

مطالعه‌ای در انتساب رسائل سعدی

آرش پوراکبر*

دانشجوی دکتری زبان و ادبیات فارسی، دانشگاه تهران

چکیده

رسائل سعدی به شش نوشتار مجزا اطلاق می‌شود که معمولاً در کلیات شیخ به‌چاپ می‌رسند و همگی به نثر نوشته شده‌اند. از زمان اولین تلاش‌ها برای چاپ نسخه مصحح علمی کلیات سعدی، پرسش از انتساب این رسائل موضوعی مهم بوده است. محققان مختلف نیز هریک به این موضوع پرداخته و به این سؤال پاسخ‌هایی داده‌اند. در این مقاله سعی شده است با استفاده از روش‌های سبک‌سنجی و انتساب تألیف، مبتنی بر تحلیلی کمی، و نیز با مقایسه این رسائل و گلستان، به پرسش درباره انتساب این رسائل پاسخ داده شود. در تحقیق پیش‌رو، با استفاده از دو نوع تحلیل مبتنی بر دو شاخص متفاوت، به‌طور دقیق به مسئله پژوهش پاسخ داده شده است. روش اول استفاده از منحنی‌های خصیصه‌نمای مندنهال است و روش دوم به‌کار گرفتن الگویی برای توصیف کمی شاخص تکرار واژگان در متن. نتیجه به‌دست‌آمده به‌طور کلی با حدس فروغی درباره انتساب این رسائل هماهنگ است. سه رساله («نصیحة الملوک»، «عقل و عشق» و «انکیانو») به قلم خود سعدی هستند؛ «تقریرات ثلاثه» و «در تقریر دیباچه» حتماً نوشته فرد دیگری هستند؛ درباره «مجالس» نیز باید گفت اگرچه به قلم سعدی نیست، احتمال اینکه سخنان شفاهی وی بر منبر باشند، زیاد است.

واژه‌های کلیدی: رسائل سعدی، سبک‌سنجی، انتساب تألیف، منحنی مندنهال، تکرار واژگان.

فصلنامه علمی - پژوهشی نقد ادبی، س. ۹، ش. ۳۵، پاییز ۱۳۹۵ (صص ۷۱-۹۹)

* نویسنده مسئول: arash.poorakbar@yahoo.com

تاریخ پذیرش: ۱۳۹۵/۸/۱۷

تاریخ دریافت: ۱۳۹۵/۴/۲۰

۱. مقدمه

حدود نیم قرن بعد از اولین تأثرات علوم انسانی از ریاضیات^۱، به دلیل اختراع رایانه، جهان معاصر به شدت متحول شد. این نوآوری باعث به وجود آمدن حوزه تازه‌ای در علوم بشری شد که به طور کلی ذیل عنوان علوم کامپیوتر شناخته می‌شود. این بار علوم انسانی چندان از جریان دنیای حاضر عقب نیفتاد و اگرچه کمی دیرتر، اما به سرعت خود را به این حوزه تازه پیوند زد و از دستاوردهای آن بهره‌برداری‌های زیادی کرد؛ تا آنجا که نه تنها این روزها کامپیوتر به مثابه ماشین و علوم کامپیوتر به منزله یک رویکرد و همچنین نحوه برخورد با این ماشین در علوم انسانی کاربرد دارند؛ بلکه این آمیختگی تا بدانجا پیش رفت که حوزه تازه‌ای با عنوان علوم انسانی دیجیتال^۲ پا به عرصه گذاشت. این حوزه به سرعت گسترش یافت و بسیاری از افراد و ایده‌ها را جذب خود کرد. فقط نگاهی به حجم مطالعات منتشرشده در این حوزه و همچنین وجود نهادهای بزرگ مطالعاتی برای اثبات این ادعا کافی است (Hockey, 2004).

همراهی این دو حوزه آغاز راهی بود که از آن ذیل عنوان‌هایی کلی همچون علوم انسانی رایانشی یا بعدتر، علوم انسانی دیجیتال می‌توان یاد کرد. شاید به عنوان یکی از متقدم‌ترین و درعین حال معروف‌ترین نمونه‌های این نوع فعالیت‌ها بتوان به روبرتو بوسا^۳ اشاره کرد. این کشیش یسوعی در سال ۱۹۴۹م فهرستی با حدود یک میلیون واژه از آثار توماس آکویناس تهیه کرد. البته، او این کار را به کمک توماس واتسون از آی.بی.ام. انجام داد. نسخه اینترنتی محصول کار او هم‌اکنون در دسترس است^۴ (همان‌جا). از نمونه‌های موفق دیگر می‌توان به مطالعات موستلر و والاس^۵ (۱۹۶۴) درباره مقالات فدرالیست و فعالیت‌های مورتون^۶ (۱۹۷۸) از جمله تحقیق درباره نامه سوم به قرنتیان^۷ منتسب به پولس اشاره کرد.

جدای از انواع بسیار متنوع فعالیت‌های انجام‌گرفته در حوزه علوم انسانی رایانشی، در اینجا تمرکز ما بر آن چیزی است که امروزه با عنوان کلی مطالعات انتساب تألیف^۸ شناخته می‌شود. در این نوع مطالعات، مسئله انتساب یک یا چند اثر به مؤلف / مؤلفان آن‌ها مطرح است.

در حالت کلی، سه گونه مسئله در مطالعات انتساب وجود دارد: اولین دسته از سؤالات به این شرح است که با فرض وجود نمونه‌ای از متون که حتماً توسط فردی از یک مجموعه از نویسندگان نوشته شده، مؤلف مورد نظر را تعیین کنید. این سؤال که به آن نسخه بسته مسئله می‌گویند، با نوع باز آن ارتباط دارد. دومین دسته از سؤالات چنین است: با فرض وجود مجموعه‌ای از متون که تصور می‌شود توسط فردی از یک مجموعه از مؤلف‌ها نوشته شده، مؤلف را تعیین کنید؛ البته، در صورتی که مؤلف در میان این گروه باشد؛ به تعبیر دیگر، پرسش دوم را می‌توان چنین خلاصه کرد: این متن را پیدا کرده‌ایم؛ بگویید مؤلفش کیست. پرسش‌های دسته دوم که به اصطلاح باز هستند، بسیار سخت‌تر و پیچیده‌ترند؛ به خصوص اگر بخواهیم مؤلف را از میان مجموعه بسیار بزرگی از نامزدها پیدا کنیم. نوع سوم مسائل حول این موضوع است که تمام خصیصه‌های مؤلف یا مؤلف‌های یک نمونه متن معین شود. به همین دلیل است که برخی محققان ترجیح می‌دهند از واژه‌های سبک‌سنجی^۹ یا پروفایل‌سازی^{۱۰} برای این تحقیقات استفاده کنند و انتساب تألیف را برای دو گونه دیگر مسائل استفاده کنند. نمونه‌های مسائل نوع سوم چنین‌اند: آیا متن یک یا چند مؤلف داشته است؟ آیا مؤلف گویش‌ور بومی بوده است؟ آیا مؤلف بریتانیایی بوده یا آمریکایی؟ آیا مرد بوده یا زن؟ (Juola, 2008). در نوشتار پیش‌رو تلاش خواهیم کرد تا به این سؤال پاسخ دهیم که آیا هریک از رسائل در دیوان سعدی به قلم خود او هستند یا خیر.^{۱۱}

۲. طرح مسئله

رسائل سعدی به مجموعه‌ای از شش نوشتار به نثر گفته می‌شود که در انتساب آن‌ها به شیخ تردید وجود دارد. البته، گاه این تعداد را هفت می‌دانند؛ زیرا مجالس هزل را نیز در این مجموعه جای می‌دهند.^{۱۲} به هر روی، در اینجا ما بر همین شش نوشتار به‌عنوان موضوع مطالعه تمرکز خواهیم کرد. عناوین این شش رساله چنین هستند: «در تقریر دیباچه»، «نصیحة الملوک»، «رساله در عقل و عشق»، «در تربیت یکی از ملوک» (یا رساله «انکیانو»)، «مجالس پنج‌گانه» و «تقریرات ثلاثه».

در برخورد اول با این مجموعه، به آسانی می‌توان دریافت که تفاوتی خصیصه‌نما هم در بین هریک از این رسائل با همدیگر و هم در میان هریک از آنها با کل مجموعه آثار پذیرفته‌شده از سعدی وجود دارد. همین موضوع باعث شده است تا سعدی‌پژوهان به نوعی به این موضوع اشاره کنند و به تحلیل آن پردازند.

فروغی درباره رساله اول با عنوان «تقریر دیباچه» می‌نویسد: «بی‌هیچ تردیدی از شیخ سعدی نیست و چون در نسخ قدیم و معتبر که در دست داریم، نبود، ناگزیر از روی نسخ چاپی و نسخه‌هایی که در قرون دهم کتابت شده تصحیح کردیم» (سعدی، ۱۳۷۲: ۹۲۶). همان‌طور که می‌بینیم، فروغی با قطعیت بسیار انتساب این رساله را به سعدی نفی می‌کند که بی‌شک، بیشتر مبتنی بر شهود و شناخت وی از سعدی است؛ اما در بیان مطلب، به نبود این رساله در نسخ معتبر اشاره می‌کند و استدلالش را بر شهود ادبی خویش بنا نمی‌کند. فروغی از «تقریر دیباچه» و مقدمه منسوب به بیستون با عنوان «ملحقات رسائل» یاد می‌کند. محمدتقی بهار در کتاب *سبک‌شناسی* خود با نوعی تردید و سردرگمی از این رساله و همچنین باقی آنها سخن می‌گوید. بی‌تردید، این سردرگمی ناشی از دور بودن سبک این رساله از سبک باقی آثار شیخ است. بهار در *سبک‌شناسی* آشکارا درباب این رسائل می‌گوید:

در انتساب این رسائل به شیخ جای تردید نیست، مگر رساله ششم که مشتمل بر بعض اغراقات است و می‌نماید که هواداران و غالبان در حق شیخ آن را وضع کرده‌اند؛ اما بعضی از این رسائل نیز در غایت متانت و سختگی است و نمی‌تواند مجعول باشد و لابد تألیف مردی دانشمند از مردم آن زمان باید باشد و چون به شیخ منسوب است، ناچار از اوست خاصه که عبارات شیرین و نمکین سعدیانه نیز گاه‌به‌گاه در آنها یافت می‌شود و شعرهایی که در آنها آمده، از خود نویسنده است قابل هست که از اشعار متوسط شمرده آید و غالباً نیز در ضمن کلیات آمده است (۱۳۳۷: ۳/ ۱۱۲-۱۱۳).

به نظر می‌رسد بهار گرفتار ایده «یا هیچ یا همه» درمورد انتساب این رسائل است و چون برخی از آنها به نظر به قلم شیخ است، لاجرم همه را به قلم شیخ می‌داند؛ از سوی دیگر می‌گوید چون این آثار به شیخ انتساب دارد، به‌ناچار از اوست. این ادعا

بسیار عجیب می‌نماید؛ زیرا بسیارند مواردی از انتساب در نسخ که در حال حاضر به قطع، از خطا بودن آنها آگاهیم.

او در رساله «تقریر دیباچه»، خود به وضوح سبک عجیب و افراطی را می‌بیند؛ اما در توصیف آن به توجیهاات نه‌چندان منطقی دست می‌زند:

این رساله هرچند بالنسبه به نثر قدیم قدری ساده و از تکلفات خام و سجع‌های بارد و صنایع معموله تاحدی عاری است، لیکن از استعمال لغات تازی برای قرینه‌سازی و ازدواج و اشتقاق و آوردن جمله‌هایی متداول و مسجوع از قبیل «تلاطم امواج هموم و تراکم افواج غموم» که معمول به زمان بوده است، خودداری ننموده و هرچه بخواهی در آوردن استدلالات قرآنی که عمده شیوه مجلس‌گویان آن زمان است، به افراط گراییده و تند رفته است (همان‌جا).

این نظر بهار نیز بسیار عجیب است؛ زیرا تمام مواردی که بهار در سبک این رساله برمی‌شمرد، دلایل واضح و روشن برای دوری سبک این رساله از سعدی است؛ اما بهار برای نتیجه‌گیری خود هیچ دلیل و استدلال قانع‌کننده‌ای نمی‌آورد.

رساله دوم «در تقریر مجالس پنج‌گانه» است. فروغی این رساله را منسوب به سعدی می‌داند و می‌گوید: «پنج مجلس است در موعظه که به نظر ما اگر از شیخ هم باشد، در منبر یا مجلسی تقریر فرموده و دیگران به‌نگارش درآورده‌اند. این مجالس در نسخه‌های بسیار قدیم که در دسترس است وجود ندارد و ممکن است افتاده باشد» (سعدی، ۱۳۷۲: ۶۹۷-۶۹۸).

محمدتقی بهار نیز می‌نویسد: «این رساله عبارت است از پنج مجلس، یعنی مفاد یا عین پنج منبر و پنج مجلس وعظ است که شیخ آنها را برای خود یا شاگردان خود نوشته است که از روی آن مجالس تذکیر کنند و مجلس گویند» (۱۳۳۷: ۳/ ۱۱۳). در اینجا نیز می‌بینیم که بهار حتی در توصیف خودش تلویحاً قائل به این است که این‌ها به قلم شیخ نیست؛ اما همچنان آنها را در شمار آثار شیخ می‌پذیرد و می‌نویسد: «شک نیست که مجالس از خود سعدی است؛ ولی نه آن سعدی بذله‌گو و حکیم که ما او را در گلستان دیده‌ایم؛ بلکه سعدی زاهد و صوفی که در حلب یا کاشغر به منبر رفته، موعظه می‌کند و مجلس می‌گوید» (همان، ۱۱۵).

این ایده را ذبیح‌الله صفا نیز طرح می‌کند و می‌نویسد: «مجالس پنج‌گانه که به شیوه مجالس صوفیان و مذکران ترتیب یافته و مسلماً تذکیرهای شیخ بر منبر وعظ بوده است و از سنخ مجالسی است که دیگر مذکران و متصوفان داشته‌اند و مجموعه‌ای از آن‌ها در دست است» (۱۳۶۶: ۳/ ۱۲۱۸).

سه رساله «نصیحة الملوک»، «عقل و عشق» و «انکیانو» را فروغی به قلم خود شیخ می‌داند و می‌نویسد: «آنچه به یقین می‌توان گفت به قلم شیخ است و آن کتاب "نصیحة الملوک" است و رساله "عقل و عشق" و رساله "انکیانو"» (سعدی، ۱۳۷۲: ۶۹۷). بهار این هر سه رساله را از سعدی می‌داند و درباب سبک و شیوه آن‌ها بحث بسیار می‌کند. وی «نصیحة الملوک» را ساده و بی‌تکلف می‌داند و در مقام مقایسه، آن را بی‌تکلف‌تر از گلستان قرار می‌دهد؛ اما فتوحی درباره «انکیانو» و «نصیحة الملوک» می‌نویسد:

دو رساله «نصیحت به سلطان انکیانو» و «نصیحة الملوک» را احتمالاً کسانی خطاب به ارباب قدرت بر ساخته‌اند و آن‌ها را به نام سعدی شهرت داده‌اند که در حافظه جمعی مردم نیز به سرعت مقبولیت یافته است؛ یکی از دلایل بر ساختن «نصیحة الملوک»‌ها به نام سعدی، صراحت و بی‌پروایی شیخ در خطاب به ارباب قدرت است (۱۳۹۱: ۷۳).

البته، جدای از صحت این حرف فتوحی، بی‌شک ایشان باید به این موضوع نیز می‌پرداختند که در صورت جعلی بودن این دو رساله، چه نیازی بود که نام ملکی معاصر سعدی در متن و عنوان آن باشد. آیا جعل این رساله به نام این سلطان نیازی داشته است؟

در نهایت، سه نوشتار به هم پیوسته با عنوان «تقریرات ثلاثه» نیز وجود دارند. فروغی درباره این رسائل می‌نویسد:

چنان‌که از مندرجات آن‌ها آشکار است، دیگران در احوال شیخ نوشته‌اند و تنها اشعاری که در این رسائل گنجانده شده، از شیخ است؛ ولی مطالب این رسائل به نظر ما مشکوک است که از شیخ باشد و در نسخ معتبر و قدیم نیست (همان، ۶۹۸).

فروغی در مقدمه خویش بر رسائل نثر سعدی، آن‌ها را به چهار دسته کلی تقسیم می‌کند و درباب انتساب هریک نظری می‌دهد. ایشان در ابتدای متن رسائل، به‌اجمال و برای جمع‌بندی نظر خویش چنین می‌نویسد:

۱. سه رساله که ظاهراً به قلم خود شیخ سعدی است: رساله نصیحة الملوک، رساله عقل و عشق، رساله انکیانو.

۲. پنج مجلس وعظ که منسوب به شیخ سعدی است.

۳. سه رساله که دیگران در احوال شیخ نوشته‌اند: سؤال صاحب دیوان از

شیخ، ملاقات شیخ با آبا، رساله شمس الدین تازیکو.

۴. شرحی که ابوبکر بیستون در تنظیم کلیات نگاشته است و دیباچه‌ای که بر کلیات نوشته شده (سعدی، ۱۳۷۲: ۸۷۰).

بهار نیز بحث خود را چنین جمع‌بندی می‌کند:

شیخ به چند شیوه نثر نوشته است و دیدیم که قسمتی از نثر شیخ بر شیوه و طریقه معمولی آن زمان - یعنی متکلفانه و بی‌مزه بود - خاصه که موضوع آن نیز دارای مزه خاصی نبود و سراسر مطالبی بود که در قالب عبارات گلستان در نمی‌آید - شیوه دیگر نثر ساده‌ای است که همه‌وقت در ایران نمونه‌هایی از او می‌بینیم چه در کتب علمی و چه در کتب تاریخی - و دو رساله پنجم و هفتم از این مقوله بود (۱۳۳۷: ۳ / ۱۲۳).

۳. روش اول

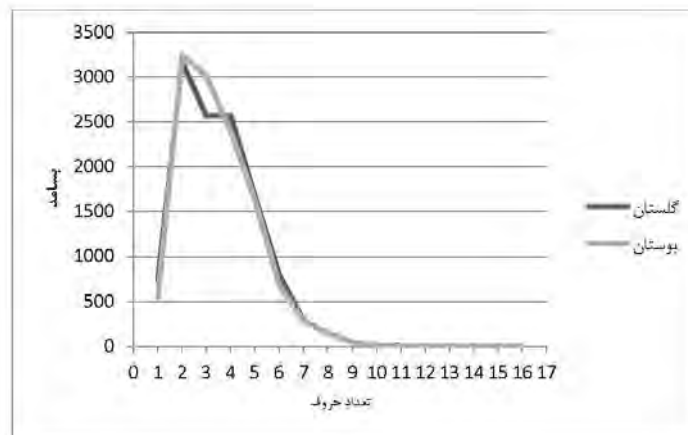
برای بررسی صحت انتساب این رسائل، از دو روش متفاوت بهره خواهیم گرفت. روش اول همانی است که توماس مندنهال^{۱۳} (۱۸۸۷) از آن باعنوان منحنی خصیصه‌نما یاد می‌کند. روش دوم مبتنی بر توضیح تکرار واژگان بر اساس چند مدل آماری است. توماس مندنهال ایده خود را درباره منحنی‌های خصیصه‌نما^{۱۴} اولین بار در سال ۱۸۸۷م در مقاله "The Characteristic Curves of Composition" بیان کرد. وی سپس در سال ۱۹۰۱م در مقاله "A Mechanical Solution of a Literary Problem" که در مجله *Popular Science Monthly* منتشر کرد، ایده‌اش را گسترش داد و آن را در آثار

شکسپیر به کار بست و از آن برای پاسخ به مسئله انتساب برخی آثار مورد تردید وی استفاده کرد.

منحنی‌های خصیصه‌نما که مندنهال از آن‌ها استفاده کرد، به شکلی بسیار ساده، بسامد واژگان متن را براساس تعداد حروف آن‌ها بازنمایی می‌کرد. روش کار او چنین بود که بسامد تمام واژگان با تعداد حروف معین را در متن محاسبه می‌کرد و در نهایت، آن‌ها را با نموداری نمایش می‌داد.

مندنهال در قدم اول نشان داد که نمودارهای حاصل از مجموعه‌های تقریباً بزرگ از متن واحد یا متن‌های مختلف از نویسنده‌ای واحد، با تقریب خوبی یکسان هستند. در واقع، او چنین نتیجه گرفت که این منحنی‌ها مشخصه‌ای ثابت در متن را بازنمایی می‌کنند. در کنار همین موضوع، مندنهال با بررسی متونی از زبان‌های مختلف نشان داد که منحنی‌های حاصل از زبان‌های مختلف تفاوت‌های ساختاری واضحی دارند که کاملاً آن‌ها را از یکدیگر متمایز می‌کند.^{۱۵} سرانجام، با استفاده از این منحنی‌ها و کاربرد آن‌ها بر آثار شکسپیر، بیکن و مارلو تحلیلی از مسئله انتساب مورد مطالعه‌اش به دست داد که اتفاقاً بسیار به یافته‌های پیشین مطالعات سبک‌شناسی و نقد ادبی نزدیک است (Mendenhall, 1901). مندنهال با این کار نه تنها راه‌حلی به اصطلاح خودش «مکانیکی» برای مسئله‌ای ادبی ارائه داد؛ بلکه حتی نشان داد که چقدر چنین راه‌حلی می‌تواند مطابق با شناخت ادبی ما و کارساز باشد.

ما نیز در مرحله اول، از همین روش برای بررسی مسئله پژوهش، یعنی انتساب رسائل سعدی، استفاده کردیم. برای برخورد با مسئله مورد بحث، به پیکره‌ای برای سنجش و مقایسه نیاز داریم. در این میان، دو متن *بوستان* و *گلستان* - یکی به نثر و دیگری به نظم - که هر دو از سعدی است، دو نقطه قابل اتکا برای مقایسه شمرده می‌شود. در مرحله اول، برای اطمینان و سنجش اینکه منحنی‌های مذکور ویژگی مشخصی را از متن بازنمایی می‌کنند، نمودار ۱۲۰۰۰ واژه ابتدایی *بوستان* و *گلستان* را مقایسه می‌کنیم^{۱۶} (شکل ۱).

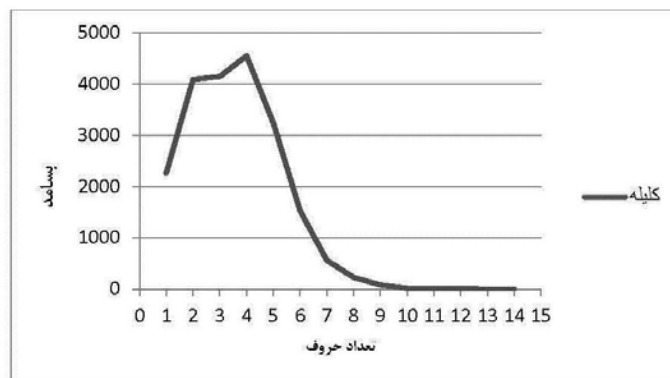


شکل ۱. گلستان و بوستان (۱۲۰۰۰ واژه ابتدایی)

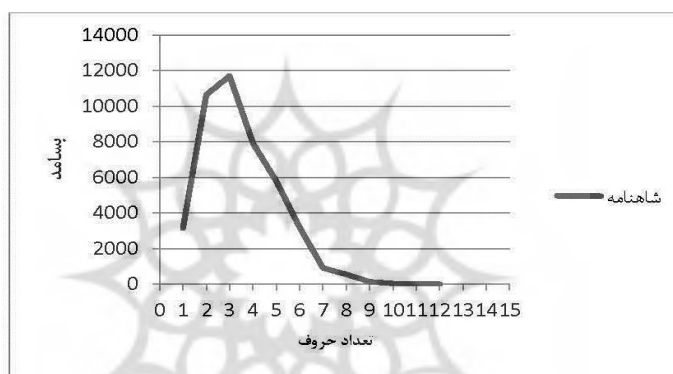
می‌بینیم که دو منحنی نمودار بالا با وجود هماهنگی بسیار زیادی که به لحاظ ساختاری با یکدیگر دارند، تفاوتی جزئی در کلمات سه‌حرفی و چهارحرفی دارند. این موضوع بسیار حائز اهمیت است؛ زیرا نه تنها می‌توانیم مطمئن باشیم که این نمودارها خصیصه‌ای ثابت از نوشتار سعدی را بازنمایی می‌کنند؛ بلکه حتی قادر به تبیین و نمایش تفاوت نظم و نثر او نیز هستند.

از سوی دیگر، برای اطمینان از اینکه آنچه این نمودارها بازنمایی می‌کنند، در مورد هر نویسنده و متن با دیگری متفاوت است، می‌توانیم نمودار حاصل از **کلیده و دمنه** باب الاسد و الثور (شکل ۲) و **شاهنامه** داستان سیاوش (شکل ۳) را با **گلستان** مقایسه کنیم.

پژوهشگاه علوم انسانی و مطالعات فرهنگی
پرتال جامع علوم انسانی



شکل ۲. کلیله و دمنه



شکل ۳. شاهنامه

منحنی حاصل از گلستان، کلیله و شاهنامه تفاوتی اساسی و ساختاری با یکدیگر دارند و این تفاوت آشکار، خود می‌تواند راهگشای ما در تمیز این متون از هم باشد. مقایسه نمودارهای کلیله، شاهنامه، بوستان و گلستان نشان می‌دهد این نمودارها تفاوت‌های معناداری با هم دارند. این تفاوت‌ها بدان معناست که می‌توانیم با قطعیت بپذیریم که این نمودارها خصیصه‌ای متعلق به مؤلف این نوشتارها را بازنمایی می‌کنند؛ یعنی نمودارهای حاصل از بخش‌هایی از یک متن متعلق به هر نویسنده که به لحاظ تعداد واژگان بزرگ باشند، با یکدیگر هماهنگ هستند و با نمودارهای مشابه از

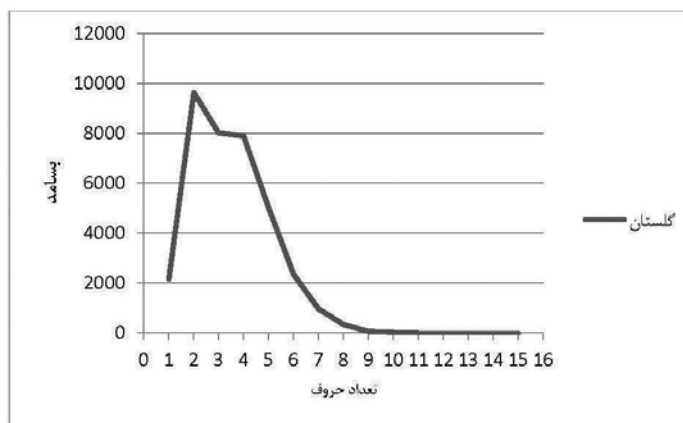
نویسندگان دیگر کاملاً متفاوت‌اند. این تفاوت دلیلی صریح است بر اینکه برای بازنمایی کمی خصیصه‌ای متعلق به یک نویسنده می‌توانیم از آن‌ها بهره ببریم.

همان‌طور که مشاهده می‌کنیم، این تفاوت‌ها به قدری هستند که می‌توان استدلالی کمی را بر آن‌ها بنا کرد؛ برای مثال، در نمودار *شاهنامه* می‌بینیم که کلمات سه‌حرفی قلّه این نمودار را می‌سازند؛ در حالی که در *کلیله و دمنه* نصرالله منشی کلمات چهارحرفی بیشترین بسامد را در متن دارند. با استفاده از این تفاوت‌ها می‌توانیم در مورد متونی که در باب انتساب تألیفشان مشکوک هستیم، به بحث و بررسی کمی بپردازیم.

حال که از توانایی‌های این نوع نمودارها در بازنمایی انتساب متن به مؤلف مشخص مطمئن شدیم، می‌توانیم با مقایسه نمودار حاصل از متن *گلستان* و هر یک از رسائل، به بررسی انتساب رسائل به شیخ بپردازیم. با اطمینان از اینکه *گلستان* نوشته سعدی است، نموداری از *گلستان* را خواهیم ساخت و شباهت یا تفاوت نمودارهای حاصل از رسائل را با این نمودار بررسی خواهیم کرد.

در تمام تحقیقات از این نوع، مشکلی مهم وجود دارد که ناهماهنگی متون مورد آزمایش با متون مبنایی برای بررسی به لحاظ ابعاد و حجم است. در این مورد، مشکل این است که *گلستان* متنی نسبتاً بلند است؛ در حالی که رسائل به نسبت بسیار کوتاه‌ترند و این می‌تواند باعث ایجاد خطا شود؛ زیرا هر قدر ابعاد متن مورد بررسی بزرگ‌تر باشد، با قطعیت بیشتری می‌توانیم از صحت و دقت این بازنمایی اطمینان حاصل کنیم و هر چه این متون کوچک‌تر و نامتناسب‌تر باشند، احتمال خطا بیشتر خواهد شد. در نهایت، در این حالت باید از روش‌های دیگر نیز بهره برد تا بتوان به دقت به صحت بازنمایی یک روش اذعان کرد.

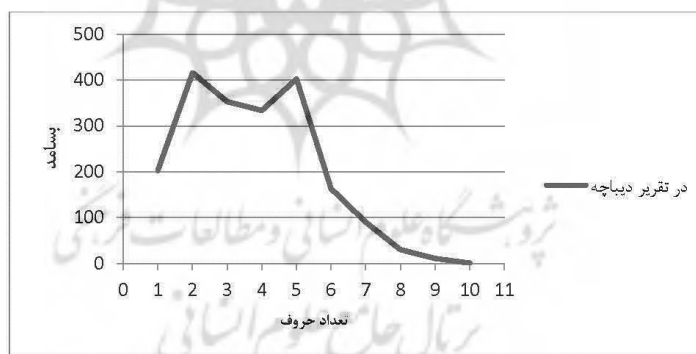
برای شروع این کار ابتدا منحنی کل متن *گلستان* را می‌سازیم (شکل ۴).



شکل ۴. گلستان

حال می‌توانیم نمودارهای تک‌تک رسائل را نیز بسازیم و هریک را با این نمودار مقایسه کنیم.

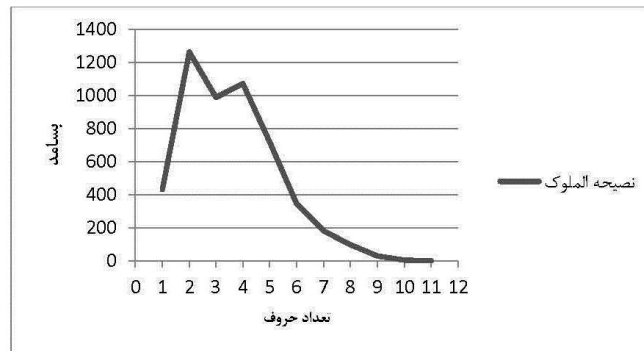
رساله «در تقریر دیباچه» به نظر می‌رسد که به قلم شیخ نباشد. این رساله بسیار با سبک شیخ در شکل و محتوای آثارش متفاوت است و فقط محمدتقی بهار بدون ذکر دلیل، انتساب آن را به شیخ پذیرفته است. حال برای روشن‌تر شدن این انتساب، نگاهی به نتایج شمارش واژگان این رساله می‌اندازیم. نمودار این رساله (شکل ۵) چنین است:



شکل ۵. در «تقریر دیباچه»

همان‌طور که در تحلیل کیفی خویش احتمال دادیم، این رساله از سبک شکلی و محتوایی سعدی بسیار دور است و فروغی نیز انتساب آن را به سعدی رد کرده است. در بررسی نمودار این رساله و مقایسه آن با *گلستان* این موضوع به وضوح مشخص است. حضور بسیار زیاد کلمات پنج‌حرفی «در تقریر دیباچه» تمایز اصلی آن با *گلستان* را می‌سازد. اگرچه حجم این رساله بسیار کمتر از *گلستان* است، تعداد این کلمات پنج‌حرفی به‌طور نسبی نیز بسیار زیاد است. این نسبت حدود ۱۹.۹۷ درصد است؛ درحالی که این عدد در مورد *گلستان* حدود ۱۳.۷۷ درصد است. از سوی دیگر، در رسائل هیچ کلمه‌ای بیش از ده حرف وجود ندارد و فقط یک کلمه ده‌حرفی در کل متن وجود دارد؛ این درحالی است که در *گلستان* با کلمات طولانی‌تر نیز روبه‌رویم. همچنین، میزان حضور کلمات ده‌حرفی در *گلستان* حدود ۰.۰۹ درصد است؛ درحالی که این عدد برای «در تقریر دیباچه» حدود ۰.۰۴ درصد است. با توجه به تمام این‌ها و تفاوت معنادار نمودار *گلستان* و رساله «در تقریر دیباچه» می‌توان با اطمینان زیاد حکم کرد که این رساله به قلم شیخ نیست و بعدها توسط فرد دیگری برای کلیات شیخ آماده شده است. اما نباید فراموش کرد که در چنین تحلیلی نیز که مبنای کار بر فرم اثر است، ردّپایی از سعدی را در این رساله می‌بینیم. همان‌طور که پیش‌تر در مقایسه نمودار *گلستان* با *شاهنامه* و *کلیله و دمنه* نیز دیدیم، بسامد بسیار زیاد کلمات دو‌حرفی از خصایص متن اوست. این موضوع را در نمودار «در تقریر دیباچه» نیز می‌بینیم. بی‌دلیل نیست که همواره در تحلیل‌های همگان، در انتساب این اثر به شیخ تردید وجود داشته است. بی‌راه نیست که جای این گمان را هم محفوظ بداریم که دست‌کم این متن تاحدی تحت تأثیر سعدی است؛ شاید دست‌کم به این دلیل که دیباچه‌ای بر آثار شیخ است، نه متنی مجزا و در انتزاع.

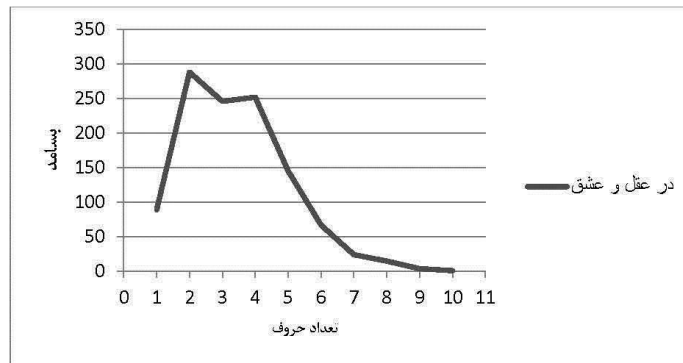
رساله بعدی «نصیحة الملوك» است. این رساله - همان‌طور که پیش‌تر بیان کردیم - قرابت بسیاری دارد به سبک سعدی در *گلستان* و اندیشه وی؛ چونان که در دیگر آثار وی مشاهده می‌کنیم. اندیشه پس این اثر بسیار به *گلستان* و بخش‌هایی از *بوستان* شباهت دارد. نمودار این رساله را مشاهده کنید (شکل ۶):



شکل ۶. «نصیحة الملوك»

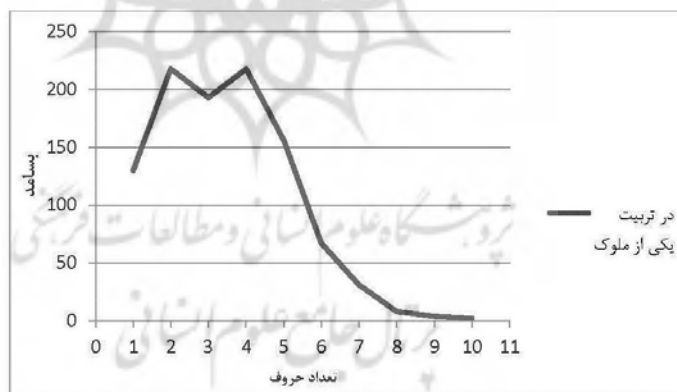
همان‌طور که می‌بینیم، نمودار «نصیحة الملوك» شباهت بسیار زیادی به نمودار گلستان دارد و تنها تفاوت، در بسامد کلمات سه‌حرفی است که در «نصیحة الملوك» مقداری کمتر است. بسامد کلمات سه‌حرفی در گلستان حدود ۲۱.۹۴ درصد و در «نصیحة الملوك» تقریباً ۱۹.۲۴ درصد است که به‌نظر می‌رسد این اختلاف جزئی چندان معنادار نیست؛ مگر اینکه تفاوت اندک در فرم «نصیحة الملوك» با گلستان را و همچنین کم بودن اشعار در این متن را نسبت به گلستان در نظر بگیریم. شباهت دیگر این نمودار با نمودار گلستان در قیاس با رساله «در تقریر دیباجه»، بسامد زیاده‌تر کلمات ده و یازده‌حرفی در آن است. بسامد کلمات ده و یازده‌حرفی در گلستان حدود ۰.۱۱ درصد و در «نصیحة الملوك» حدود ۰.۱۷ درصد است؛ درحالی که این رقم در رساله «در تقریر دیباجه» بسیار کمتر از این بود. هماهنگی بسیار زیاد نمودار «نصیحة الملوك» و گلستان این ایده را بسیار تقویت می‌کند که این رساله به قلم خود شیخ باشد.

رساله بعدی «رساله در عقل و عشق» است (نک: شکل ۷):



شکل ۷. «در عقل و عشق»

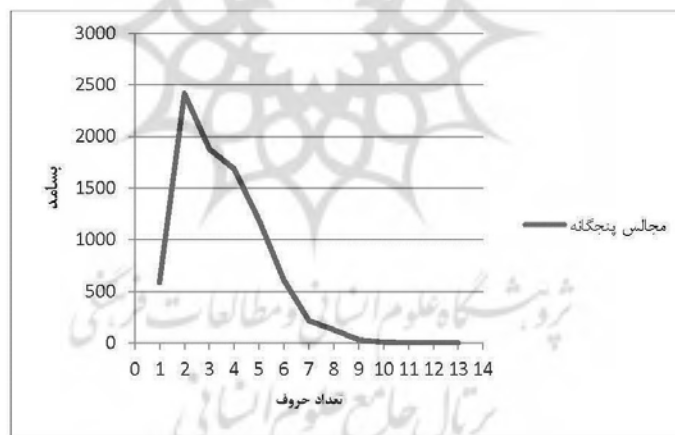
اگرچه این رساله از باقی رساله‌ها کوتاه‌تر و از گلستان بسیار کوتاه‌تر است که این موضوع کار مقایسه را دشوار می‌کند، می‌توان هماهنگی کلی این نمودار را با نمودار گلستان به‌وضوح مشاهده کرد. در واقع، یکی از دلایل این هماهنگی جدای از سبک درونی نوشتار، کنار هم قرار گرفتن نظم و نثر به‌سیاق گلستان در این رساله است. به‌طور کلی می‌توانیم شباهت این دو نمودار را دال بر شباهت خصیصه‌نمایانه این دو متن در نظر بگیریم و چنین استنباط کنیم که این رساله نیز به قلم خود شیخ است. رساله بعدی «رساله در تربیت یکی از ملوک» است. این رساله نیز متنی بسیار کوتاه دارد (نک: شکل ۸):



شکل ۸. «در تربیت یکی از ملوک»

شباهت کلی این نمودار با نمودار **گلستان** مشخص است. تفاوت چشمگیر در این دو نمودار، کمبود نسبی کلمات دوحرفی و سه حرفی است. این موضوع باعث شده است کلمات چهارحرفی در این نمودار یکی از دو قله نمودار را بسازند؛ درحالی که در **گلستان** چنین نیست. باقی رفتارهای نمودار بسیار شبیه به **گلستان** است، جز همین مورد. میزان ظهور کلمات دوحرفی در **گلستان** ۲۶.۲۲۲۵ درصد و کلمات سه حرفی ۲۱.۹۴۵۸ درصد است. در این رساله، این اعداد به ترتیب به ۲۱.۲۲۶۸ درصد و ۱۸.۷۹۲۵ درصد تغییر می کنند. همان طور که مشاهده می کنیم، این تغییر شباهت کلی دو نمودار را چندان به هم نریخته است؛ اما باید توجه کرد که این اختلاف چشمگیر است و نشان دهنده انحراف زیادی از سبک سعدی. تحلیل این اختلاف و دلایل آن را به نتیجه گیری موکول می کنیم.

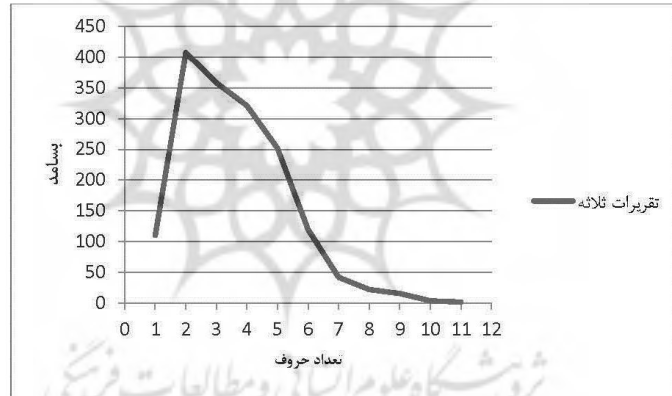
رساله بعدی با عنوان «مجالس پنج گانه» شناخته می شود. این رساله تقریباً بلند است؛ به همین دلیل، انتظار تطابق و هماهنگی نمودار در این رساله بیشتر است؛ زیرا هر قدر طول متن بیشتر باشد، دقت این مقایسه نیز افزایش خواهد یافت. نمودار این متن به این شرح است:



شکل ۹. «مجالس پنج گانه»

این نمودار به طور کلی با نمودار *گلستان* هماهنگ است؛ اما تمایزی با آن دارد که همان بسامد کلمات چهارحرفی است. بسامد کلمات چهارحرفی در *گلستان* ۲۱.۵۹ درصد است؛ در حالی که در «مجالس پنج‌گانه» ۱۹.۲۴۰۴ درصد است. بخش عمده‌ای از این شباهت را باید حاصل ترکیب نظم، نثر و عبارات عربی در این متن بدانیم. این رخداد باعث شده است این متن در فرم ظاهری به *گلستان* شبیه شود. تحلیل این موضوع را به بخش نتیجه‌گیری موکول می‌کنیم.

رساله آخر «تقریرات ثلاثه» نام دارد. این رساله از سه بخش جداگانه تشکیل شده است؛ اما به دو دلیل این سه را در کنار هم و ذیل عنوان «تقریرات ثلاثه» بررسی خواهیم کرد: اول اینکه، این سه بخش همواره ذیل همین عنوان در کلیات شیخ حضور داشته‌اند که پیوند و قرابت این سه بخش را نشان می‌دهد؛ همچنین، در کنار هم قرار گرفتن این سه متن باعث افزایش طول رساله و زیاد شدن دقت بازنمایی خواهد شد. نمودار حاصل از این متن چنین است:



شکل ۱۰. «تقریرات ثلاثه»

همان‌طور که می‌بینیم، این نمودار به طور کلی به نمودار *گلستان* شباهت دارد؛ اما برخی تغییرات جزئی را به وضوح می‌توان در آن ردیابی کرد. این تمایز نیز در کلمات

چهارحرفی است. بسامد این کلمات در «تقریرات ثلاثه» ۱۹.۴۰۷۴ درصد است که از گلستان کمتر است. توضیح و تحلیل این موضوع را در بخش نتیجه گیری خواهیم دید.

۴. روش دوم

روش دوم تلاش برای بازنمایی کمی میزان تکرار واژگان در یک پیکره است. به لحاظ شهودی، میزان تکرار کلمات در متن نماینده ویژگی ای سبکی است که تا حد زیادی مختص همان نویسنده است؛ به این شرح که برخی نویسندگان از کلمات ساده و به کرات بهره می برند و در پرورش ایده خود کمتر از واژگان کم بسامد بهره می گیرند؛ در حالی که نویسندگان دیگر از کلمات کم بسامد و شاذ برای پرورش ایده خود استفاده می کنند.

در اینجا تلاش می کنیم تا این تکراری بودن-کلمات در متن را بازنمایی کنیم. در ابتدا مجموعه ای از اعداد را فرض می کنیم که در این مجموعه، هر عدد بسامد یک کلمه مشخص را در پیکره مورد نظر نشان می دهد؛ بدین ترتیب که مجموعه مورد نظر به ازای هر کلمه در متن مورد بررسی، یک عدد را شامل می شود که این عدد بسامد کلمه مذکور را در متن نشان می دهد؛ بنابراین، تعداد اعضای این مجموعه به تعداد کلمات مستقل موجود در متن بدون در نظر گرفتن تکرارهای آن است.

در این مجموعه، C_n را چنین تعریف می کنیم:

$$\begin{aligned}
 1) \quad C_1(x) &= \frac{T_x}{I_x \log T_x} = \frac{Ave_x}{\log T_x} \\
 2) \quad C_2(x) &= \frac{T_x}{I_x \sqrt{T_x}} = \frac{Ave_x}{I_x \sqrt{T_x}} \\
 3) \quad C_3(x) &= \frac{T_x}{I_x \sigma_x} = \frac{Ave_x}{\sigma_x} \\
 4) \quad C_4(x) &= \frac{T_x}{I_x \log \text{Var}(x)} = \frac{Ave_x}{\log \text{Var}(x)} \\
 5) \quad C_5(x) &= \frac{T_x}{I_x \log \sigma_x} = \frac{Ave_x}{\log \sigma_x} \\
 6) \quad C_6(x) &= \frac{T_x}{I_x \log \text{dif}(x)} = \frac{Ave_x}{\log \text{dif}(x)} \\
 7) \quad C_7(x) &= \frac{T_x}{I_x \sqrt{\text{dif}(x)}} = \frac{Ave_x}{\sqrt{\text{dif}(x)}}
 \end{aligned}$$

$$8) C_8(x) = \frac{T_x}{I_x \ln T_x} = \frac{Ave_x}{\ln T_x}$$

$$9) C_9(x) = \frac{T_x}{I_x \ln \text{Var}(x)} = \frac{Ave_x}{\ln \text{Var}(x)}$$

$$10) C_{10}(x) = \frac{T_x}{I_x \ln \sigma_x} = \frac{Ave_x}{\ln \sigma_x}$$

$$11) C_{11}(x) = \frac{T_x}{I_x \ln \text{dif}(x)} = \frac{Ave_x}{\ln \text{dif}(x)}$$

$$12) C_{12}(x) = \frac{T_x}{I_x \ln \text{dif}(x) \ln T_x} = \frac{Ave_x}{\ln \text{dif}(x) \ln T_x}$$

که در آن‌ها، T_x تعداد کل کلمات متن، I_x تعداد کلمات مستقل متن، $\text{dif}(x)$ اختلاف کمترین و بیشترین بسامد در مجموعه مورد نظر، و Ave_x همان میانگین تکرار کلمات یک متن یا $\frac{T_x}{I_x}$ است.

مقادیر هریک از این متغیرهای اولیه برای متون مورد بررسی ما چنین است:

جدول ۱. متغیرها

	شاهنامه (داستان سیاوش)	کلیله اشیر و گاو	گلستان	در تقریر دبیاچه	تربت یکی از ...	عقل و عشق	مجالس بیت‌گانه	تقریرات ناله	نصیحة الملوک	بوستان (باب اول)
حداقل (Min)	۱	۱	۱	۱	۱	۱	۱	۱	۱	۱
حداکثر (Max)	۲۱۵۳	۲۲۳۹	۲۰۵۷	۳۰۰	۱۲۹	۸۰	۵۵۲	۱۰۹	۴۲۸	۵۰۱
تعداد کل (T)	۴۴۰۰۵	۲۰۶۷۵	۳۶۵۸۲	۲۰۱۳	۱۰۲۷	۱۹۳۲	۸۷۷۰	۱۶۵۴	۵۱۳۹	۱۲۹۹۲
کلمات مستقل (I)	۴۶۳۶	۵۱۰۹	۷۸۷۹	۱۰۱۴	۴۸۲	۵۸۳	۳۰۱۴	۶۹۶	۱۸۳۶	۳۶۶۷
میانگین (Ave)	۹,۴۹۲,۰۱۹	۴,۰۴۶,۷۸	۴,۶۴۲,۹۷۵	۱,۹۳۹,۳۰۶	۲,۱۳۰,۷۰۵	۱,۹۴۱,۶۸۱	۲,۹۰۹,۷۵۴	۲,۳۷۶,۴۳۷	۲,۷۹۹,۰۲	۳,۵۴۲,۹۵۱
تفاوت حداقل و حداکثر (Dif)	۲۱۵۲	۲۲۳۸	۲۰۵۶	۱۹۹	۱۲۸	۷۹	۵۵۱	۱۰۸	۴۲۷	۵۰۰
واریانس (Var)	۲۸۴۹,۹۱۵	۱۲۵۷,۹۵۸	۱۳۷,۲۶۱	۵۲,۹۹۴۳۸	۴۳,۳۸۴۲۱	۲۱,۳۳۳۷۱	۲۲۷,۴۰۰۴	۴۱,۱۲۷۱۶	۱۸۱,۹۸۳	۲۸۷,۵۱۹۴
انحراف معیار (Std)	۵۳,۳۸۴۶	۳۵,۴۶۷۷	۳۷,۰۷۱۰۳	۷,۲۷۹۷۲۴	۶,۵۸۶۸۲۱	۴,۶۰۸۰۰۵	۱۵,۰۷۹۸	۶,۴۱۳,۰۴۶	۱۳,۴۹۰۲۳	۱۶,۹۵۶۴

و همچنین مقادیر همه دوازده تابعی که در بالا بر روی این مجموعه‌ها تعریف کردیم، بدین قرارند:

جدول ۲. خروجی های دوازده تابع

	گلستان	بوستان	کلیله (نبیر و گاو)	شامنامه (داستان سیاوش)	نصیحه الملوک	تقریر دیباجه	مجالس بنج گانه	در تربیت یکی...	عقل و عشق	تقریرات تلاوه
C1	1.017467	0.861261	0.937743	2.044151	0.754274	0.586985	0.737955	0.707506	0.635815	0.73836
C2	0.024275	0.031083	0.028144	209.7737	0.039045	0.043224	0.031071	0.066487	0.05771	0.058433
C3	0.125245	0.208945	0.114098	0.177804	0.207485	0.266398	0.192957	0.32348	0.421371	0.370563
C4	1.479564	1.441005	1.305554	2.747462	1.238483	1.124738	1.234625	1.301312	1.463183	1.472272
C5	2.959128	2.882009	2.611107	5.494924	2.476966	2.249475	2.469251	2.602624	2.926365	2.944544
C6	1.401432	1.312705	1.208045	2.848025	1.064093	0.843597	1.052666	1.01115	1.023215	1.168687
C7	0.102396	0.158446	0.085542	0.204615	0.135454	0.137474	0.120717	0.18833	0.218456	0.228673
C8	0.44188	0.374041	0.407257	0.887763	0.327577	0.254924	0.32049	0.307266	0.276131	0.320666
C9	0.642567	0.62582	0.566995	1.193208	0.537866	0.488467	0.536191	0.565153	0.635452	0.6394
C10	1.285133	1.251641	1.133989	2.386415	1.075733	0.976935	1.072382	1.130305	1.270904	1.278799
C11	0.608634	0.5701	0.524647	1.236882	0.46213	0.36637	0.457167	0.439137	0.444377	0.507555
C12	0.057925	0.060187	0.052799	0.115682	0.054084	0.04816	0.050354	0.063327	0.063196	0.068487

درواقع، هریک از این مقادیر به نوعی بازنماینده تکرار واژگان در متن هستند. برای مقایسه دو متن نیاز داریم تا اختلاف آن‌ها را بسنجیم. این اختلاف را به ازای هر مشخصه از طریق رابطه زیر محاسبه می‌کنیم:

$$\Delta_n(x, y) = \frac{C_n(y) - C_n(x)}{C_n(y) + C_n(x)}$$

نتیجه اختلاف هر دوازده مشخصه از همه این متون با گلستان را در جدول‌های زیر مشاهده می‌کنید. این اعداد به نوعی بیانگر تفاوت گلستان با هریک از این متون در مشخصه مورد نظر هستند. هر قدر این عدد به صفر نزدیک‌تر باشد، این اختلاف کمتر است و می‌توان نتیجه گرفت که مشخصه مورد نظر در توصیف پراکندگی در دو متن شباهت دارد.

جدول ۳. اختلاف

۷	بوستان	کلیده (شیر و گاو)	شاهنامه (داستان سیاوش)	نصیحه الملوک	در تقریر دیباچه	محالس پنج‌گانه	در عقل و عشق	در تربیت یکی...	تقریرات ثلاثه
Δ1(۷ گلستان)	-0.08314	-0.04078	0.33534	-0.14855	-0.2683	-0.15923	-0.23085	-0.17969	-0.15896
Δ2(۷ گلستان)	0.122981	0.073809	0.999769	0.23326	0.28073	0.122791	0.407819	0.465085	0.412995
Δ3(۷ گلستان)	0.250456	-0.04657	0.173434	0.247167	0.360412	0.212796	0.541744	0.441774	0.494784
Δ4(۷ گلستان)	-0.0132	-0.06248	0.29995	-0.0887	-0.13625	-0.09024	-0.00557	-0.0641	-0.00247
Δ5(۷ گلستان)	-0.0132	-0.06248	0.29995	-0.0887	-0.13625	-0.09024	-0.00557	-0.0641	-0.00247
Δ6(۷ گلستان)	-0.03269	-0.07411	0.340418	-0.13682	-0.24848	-0.14212	-0.15599	-0.16177	-0.09056
Δ7(۷ گلستان)	0.214881	-0.08968	0.332949	0.138987	0.146238	0.082115	0.361724	0.295584	0.381422
Δ8(۷ گلستان)	-0.08314	-0.04077	0.33534	-0.14855	-0.26831	-0.15923	-0.23084	-0.17969	-0.15896
Δ9(۷ گلستان)	-0.0132	-0.06248	0.29995	-0.0887	-0.13625	-0.09024	-0.00557	-0.0641	-0.00247
Δ10(۷ گلستان)	-0.0132	-0.06248	0.29995	-0.0887	-0.13625	-0.09024	-0.00557	-0.0641	-0.00247
Δ11(۷ گلستان)	-0.03269	-0.07411	0.340419	-0.13682	-0.24847	-0.14212	-0.15599	-0.16177	-0.09056
Δ12(۷ گلستان)	0.019151	-0.0463	0.332688	-0.03429	-0.09205	-0.06992	0.043518	0.044552	0.083552

برای آزمون معنادار بودن این مشخصه‌ها کافی است به اختلاف *گلستان*، *شاهنامه* و *کلیده* (سه اثر که انتساب و تمایز آن‌ها قطعی است) نگاهی بیندازیم. این موضوع را نیز در نظر بگیریم که *کلیده* در فرم خود بسیار شبیه به *گلستان* است و *شاهنامه* تفاوت بسیار با آن دارد.

مقیاس این اعداد در مورد *شاهنامه* چیزی حدود سه‌برابر است که کاملاً با واقعیت متون هماهنگ است. نخست اینکه *شاهنامه* نوشته فردوسی است و تمایز انتساب واضح؛ از سویی سبک *شاهنامه* تفاوتی به کلی بارز دارد که نمود آن را در این اعداد می‌بینیم؛ از دیگر سوی، اگرچه *کلیده* در مقایسه با *بوستان* برای مثال تفاوت بیشتری با *گلستان* دارد، این تفاوت در مقایسه با *شاهنامه* بسیار کمتر است. دلیل این موضوع فرم و سبک *کلیده* است که نزدیکی بسیاری به *گلستان* دارد و این موضوع در گفتمان ادبیات و سبک‌شناسی و در شهود بارز و واضح است.

حال با استفاده از این خصیصه‌ها تلاش می‌کنیم تا برابندی نهایی از این تمایزات به دست آوریم. برای این کار مجموع اختلافات (میزان اختلاف نهایی یا Δ^S) را با در نظر گرفتن علامت آن‌ها و همچنین میانگین این اختلافات را محاسبه می‌کنیم.

$$\Delta^S(x, y) = \sum_{n=1}^{12} \Delta_n(x, y) = \sum_{n=1}^{12} \frac{C_n(y) - C_n(x)}{C_n(y) + C_n(x)}$$

نتایج را برای متون مورد بررسی در جدول زیر می‌بینیم:

جدول ۴. اختلاف نهایی

Y	بوستان	کلیله (شیر و گاو)	شاهنامه (داستان سیاوش)	نهیجه الملوک	تقریر دیاچه	عقل و عشق	در تربیت یکی...	مجالس پنج‌گانه	تقریرات ثلاثه
$\Delta^S(\text{گلستان}, y)$	۰.۳۲۲۹۸۸	-۰.۵۸۸۴۲	۴.۳۹۰۱۵۹	-۰.۳۴۰۴۱	-۰.۸۸۳۲۱	۰.۵۵۸۸۷۲	۰.۳۰۷۶۷۸	-۰.۶۱۵۸۸	۰.۸۶۳۸۳۷
$\Delta^S(\text{گلستان}, y) / n = 12$	0.026916	-0.04904	۰.۳۶۵۸۲۷	-۰.۰۲۸۳۷	-۰.۰۷۳۶	۰.۰۴۶۵۷۳	۰.۰۲۵۶۴	-۰.۰۵۱۳۳	۰.۰۷۱۹۸۶

همان‌طور که می‌بینیم، این عدد برای بوستان اندک است که البته، قابل پیش‌بینی است. اگرچه بوستان نیز نوشته سعیدی است، تفاوتی اساسی در فرم خود با گلستان دارد. گلستان به نثر و بوستان به نظم نوشته شده و همین موضوع باعث اختلاف اندک این دو متن در میزان تکرار واژگان است. در واقع، همیشه در متون نظم به دلیل تنگنای وزن و قافیه کلمات بیشتر تکرار می‌شوند؛ برخلاف نثر که نویسنده در آن صرفاً تحت تأثیر ذهنیت خویش و سبک شخصی خود، کلمات را تکرار می‌کند.

با نگاهی به کلیله می‌توانیم سند و دلیلی دیگر برای صحت و دقت روش خود بیابیم. کلیله اگرچه نوشته نصرالله منشی است، در فرم و ظاهر خود شباهت بسیار به گلستان دارد. جدای از اینکه هر دو متون نثری هستند که در میان خود اشعاری دارند و از عبارات عربی نیز در مطای کلام بهره می‌برند، به نوعی می‌توان گفت شهود ادبی ما حاکی از این است که برخورد سعیدی با کلمات با برخورد نصرالله با کلمات شباهتی دارد. همین امر باعث می‌شود اختلاف نهایی (Δ^S) برای کلیله بیشتر باشد؛ اما اندازه آن بسیار کمتر از شاهنامه باشد. از سوی دیگر، اتفاق جالب اینجاست که این عدد برای

کلیده مقداری منفی است. این موضوع نشان می‌دهد کلمات در **گلستان** بیش از **کلیده** تکرار می‌شوند که کاملاً با شهود ادبی ما به این دو متن هماهنگ است.

حال با در نظر گرفتن این موضوعات، به نتایج حاصل برای رسائل نگاهی کلی می‌افکنیم و تحلیل نهایی این اطلاعات را به بخش بعد موکول می‌کنیم. در صورت مرتب کردن این متون بر اساس بزرگی عدد اختلاف (بدون توجه به علامت)، رساله اول، «در تربیت یکی از ملوک» است. اختلاف نهایی برای این رساله برابر با ۰.۳۰۷۶۷۸ است که از اختلاف نهایی **بوستان** نیز کمتر است. این بدان معناست که میزان تکرار در این رساله از **گلستان** اندکی بیشتر است که البته، جدای از سبک و بحث انتساب تألیف، تاحدی به سبب حجم نوشتار نیز توجیه‌پذیر است. رساله بعدی «نصیحة الملوک» است که اختلاف نهایی برای آن معادل ۰.۳۴۰۴۱- است. این اختلاف نیز اندک به نظر می‌رسد و نزدیک به همین عدد برای **بوستان**. باید توجه کرد که این عدد برای «نصیحة الملوک» منفی است؛ یعنی در مقام مقایسه، **گلستان** متنی پرتکرارتر از «نصیحة الملوک» است. باید اشاره کرد که این موضوع نیز با شاهدی در متن هماهنگی دارد. در **گلستان** شکل قالب برای بخش‌های متن حکایت است؛ یعنی هر بخش از متن حکایتی است که آغاز و انجامی دارد و همین آغاز و انجام چونان قالبی همیشگی (درست مثل وزن) مایه افزایش تکرار. در حالی که در «نصیحة الملوک» هر جمله و عبارت متمایز از دیگری و مستقل از آن قرار دارد که بی‌شک، مایه تکرار کمتر واژگان در متن است. رساله‌های بعدی به ترتیب بزرگی اختلاف نهایی، «عقل و عشق»، «مجالس پنج‌گانه»، «تقریرات ثلاثه» و «در تقریر دیباچه» هستند.

۵. نتیجه

۱. «در تقریر دیباچه»

همان‌طور که پیش‌تر دیدیم، نمودار حاصل از متن این رساله تفاوت فاحشی با نمودار **گلستان** دارد. این موضوع تفاوت و تمایز بسیار این رساله را با سبک سعدی در **گلستان** و حتی **بوستان** - به جهت شباهت نمودار **بوستان** و **گلستان** - نشان می‌دهد. نتیجه آزمون

پراکندگی و تکرار واژگان نیز در مورد این رساله همین موضوع را تأیید می‌کند. میزان اختلاف نهایی ((دیباچه, گلستان^S)) برای این متن ۰.۸۸۳۲۱- بود. مقداری که بزرگی آن حتی از نتیجه حاصل برای کلیله نصرالله منشی نیز بیشتر بود. از مجموع تمام این داده‌ها می‌توان چنین استنباط کرد که این رساله نباید به قلم شیخ باشد. این نتیجه با حدس کیفی ما مبتنی بر شهود ادبی و نظر دیگر محققان نیز هماهنگ است.

۲. «نصیحة الملوك»

شباهت کلی این نمودار و نمودار گلستان بسیار زیاد است و نزدیکی سبکی این دو متن را نشان می‌دهد. با توجه به نتایج آزمون پراکندگی و تکرار واژگان نیز می‌توانیم تصویر بهتری به دست آوریم. میزان اختلاف نهایی این رساله و گلستان ((نصیحة الملوك, گلستان^S)) معادل ۰.۳۴۰۴۱- بود. این عدد به لحاظ بزرگی، به مقدار اندکی از همین عدد برای بوستان بیشتر است. در واقع، می‌توان گفت شباهت سبکی این متن و گلستان بسیار زیاد است و در کنار گلستان و بوستان قرار می‌گیرد. همان‌طور که پیش‌تر اشاره شد، باید به علامت میزان اختلاف نهایی هم توجه کرد که بسیار مهم است. میزان اختلاف نهایی برای «نصیحة الملوك» منفی است که نشان می‌دهد واژگان «نصیحة الملوك» اندکی کمتر تکراری هستند که کاملاً با شهود ادبی نیز هماهنگ است؛ زیرا این رساله ساختار حکایتی گلستان را ندارد و خلاصه و زبده مطلب به‌طور جداگانه در آن نقل است. این اختلاف اندک خود دلیل بزرگی در انتساب این رساله به سعدی است. با در نظر گرفتن تمام این موارد می‌توان با قاطعیت اعلام کرد که این رساله به قلم خود شیخ است و در انتساب آن به سعدی تردیدی نیست.

۳. رساله «انکیانو» یا «در تربیت یکی از ملوک»

با مقایسه نمودار این رساله و گلستان، آنچه از همه مهم‌تر به نظر می‌رسد، شباهت کلی دو نمودار است. تفاوت اندکی که در کلمات دو و سه حرفی وجود دارد، به دلیل حجم کم این رساله است، نه تفاوت سبکی و خصیصه‌نمایانه آن با گلستان. با توجه به این موضوعات می‌توان این نمودار را دلیلی در اثبات انتساب این رساله به شیخ دانست.

برای تأیید این موضوع می‌توانیم از نتایج تحلیل تکرار واژگان استفاده کنیم. در تحلیل تکرار و پراکندگی واژگان، میزان اختلاف نهایی ((انکیانو، گلستان^S)) برای این رساله معادل ۰.۳۰۷۶۷۸ به دست آمد. کوچک بودن این عدد و حتی کوچک‌تر بودن آن از همین عدد برای **بوستان** دلیل روشن و واضحی است برای نزدیکی سبک این رساله و **گلستان** که مهر تأیید دیگری است بر صحت انتساب این رساله. با در کنار هم قرار دادن این موضوعات می‌توانیم چنین نتیجه بگیریم که صحت انتساب این رساله به شیخ قطعی است و رساله حاصل قلم شیخ است.

۴. «عقل و عشق»

نمودار این رساله شباهت بسیار زیادی به نمودار **گلستان** دارد. کوتاه بودن این رساله نشان‌دهنده شباهت بسیار زیاد سبکی **گلستان** و این رساله است. این شباهت سبکی می‌تواند دلیلی بر صحت انتساب این رساله به سعدی باشد. برای اطمینان از این منظور، به نتایج تحلیل تکرار واژگان رجوع می‌کنیم. در تحلیل تکرار، میزان اختلاف نهایی ((**عشق و عقل، گلستان^S**)) برای این متن معادل ۰.۵۵۸۸۷۲ به دست آمده است. اگرچه این عدد در مقایسه با **بوستان** اندکی بیشتر است، باید توجه کرد که این اختلاف آن قدر نیست که بتوان انتساب این رساله به شیخ را رد کرد. شاید بتوانیم چنین نتیجه بگیریم که این رساله به قلم شیخ است؛ اما حاصل جنبه و ساحتی کاملاً متفاوت با آنچه از شیخ می‌شناسیم. در واقع، همان‌طور که از مضمون رساله نیز پیداست، سعدی «عقل و عشق» با سعدی **گلستان** تفاوت بسیار دارد؛ اما میزان این تفاوت با بررسی کمی ما آن قدر کم است که می‌توان انتساب این رساله به شیخ را پذیرفت.

۵. «تقریرات ثلاثه»

همان‌طور که پیش‌بینی می‌کردیم، نمودار این رساله تفاوت فاحشی با نمودار **گلستان** دارد. این تفاوت فاحش نه تنها نشان‌دهنده تمایز سبکی بسیار زیاد این متن با **گلستان** است؛ بلکه می‌تواند دلیل قاطعی در رد انتساب این رساله به شیخ باشد. برای تأیید این موضوع می‌توانیم از نتایج تحلیل تکرار واژگان کمک بگیریم. در تحلیل میزان تکرار

واژگان، میزان اختلاف نهایی ((تشریحات ثلاثه، گلستان^S)) برای این رساله معادل ۰.۸۶۳۸۳۷ است. این عدد حدود ۲.۶ برابر بوستان و حتی ۱.۵ برابر بزرگی این عدد برای کلیله است. این اختلافات تمایز بسیار این رساله را با گلستان و دوری آن را از سبک سعدی نشان می‌دهد. این نتایج دلیل قاطعی در رد انتساب این اثر به سعدی است.

۶. «مجالس پنج‌گانه»

نمودار این رساله شباهت بسیار زیادی به گلستان دارد؛ اما تفاوت بزرگی نیز در آن وجود دارد. بسامد کلمات چهارحرفی در این رساله به مقدار زیادی کمتر از گلستان است. با توجه به اینکه «مجالس پنج‌گانه» رساله بلندی است، این تفاوت می‌تواند معنادار باشد؛ اما به دلیل اینکه این تفاوت اندک است، ناگزیر باید نتایج تحلیل تکرار واژگان را نیز برای بررسی انتساب این رساله به سعدی مورد توجه قرار دهیم. میزان اختلاف نهایی ((مجالس، گلستان^S)) برای این رساله در آزمون تکرار واژگان معادل ۰.۶۱۵۸۸- است؛ درحالی که این عدد برای کلیله ۰.۵۸۸۴۲- است. همان‌طور که می‌بینیم، بزرگی این عدد برای رسائل از کلیله نصرالله منشی هم بیشتر است؛ این درحالی است که «مجالس پنج‌گانه» بلندترین رساله در میان باقی رسائل است و در صورت نوشته شدن این رساله به قلم سعدی، این طول زیاد متن باید باعث شبیه‌تر شدن متن و سبک آن به گلستان شود؛ درحالی که می‌بینیم نه تنها اختلاف کمتر نشده، بلکه بیشتر نیز شده است؛ بنابراین، می‌توان گفت مجالس به قلم شیخ نیستند. اما با در نظر گرفتن این قول مشهور که این مجالس خطابه‌های سعدی بر منبر بوده‌اند، فرض اولیه ما، اینکه مجالس به قلم شیخ نیستند، صحت دارد. این موضوع می‌تواند به این دلیل باشد که این عبارات حاصل سخنان شفاهی سعدی است. اما باید این مسئله را هم در نظر گرفت که احتمالاً بعد از کتابت این مجلس توسط فردی غیر از شیخ، دستبندی هم در آن‌ها صورت گرفته که این دگرگونی سبکی را ایجاد کرده است.

پی‌نوشت‌ها

۱. شاید به‌عنوان یکی از مقدم‌ترین نمونه‌های ایده‌پردازی برای برخورد ریاضیاتی با مسئله انتساب متون، بتوان به آگوستوس دمرگان، ریاضی‌دان و منطق‌دان بریتانیایی، اشاره کرد. در میان نوشته‌ها و نامه‌های باقی‌مانده از وی، نامه‌ای به تاریخ هجدهم آگوست ۱۸۵۱ خطاب به شخصی به‌نام دبلیو هیلد بسیار درخور توجه است؛ برای مثال، در بخشی از این نامه چنین می‌خوانیم: «امیدوارم این کار را انجام بدهی: نگاهی به هر بخشی از رسائل پولس که با کلمه $\Pi\alpha\tilde{\upsilon}\lambda\omicron\varsigma$ آغاز می‌شود ببنداز- منظورم متن یونانی است- و توجهی به معنای عبارات نکن. سپس همین کار را با رساله‌ای به عبریان انجام بده و تلاش کن تا در ذهن خود در نظر بگیری که آیا دومی شامل لغات طولانی‌تری نسبت به اولی نیست. همواره در ذهن من بوده است که صرف مقداری پول مسائل تألیف از این دست را حل خواهد کرد» (De Morgan, 2010: 215-216).

2. digital humanities

3. Roberto Busa

4. <http://www.corpusthomaticum.org/it/index.age>

4. Mosteller & Wallace

5. Morton

7. *Third Letter to the Corinthians*

8. authorship attribution studies

9. stylometry

10. profiling

۱۱. برای دیدن گزارشی از روش‌ها و ابزارهای مورد استفاده در این گونه مطالعات نک: Stamatatos, 2009; Rudman, 1998; Love, 2002

۱۲. محمدبن ابی‌بکر بیستون در مقدمه خویش بر کلیات می‌نویسد: «پس بدان ای عزیز من که جمع‌آورنده دیوان شیخ در اصل وضع بنیاد بر بیست‌ودو کتاب و رساله کرده بود: شانزده کتاب و شش رساله و بعضی به هفت رساله نوشته بودند؛ چنان‌که بیست‌وسه می‌شد، سبب آنکه رساله مجلس هزل هم در اول اضافه رسالات کرده بودند» (سعدی، ۱۳۷۲: ۹۲۵).

13. Mendenhall

14. characteristic curves

۱۵. یکی از تمایزات قابل توجه قله‌های این نمودارها بود که نشان می‌دهد کدام دسته از کلمات بیشترین بسامد را در آن زبان دارند که در مورد زبان‌های مختلف به‌طور معمول بین دو، سه یا چهارحرفی‌ها قرار دارد. این قله در ساختار اساسی نمودار نقش مهمی دارد (برای دیدن توضیحات مربوط به این موضوع ر.ک: Mendenhall, 1901).

۱۶. نیاز به توضیح نیست که در این پژوهش، معیار شمارش و پایش هر متن، صورت نوشتاری آن بوده و تعداد حروف در صورت نوشتاری مبنای قضاوت است؛ برای مثال، کلمه «خواهر»

پنج حرفی در نظر گرفته می‌شود. همچنین، در این نوع پایش متن، تمام کلمات (اسم، حرف و غیره) مورد بررسی قرار می‌گیرند و کلمات مرکب نیز به تبع، یک کلمه حساب می‌شوند.

منابع

- بهار، محمدتقی (۱۳۳۷). *سبک‌شناسی یا تاریخ تطور نثر فارسی*. ۳ ج. تهران: امیرکبیر.
- سعدی، مصلح‌بن عبدالله (۱۳۵۹). *بوستان سعدی (سعدی‌نامه)*. تصحیح غلامحسین یوسفی. تهران: انجمن استادان زبان و ادبیات فارسی.
- _____ (۱۳۶۸ الف). *بوستان سعدی «سعدی‌نامه»*. تصحیح غلامحسین یوسفی. تهران: خوارزمی.
- _____ (۱۳۶۸ ب). *گلستان*. تصحیح غلامحسین یوسفی. تهران: خوارزمی.
- _____ (۱۳۷۲). *کلیات سعدی*. به اهتمام محمدعلی فروغی. تهران: امیرکبیر.
- صدیقیان، مهین‌دخت (۱۳۷۸). *فرهنگ واژه‌نمای غزلیات سعدی به‌انضمام فرهنگ بسامدی (براساس متن غزلیات سعدی، تصحیح حبیب یغمائی)*. ۳ ج. تهران: پژوهشگاه علوم انسانی و مطالعات فرهنگی.
- صفا، ذبیح‌الله (۱۳۶۶). *تاریخ ادبیات در ایران و در قلمرو زبان فارسی*. ۵ ج. تهران: فردوس.
- فتوحی، محمود (۱۳۹۱). «سبک‌شناسی رسائل سعدی». *سعدی‌شناسی*. دفتر پانزدهم ویژه قصیده. صص ۵۳-۷۸.
- نصرالله منشی، ابوالمعالی (۱۳۷۱). *ترجمه کلیله و دمنه*. تصحیح و توضیح مجتبی مینوی طهران. تهران: امیرکبیر.
- Bahār, M.T. (1958). *Sabkshenāsi yā Tārikh-e Tatavvo-re Nasr-e Fārsi*. 3V. Tehran: AmirKabir. [in Persian]
- Sa'di, M.A. (1980). *Bustān-e Sa'di (Sa'di Nāme)*. Gh. Yusefi (Ed.). Tehran: Anjoman-e Ostādān-e Zabān va Adabiyāt-e Fārsi. [in Persian]
- _____ (1989 a). *Bustān-e Sa'di (Sa'di Nāme)*. Gh. Yusefi (Ed.). Tehran: Khārazmi. [in Persian]
- _____ (1989 b). *Golestān*. Gh. Yusefi (Ed.). Tehran: Khārazmi. [in Persian]
- _____ (1993). *Kolliyāt-e Sa'di*. M.A. Forughi (Ed.). Tehran: AmirKabir. [in Persian]
- Sedighiyān, M. (1999). *Farhang-e VāzheNāme-ye Ghazaliyāte Sa'di be Enzemām-e Farhang-e Basāmadi*. 3V. Tehran: Pazhuheshgāh-e Olum-e Ensāni. [in Persian]
- Safā, Z. (1987). *Tārikh-e Adabiyāt dar Iran va dar Ghalamro'e Zabān-e Fārsi*. 5V. Tehran: Ferdows. [in Persian]

- Fotuhi, M. (2012). "Sabkshenāsi-ye Rasā'el-e Sa'di". *Sa'diShenāsi. Daftare Pānzdahom*. [in Persian]
- Nasrollāh Monshi, A. (1992). *Tarjome-ye Kelile va Demne*. M. Minovi (Ed.). Tehran: AmirKabir. [in Persian]
- DeMorgan, S.E. (2010). *Memoir of Augustus DeMorgan with Selections from His Letters*. Cambridge: Cambridge University Press.
- Hockey, S. (2004). "The History of Humanities Computing" in J. Unsworth, R. Siemens & S. Schreibman. *A Companion to Digital Humanities*. Malden: Blackwell.
- Juola, P. (2008). "Authorship Attribution". *Foundations and Trends in Information Retrieval. Vol. 1. No. 3*. pp. 233-334.
- Love, H. (2002). *Attributing Authorship: An Introduction*. Cambridge: Cambridge University Press.
- Mendenhall, T.C. (1887). "The Characteristic Curves of Composition". *Science*. IX. pp. 237-249.
- Mendenhall, T.C. (1901). "A Mechanical Solution of a Literary Problem". *Popular Science Monthly. Vol. 60*. December.
- Morton, A.Q. (1978). *Literary Detection: How to Prove Authorship and Fraud in Literature and Documents*. New York: Scribner's.
- Mosteller, F. & D.L. Wallace (1964). *Inference and Disputed Authorship: The Federalist*. Reading, MA: Addison-Wesley.
- Rommel, T. (2004). "Literary Studies" in J. Unsworth, R. Siemens & S. Schreibman. *A Companion to Digital Humanities*. Malden: Blackwell.
- Rudman, J. (1998). "The State of Authorship Attribution Studies: Some Problems and Solutions". *Computers and the Humanities. Vol. 31*. pp. 351-365.
- Stamatatos, E. (2009). "A Survey of Modern Authorship Attribution Methods". *Journal of the American Society for Information Science and Technology. 60. No. 3*. pp. 538-556.