A Corpus-assisted Discourse Analysis of 'The Star Online' Columns

Siti Aeisha Joharry
Akademi Pengajian Bahasa, UiTM Shah Alam

Habibah Ismail
Fakulti Pengajian Bahasa Utama, Universiti Sains Islam Malaysia

aeisha@uitm.edu.my

Introduction

Languages used by journalists is often characterised by the writers' personal point of view. This is especially true for columnists whose writings are based on their subjective opinions, usually published in a series (McNair, 2008). Consequently, it is important for columnists to create and maintain relationships with the target audience and in so doing, columnists use the power of language to communicate and interact with their readers. Not only are readers engaged when they respond to columns with their comments, other forms of feedback are now provided such as options to click the 'like' button or from a scale of 'emojis', usually at the end of the article –increasing the challenge for writers to produce a good piece of writing. According to McNair (2008, p109), this is not only entertaining for the reader, the thought "that we will disagree with a columnist whose opinions occupy the other end of the ideological spectrum from our own is often what compels us to read". However, journalistic commentary may require critical review as we are aware of the influence and how changes in the media environment are impacting its demand and supply (Višňovský & Radošinská, 2017).

From a linguistic point of view, writing is perceived as an interactive means whereby social relationships are managed via examining discourse features of how authors project their perceptions and maintain their reader's attention (Hyland, 2005, p. 11). This, according to Hyland is achieved through use of metadiscourse, which refers to how language works in achieving certain communicative purposes for users (ibid, p. 24), and therefore have been shown to be heavily contextualised (Noorian & Biria, 2010; Dafouz-Milne, 2008). Metadiscourse features are usually examined via functional analyses that investigates "the use of language in relation to its surrounding co-text and the purpose of the writer in creating a text as a whole" (Hyland, 2005, p. 24), which is explained next.

Methodology

This study adopts the corpus linguistics approach to investigate naturally-occurring language in online newspaper columns. More specifically, this paper aims to answer the following questions: 1) what are the typical patterns of language used by the columnists and 2) how are these patterns significant to creating the columnists' style of writing. *The Star Online* (online version of the newspaper *The Star*) is chosen mainly for its wide readership.[2] For the purpose of this study, three columnists were chosen, namely Syahredzan Johan, June HL Wong, and M. Veera Pandiyan. Each columnist had more than 130 articles published on the portal between the years 2010 and 2019, which constituted to a total of 513 articles (464,461 words across all three writers).[3] Using corpus techniques, separate word lists are firstly generated using WordSmith tools 6.0 (Scott, 2012) to explore statistical findings for each columnist's

collection of texts, or 'corpus'. By using the consistency analysis feature, frequent words that are used among each corpus in contrast to another, are identified.

Discussion

As mentioned earlier, our first analysis is based on comparison of wordlists using the consistency analysis function in WordSmith, and is presented in Table 1 below.

Table 1: Frequency analysis of words occurring in each corpus

N	Word	Total	Texts	Syahredzan	June	Veera
1	the	29704	513	7039	9301	13364
2	and	13540	513	2018	6424	5098
3	to	12919	513	2836	5290	4793
4	of	12815	513	3037	3990	5788
5	a	9581	513	2001	3994	3586
6	in	9089	513	1660	3438	3991
7	#	7610	513	1217	2292	4101
8	is	5950	511	1689	2280	1981
9	that	5552	513	1370	2537	1645
10	it	5074	513	906	2446	1722
11	for	4375	512	804	1738	1833
12	was	3797	474	431	1827	1539
13	as	3561	512	667	1376	1518
14	s	3286	451	165	1654	1467
15	by	3188	511	767	1077	1344
16	be	3100	505	1032	931	1137
17	on	3053	503	472	1322	1259
18	I	2909	324	163	2380	366
19	are	2848	498	771	1156	921
20	with	2834	506	484	1153	1197
21	not	2548	494	983	805	760
22	but	2405	488	349	1097	959
23	have	2272	497	601	941	730
24	we	2263	415	631	1187	445
25	or	2105	476	739	658	708
26	from	1960	485	331	776	853
27	he	1891	368	189	826	876
28	has	1854	473	355	570	929
29	who	1813	464	283	770	760
30	this	1706	477	578	737	391
31	they	1670	439	359	845	466
32	so	1662	417	238	1178	246
33	his	1655	364	202	729	724
34	at	1509	466	294	596	619
35	their	1497	416	237	760	500
36	there	1474	463	347	603	524
37	were	1393	410	167	572	654
38	an	1378	454	298	526	554
39	which	1307	447	268	508	531
40	what	1307	421	231	801	275
41	my	1253	231	39	1088	126
42	our	1241	361	284	677	280
43	people	1240	413	261	481	498
44	when	1219	424	248	554	417
45	can	1205	417	279	620	306
46	all	1180	425	185	627	368
47	one	1157	429	220	469	468
48	also	1140	442	306	362	472
49	if	1135	418	355	498	282
50	more	1099	417	154	463	482
51	will	1092	341	403	501	188
52	had	1078	369	90	532	456
53	been	1073	432	239	310	524
54	about	1047	397	151	450	446
55	would	1009	399	282	387	340

Table 1 shows a consistency analysis of highly frequent shared words across the three corpora. By using this technique, all three corpora are compared against each other to examine how each columnist over or under-uses certain words. It can be seen that June over-uses most of these words, particularly functional words and pronouns (except for *he* that is higher in Veera's). The use of personal pronoun *I* is seen to be more staggering here: 14 more times in June compared to Syahredzan, and almost 7 more times in Veera, which indicates a marked style for June. Other interesting findings include Syahredzan's use of the negative word *not* more salient than in June's and Veera's, while Veera has more use of numbers (#), passive forms indicated by the lexical item *by* and as discussed earlier, use of the third person pronoun *he*.

Given the high frequency of personal pronouns in June's columns, we shall now focus on the prominent use of 'I' in her writing by following the interactional category of the metadiscourse model proposed by Hyland (2005).[4] The interactional dimension of Hyland's (2005) model refers to the ways writers conduct interaction by explicitly asserting and inviting

readers to respond to their views (p49). Hyland relates these acts as an expression of the writer's 'voice', or "community-recognized personality", which are usually: evaluative and engaging, expressing solidarity, anticipating objections and responding to an imagined dialogue with others (pp. 49-50).

For the purpose of this extended abstract, use of the first person pronoun *I* is examined more closely –using collocational analysis – to discern the typical patterns in which the word is often associated. In June's columns, *I* was found to express self-mention particularly as having experienced something in the past (*I have/was/would/had*), something that she is currently experiencing at the time (*I am/think/can/do/know*) or an action she intends to do in the future (*I will*). These not only demonstrate the explicit author presence in her texts, but as Hyland (2001) notes, it generally represents the writer's decision/choice to adopt a particular stance and a contextually situated authorial identity as well as making reference to shared knowledge with the reader. Due to limited space, we will only discuss the highly frequent occurrence of *I + was*.

Figure 1: Concordance lines for 'I + was' in June's columns

Figure 1 presents examples of 'I was' in context. Firstly, it can be seen that the phrase occurs with several attitude markers (e.g. *I was struck by how South Koreans…, I was amazed*), indicating affective expressions like surprise, frustration and so on. The use of adverbs like *truly* in *I was truly gobsmacked*, and *so* (*I was so intrigued…*) are also found to amplify the writer's expression further. In terms of engagement markers, Hyland (2005, p. 54) states that writers may use certain linguistic devices like reader pronoun (inclusive *we*), question tags, directives and modal verbs to achieve two purposes: to include reader participation in an argument as well as rhetorically positioning the audience to be involved in the discourse. Further inspection of the lines reveal how June engages with her audience, particularly through use of questions (e.g. *Instead, I was amused because I was accused of sleeping my way up at 24, and 30 years on, I was still doing the same thing? I must be one heck of an evergreen femme fatale!*), and references to shared knowledge (e.g. *And that was my first rather bad day on a bicycle. I was ready to abandon it for the rest of holiday but since there was no refund, I had to get back on it*).

As McNair (2008) mentions, the columnist stresses the 'I' as opposed to the detached objectivity of the reporter or correspondent. More importantly, he points out that in the case of commentary columns that cover a range of topics from politics and economics to sport, culture or lifestyle; such as June's, "the idea of the columnist as the journalist of the 'I' is most obvious, since these columns largely comprise whimsical accounts of what 'I' did, or wore, or felt on a particular occasion" (McNair, 2008, p. 110).

Conclusion

This study has briefly examined the use of interactional types of metadiscourse among three Malaysian columnists of *The Star Online*. Findings reveal that June employs the most significant use of self-mentions, indicated by the marked use of the personal pronoun 'I'. The style in which she writes also depicts a typical commentary columnist – colloquial and interactive – mainly expressing what the writer thinks about a certain piece of news (McNair, 2008, p. 109). One explanation could be that the selected type of discourse (or topics) in June's columns mostly revolve around everyday musings. Future work should analyse similar types of columns where over or under-use of certain metadiscourse features could better illustrate the distinctive style among writers. Finally, the study demonstrates the use of corpus-assisted discourse analysis approach, which has shown to be particularly useful in showing stylistic differences among online columnists.

References

Dafouz-Milne, E. (2008). The pragmatic role of textual and interpersonal metadiscourse markers in the construction and attainment of persuasion: A cross-linguistic study of newspaper discourse. *Journal of Pragmatics*, *40*: 95-113.

Hyland, K. (2001). Humble servants of the discipline? Self-mention in research articles. *English for Specific Purposes*, *20*(3): 207-26.

Hyland, K. (2005). *Metadiscourse*: *Exploring Interaction in Writing*. London and New York: Continuum.

McNair, B. (2008). I, Columnist. In Franklin, B. (Ed.), *Pulling Newspapers Apart*: *Analysing Print Journalism* (pp. 106-114). London and New York: Routledge.

Noorian, M. & Biria, R. (2010). Interpersonal metadiscourse in persuasive journalism: A study of texts by American and Iranian EFL Columnists. *Journal of Modern Languages*, *20*: 64-79.

Scott, M. (2012). WordSmith Tools. (Version 6.0). [Computer Software]. Stroud: Lexical Analysis Software. Available from http://lexically.net/wordsmith/downloads/

Višňovský, J., & Radošinská, J. (2017). Online journalism: Current trends and challenges. In Pena Acuna, B. (Ed.), *The Evolution of Media Communication* (pp. 3- 22). Rijeka: InTech.