۷۹). «نقد دستور خط فارسی متن پیشنهادی فرهنگستان زبان و ادب فارسی»، **مجله علوم انسانی دانشگاه الزهراء (س)**، شماره ۳۴، صص ۱۵۳-۱۲۵.

فرهمندزاد، علی (۱۳۹۰). **آموزش املا و راهکارهای آن**، دریافت از وبگاه <http://amouzsh3.persianblog.ir> ۲۱/۰۹/۱۳۹۱.

فرهنگستان زبان و ادب فارسی (۱۳۸۹). **دستور خط فارسی**، دریافت از وبگاه <http://www.persianacademy.ir> ۲۲/۰۹/۱۳۹۵.

کبیری، سعدالله (۱۳۸۶). «خط فارسی بهترین نیست»، **مجله حافظ**، شماره ۴۵، صص ۴۹-۴۲. کشاورز، پریسا و عظیمه ستاری (۱۳۹۰). **کاربرد خط فارسی در تلویزیون**، تهران: مرکز تحقیقات صداوسیما جمهوری اسلامی ایران.

مدرس خیابانی، شهرام (۱۳۸۹). **آسیب‌شناسی زبان در گزارش‌های فوتبال سیما**، تهران: مرکز تحقیقات صداوسیما جمهوری اسلامی ایران.

مدرس خیابانی، شهرام (۱۳۹۱). **آسیب‌شناسی زبان در مجموعه‌های طنز سیما**، تهران: مرکز زبان صداوسیما جمهوری اسلامی ایران.

مهریار، لیل (۱۳۸۲). **بررسی میزان کاربرد واژگان مصوب فرهنگستان زبان در برنامه‌های سیما جمهوری اسلامی ایران**، پایان‌نامه کارشناسی ارشد، تهران: دانشگاه آزاد اسلامی واحد تهران مرکز.

میرزایی، علی (۱۳۸۷). **بررسی جنبه‌های نامأنوس زبان فارسی در فیلم‌های دوبله‌شده سیما جمهوری اسلامی ایران**، پایان‌نامه کارشناسی ارشد، اصفهان: دانشگاه اصفهان.

نصیری‌پور، ملیحه (۱۳۹۱). **مقایسه شیوه املا مصوب فرهنگستان زبان و ادب فارسی و شیوه املا غالب در متون فارسی معاصر**، پایان‌نامه کارشناسی ارشد، تهران: دانشگاه پیام نور.

نفیسی، سعید (۱۳۴۲). «تاریخ خط در ایران»، **مجله کاوش**، شماره ۱، صص ۲۶-۱۸. هندی، سعید (۱۳۸۱). «دستور خط فارسی: شیوه‌ای در نگارش کلمه‌های مرکب»، **رشد آموزش زبان و ادب فارسی**، شماره ۶۳، صص ۳۱-۲۷.

Biber, D., Johansson, S., Leech, G., Conrad, S., and Finegan, E. (1999) **Longman Grammar of Spoken and Written English**. London: Pearson.

Baker, P., Hardie, A. and T. McEnery (2006). **A Glossary of Corpus Linguistics**. Scotland: Edinburgh University Press.

Cordella, M (2006). "Discourse Analysis and The Subtitles of Documentaries: The case of the Children of Russia", **Odisea**, 7, pp 77-88.

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

Linde, Z. D and Kay, N (1999). **The Semiotics of Subtitling**, Manchester: St. Jerome Publishing.

McEnery, C. T. and A. Hardie (2012). **Corpus Linguistics: Method, Theory and Practice**, London: Cambridge University Press.

Stubbs, M. (2000) "Using Very Large Corpora Text Collections to Study Semantic Schemes",
in C. Heffer and H. Saunston (eds) **Words in Context: A Tribute to John Sinclair on
his Retirement, English Language Research Discourse Analysis Monograph 18.**
Birmingham: University of Birmingham.