Man Guo and Qingshun He*

A corpus-based study of the *ADJ-looking* adjectivization in English

<https://doi.org/10.1515/jwl-2021-0017>

Received August 12, 2021; accepted December 6, 2021; published online January 7, 2022

Abstract: This article aims to conduct a corpus-based study of the diachronic and synchronic distributions of a special type of participle adjectivization, the *ADJ-looking* adjectivization. The study based on the Corpus of Historical American English (COHA) finds that this process of adjectivization consists of two phases: (1) The downward rank-shift from the *look ADJ* construction to the *ADJ looking* adjectivization is a process of metaphorization; (2) The transcategorization from the *ADJ looking* adjectivization to the *ADJ-looking* adjectivization is a process of lexicalization. The study based on the Corpus of Contemporary American English (COCA) finds that *ADJ-looking* adjectivizations are mode and register sensitive but not discipline sensitive. The modifier use prefers to occur more in hard science texts to increase the complexity of nominal groups and the predicative use prefers to occur more in soft science texts to increase the grammatical intricacy of sentences. The reason for the non-sensitivity across disciplines is that evaluative adjectives tend to occur in neither soft nor hard science texts.

Keywords: *ADJ-looking* adjectivization; corpus-based; lexicalization; metaphorization; modifier use; predicative use

1 Introduction

By what grammatical steps might a clause complex that straightforwardly realizes a semantic content as in (1a) be shifted to a simple clause realizing the same semantic content as in (1b)?

- (1) a. They shredded the documents before they departed for the airport.
- b. Their shredding of the documents preceded their departure for the airport.

(Halliday and Matthiessen 1999: 343)

***Corresponding author: Qingshun He**, School of Foreign Languages, Sun Yat-sen University, Guangzhou, China, E-mail: heqsh5@mail.sysu.edu.cn

Man Guo, School of Foreign Languages, Sun Yat-sen University, Guangzhou, China, E-mail: gman@mail.sysu.edu.cn

The two clauses in the clause complex in (1a) are shifted to two nominal groups in (1b). This shift is referred to as nominalization in the Hallidayan sense (Halliday 1985, 1994; Halliday and Matthiessen 1999, 2004, 2014). In the process of nominalization, the conjunction group *before* realizing the logico-semantic relation between the two clauses is correspondingly shifted to the verbal group *preceded* realizing process in the simple clause in (1b). This shift is referred to as verbalization (He and Yang 2018). With the nominalization of the verbal groups, the personal pronoun *they* has been shifted to the processive pronoun *their* to function as the modifiers of the head nouns *shredding* and *departure*. This process of downward rank-shift is referred to as a type of adjectivization of nominal groups identified by He (2019b). It is from nominalization, verbalization, and adjectivization that the clause complex in (1a) has been shifted to a simple clause in (1b).

Besides nominal groups, participles also have the potential to be adjectivized. When a participle modifying the head noun (e.g. *the moving blackboard*, *the sleeping lion*) can be intensified or accept degrees of comparison (e.g. *the very moving story*, *the most touching poem*) (Biber et al. 1999), adjectivization occurs. This type of adjectivization can have predicative use. For example:

- (2) a. It was entertaining, instructive and often *very touching*. (COHA_1972)
- b. I haven't even mentioned the *most touching* part. (COHA_2005)

According to Halliday (2004), participles functioning as modifiers can themselves be further modified, resulting in complex modifiers (e.g. *the fast-moving car*, *the well-boiled water*). In *the fast-moving car*, for example, *fast* is the modifier of *moving* which can itself function as the modifier of the head noun *car* (e.g. *the moving car*). Similarly, in *the well-boiled water*, *boiled* can itself function as the modifier of the head noun *water*. However, not all complex modifiers can be regarded as being further modified. Some participles can only function in complex modifiers (e.g. *a good-looking boy*, *a fund-raising activity*). In *a good-looking boy*, for example, *looking* cannot be considered as the modifier of *boy*, and in *a fund-raising activity*, *fund* has no potential to function as the modifier of *raising*. These complex modifiers violate the principle of right-headedness (see Williams 1981).

Some complex modifiers can be intensified or accept degrees of comparison and have predicative use. For example, *fast-moving* can be intensified as in *the very fast-moving car*, and *the car is very fast-moving* is also acceptable. However, *fund-raising* in *a fund-raising activity* cannot be intensified by *very* or have predicative use. According to He (2019a), adjectivizations function as Classifier or experiential Epithet (Epithet₂) in nominal groups. In *the fast-moving car*, *fast-moving* functions as experiential Epithet and in *a fund-raising activity*, *fund-raising* functions as Classifier. Classifier adjectives cannot be intensified or accept degrees of comparison.

According to He and Guo (2021), a participle cannot be adjectivized by itself. Rather, it is the auxiliary verb *be* and the present or past participle together that have the potential to be adjectivized. For example, the verbal groups *is pleased* in (3a) and *is worrying* in (3b) can be adjectivized as *pleased* (e.g. *the pleased city*) and *worrying* (e.g. *the worrying advertising industry*) respectively. Seen from this perspective, the adjectivization of participles is also a process of downward rank-shift as the adjectivization of nominal groups.

- (3) a. The city *is pleased* by what it has seen so far. (COCA_ACAD)
- b. The enormous cost of sports programming *is worrying* the advertising industry. (COCA_NEWS)

The *ADJ-V-ing* adjectivization is also a type of participle adjectivization that is getting increasingly prevalently used in the Present-Day English (Sun and Baayen 2021; Vartiainen 2016). Other popular sense-perception ascriptive verbs are *sound*, *smell*, *feel* and *taste* (Halliday and Matthiessen 2014: 269). In the present research, we will take the *ADJ-looking* adjectivization as our research target because the frequency of *looking* in this type of adjectivization is much higher than those of the other four sense-perception verbs in the corpus, accounting for 86.37% of the total five *ADJ-V-ing* adjectivizations.¹

The *ADJ-looking* adjectivization should have originated from the *look ADJ* construction. According to Vartiainen (2016), the *ADJ looking* complex came into being in the 18th century mainly functioning as noun modifier. From the early 19th century, the hyphenated *ADJ-looking* and its predicative use started to develop. A hyphen is added in between to reduce the possibility of parsing the entire construction into two separated units (Sanchez-Stockhammer 2018; Sun and Baayen 2021). For example:

- (4) a. I saw a *good looking* woman sit at work. (1765)
- b. He was a *very shabby looking* Fellow. (1778)
- Vartiainen (2016: 13)
- (5) a. I remember that she was *fine-looking*. (COHA_1855)
- b. They say she's *very good-looking*. (COHA_1895)

It should be noted that in the *look ADJ* construction, the *ADJ* is always evaluative because *look* construes a relational process and evaluative adjectives are always gradable. After the *look ADJ* construction has been adjectivized as *ADJ-looking*, this

¹ We retrieved 16,934 *ADJ-looking* adjectivizations, 927 *ADJ-sounding* adjectivizations, 987 *ADJ-smelling* adjectivizations, 527 *ADJ-feeling* adjectivizations and 231 *ADJ-tasting* adjectivizations in the COHA.

adjectivization is still evaluative. It, therefore, functions as the interpersonal Epithet in the nominal group. For example, the adjectives *good-looking* and *young* in *the good-looking young boy* realize the interpersonal Epithet and the experiential Epithet respectively.

Nominalization in the Hallidayan sense is a characteristic feature of scientific writing (Biber 2006; Biber et al. 1999; Galve 1998; Halliday 1998, 2004), because it gives scientific writing “an appearance of solidity, stability, and fixed factuality” (Banks 2005: 350). According to He and Yang (2018), nominalizations are more prevalent in written texts than in spoken texts with regard to mode distribution, and more prevalent in academic texts than in fiction texts with regard to register distribution. However, it is verbalization rather than nominalization that is discipline sensitive. Verbalizations are more prevalent in hard science texts than in soft science texts. Since nominalization, verbalization, and adjectivization are the three resources for creating ideational grammatical metaphor² (He 2019b), adjectivizations should have a similar distribution pattern to nominalizations and verbalizations.

However, of the two types of adjectivization (i.e. nominal group adjectivization and participle adjectivization), nominal group adjectivizations like nominalizations and verbalizations tend to occur more in formal academic texts with regard to register distribution. Participle adjectivizations, however, are oppositely distributed because participle adjectivizations function mostly as interpersonal Epithet, and “the speaker involvement is not a characteristic feature of academic writing” (He and Guo 2021: 13). As a special type of participle adjectivization, *ADJ-looking* adjectivizations should have a similar distribution pattern to present or past participle adjectivizations. That is, they should occur more in the relatively non-formal fiction texts rather than in the formal academic texts.

In this study, we will investigate the diachronic and synchronic distributions of the *ADJ-looking* adjectivizations. For this purpose, we will employ real language data to conduct a corpus-based study of the *look ADJ* constructions, the *ADJ looking* adjectivizations, and the *ADJ-looking* adjectivizations. We can hereby work on the following two hypotheses: (1) The *look ADJ* constructions have the potential to shift to the *ADJ looking* adjectivizations which will be further lexicalized as the *ADJ-looking* adjectivizations, and the modifier use of the *ADJ-looking* adjectivizations has the potential to shift to the predicative use; (2) *ADJ-looking* adjectivizations are more prevalent in spoken texts than in written texts with regard to mode distribution, more prevalent in fiction texts than in academic texts with regard to register distribution, and more prevalent in soft science texts than in hard science texts with regard to discipline distribution.

² Ideational grammatical metaphor arises from the semogenic rank-shift (Halliday and Matthiessen 1999; Yang 2020). For example, the clause *the engine failed* can be shifted to the nominal group *the engine failure*.

We will introduce the corpora and data collection in Section 2. Then we will present the diachronic and synchronic distributions of the *ADJ-looking* adjectivizations in Sections 3 and 4 respectively. The findings of this study will be discussed in Section 5.

2 Methodology

2.1 Corpora

The corpora we used in the present research are the COHA and the COCA. The COHA is the largest structured corpus (406 million words) of historical English, covering a time span of 200 years from 1810 to 2009. This permits the “analysis of historical change on a scale not imaginable with earlier corpus designs” (Biber and Gray 2013: 108). Regarding the study of the *ADJ-looking* adjectivization, we can use this corpus to explore the diachronic distribution of this construction across decades over the past 200 years.

We use the COCA because this corpus is the only large, register-balanced corpus of American English. The COCA has been expanded by 20 million words each year since the early 1990s. It has reached one billion words till 2020 and new registers have been added. Regarding the present study, we will only use 600 million words distributed over five sections, i.e. Spoken, Fiction, Popular Magazines, Newspaper, and Academic (around 20% in each). The academic register is further organized with regard to disciplines (e.g. History, Education, Geography/Social Science, Law/Political Science, Humanities, Philosophy/Religion, Science/Technology, Medicine, Miscellaneous, and Business). This can facilitate the investigation of the synchronic distribution of the *ADJ-looking* adjectivizations across modes, registers, and disciplines.

Both the COHA and the COCA are available online.³ They also provide convenient search tools. We can write search queries following the query syntaxes to retrieve the *ADJ-looking* adjectivizations.

2.2 Data collection

According to the query syntaxes provided by the corpora, we wrote the following search queries:

- SQ 1. -[j*] [look] [j*]

³ <https://www.english-corpora.org/>.

- SQ 2. [j*] looking [nn*]
- SQ 3. [jjr*] looking [nn*]
- SQ 4. more|most|very [j*] looking [nn*]
- SQ 5. [vb*] [j*] looking [y*]⁴

This first search query can be described as any construction consisting of *look* and an adjective not following any adjective. The second to the fourth search query can be described as any noun preceded by an adjective or a comparative or superlative adjective and *looking*. The fifth search query can be described as any construction consisting of an adjective and *looking* following a linking verb and followed by a punctuation mark. Using these five search queries we can retrieve concordance lines as in (6):

- (6)
- a. We will have to *look sharp*. (COHA_1919)
 - b. That's a *suspicious looking locket* about her neck. (COHA_1821)
 - c. The judge was the *younger looking man*. (COHA_1859)
 - d. Each class is in charge of an older and *more responsible looking teacher*. (COHA_1938)
 - e. And on top of that, Mason Pew *was good looking*. (COHA_1988)

Note that the occurrences retrieved using the second search query subtract those retrieved using the fourth search query will result in the modifier *ADJ looking* adjectivizations without a degree modifier. We then wrote the following search queries to retrieve the *ADJ-looking* adjectivizations.

- SQ 6. *-looking [nn*]
- SQ 7. more|most|very *-looking [nn*]
- SQ 8. [vb*] *-looking -[nn*]

The sixth search query can be described as any *ADJ-looking* adjectivization followed by a noun. The seventh search query can be described as any *ADJ-looking* adjectivization with intensifier or degree modifier followed by a noun. The eighth search query can be described as any *ADJ-looking* adjectivization following any form of the relational verb *be*. For example:

⁴ [j*] adjective (e.g. *good, happy, electrical*)
 [nn*] common noun (e.g. *time, people, headquarters*)
 [jjr*] comparative adjective (e.g. *better, happier*)
 | or
 [vb*] BE verb (e.g. *is, were, being*)
 [y*] punctuation (e.g., ., !, ?)

- (7) a. His deputy was an *aimless-looking goof* with fresh acne scars.
(COHA_2006)
- b. It was a very thin, *very flat-looking weapon*. (COHA_2005)
- c. Mr. Geller *is soft-looking* and bottom-heavy like an old pear.
(COHA_1992)

Similarly, subtracting the occurrences retrieved using the seventh search query from the occurrences retrieved using the sixth search query will result in the modifier *ADJ-looking* adjectivizations without intensifier or degree modifier. The data we retrieved from the two corpora are shown in Tables 1 and 2.

3 Diachronic distribution of the *ADJ-looking* adjectivizations

In this section, we investigated the diachronic distribution of the *ADJ-looking* adjectivizations.

3.1 Comparing the *ADJ looking* and the *ADJ-looking* adjectivizations

We added up the occurrences retrieved using the second and the fifth search queries for the frequency of the *ADJ looking* adjectivizations. Similarly, we added up the occurrences retrieved using the sixth and eighth search queries for the total frequency of the *ADJ-looking* adjectivizations. The size of the COHA is not evenly divided across decades. To facilitate comparison, we converted the raw frequencies in each sub-corpora to the normalized frequencies of per million words⁵ and then further converted the normalized total frequency of the *ADJ looking* adjectivizations (111.81) and that of the *ADJ-looking* adjectivizations (451.65) to that of the *look ADJ* constructions (1,678.34) (see Figure 1).

Figure 1 shows that the diachronic distribution of the *look ADJ* constructions maintains an overall trend of increase, whereas that of the *ADJ looking* adjectivizations shows a general trend of decrease. The diachronic distribution of the *look ADJ* constructions and that of the *ADJ looking* adjectivizations are not necessarily correlated. This is not in agreement with our hypothesis that the *look ADJ* constructions will shift to the *ADJ looking* adjectivizations. It is also shown that the

⁵ In this research, all the raw frequencies were normalized to the frequencies of per million words and then to equal totality for comparison.

Table 1: Data retrieved from the COHA.

Period	Look ADJ	ADJ looking			ADJ-looking			Total
		Modifier	Degree modifier	Predicative	Modifier	Degree modifier	Predicative	
1810s	52	12	1	0	2	0	0	67
1820s	210	90	6	0	36	2	0	344
1830s	479	143	7	0	254	14	5	902
1840s	721	214	15	5	409	22	17	1,403
1850s	947	132	12	2	531	27	12	1,663
1860s	1,193	104	7	0	432	22	19	1,777
1870s	1,332	78	5	3	508	29	15	1,970
1880s	1,625	98	10	3	505	28	14	2,283
1890s	1,218	69	8	4	557	32	22	1,910
1900s	1,533	85	5	2	473	16	25	2,139
1910s	1,688	128	8	4	485	29	34	2,376
1920s	1,819	141	12	10	538	21	40	2,581
1930s	2,360	97	8	9	554	17	34	3,079
1940s	2,759	70	5	2	550	16	46	3,448
1950s	2,636	58	5	4	524	25	37	3,289
1960s	2,552	50	4	4	482	18	41	3,151
1970s	2,846	68	3	2	574	16	32	3,541
1980s	3,273	44	2	10	576	18	44	3,967
1990s	4,180	61	1	8	582	23	48	4,903
2000s	4,432	54	6	14	681	24	49	5,260
Total	37,855	1,796	130	86	9,253	399	534	50,053

Table 2: Data retrieved from the COCA.

Section	Look ADJ	ADJ looking		ADJ-looking		Total
		Modifier	Predicative	Modifier	Predicative	
Spoken	11,572	315	63	772	101	12,823
Fiction	27,435	468	50	3,679	257	31,889
Magazine	11,473	148	16	2,467	137	14,241
Newspaper	6,516	94	6	1,213	91	7,920
Academic	1,538	57	5	459	29	2,088
Total	58,534	1,082	140	8,590	615	6,8961

Figure 1: Diachronic distributions of the look ADJ, ADJ looking and ADJ-looking constructions.

diachronic distribution of the *ADJ-looking* adjectivizations is leveling off since the 1850s and this distribution trend is significantly negatively correlated with that of the *ADJ looking* adjectivizations at the 0.05 level ($r = -0.461$; $p = 0.041 < 0.05$), indicating that the *ADJ looking* adjectivizations have the potential to shift to the *ADJ-looking* adjectivizations.

The *look ADJ* construction functions as process and attribute in relational clauses, and so the *ADJ looking* adjectivization and the *ADJ-looking* adjectivization maintain the same nature and can be intensified or accept degrees of comparison. The question is whether the comparative *ADJ-looking* adjectivization originated directly from the comparative *ADJ looking* adjectivization (e.g. *good looking* → *better looking* → *better-looking*) or from a further word-formation process (e.g. *good looking* → *good-looking* → *more good-looking*).

In the following, we took the adjective *good* as an example to compare the diachronic distribution of the *better-looking* adjectivization and that of the *more good-looking* adjectivization. The reason for choosing *good* is that the frequency of the *good-looking* adjectivization is the highest (3,106) of all the *ADJ-looking* adjectivizations (16,934), accounting for over 18.34%. Another reason is that *good* has the comparative form *better* (e.g. *a better-looking boy*).

The data we retrieved from the COHA regarding the comparative *good-looking* adjectivization are shown in Table 3, and the diachronic distributions of the normalized frequencies of the data are shown in Figure 2.

It can be seen from Table 3 and Figure 2 that of all the five groups of data *good-looking* adjectivizations count the most. This type of adjectivization keeps a general trend of increase over the past 200 years. The diachronic distribution of the *good looking* adjectivization, however, does not show an obvious trend of increase or decrease. Both the *better looking* and *better-looking* adjectivizations show a general trend of increase. Bivariate correlation test shows that the diachronic distribution

Table 3: Diachronic distributions of comparative *good-looking* adjectivizations.

Period	<i>Good looking</i>	<i>Good-looking</i>	<i>Better looking</i>	<i>Better-looking</i>	<i>More good-looking</i>	Total
1810s	1	3	0	0	0	4
1820s	10	25	1	0	0	36
1830s	18	61	3	2	0	84
1840s	16	41	0	4	0	61
1850s	8	66	4	1	0	79
1860s	12	58	1	5	0	76
1870s	11	101	3	6	0	121
1880s	10	96	4	4	0	114
1890s	26	103	9	0	0	138
1900s	39	133	3	6	0	181
1910s	62	185	18	8	0	273
1920s	46	224	12	10	0	292
1930s	21	195	11	11	1	239
1940s	34	178	9	12	0	233
1950s	32	174	9	12	0	227
1960s	25	229	12	17	0	283
1970s	25	196	18	21	1	261
1980s	32	215	13	23	2	285
1990s	47	268	23	21	3	362
2000s	40	172	12	18	1	243
Total	515	2,723	165	181	8	3,592

Figure 2: Diachronic distributions of comparative *good-looking* adjectivizations.

of the *better-looking* adjectivization and that of the *better looking* adjectivization are significantly correlated at the 0.01 level ($r = 0.610$; $p = 0.004$), indicating that the *better looking* adjectivization does not have the potential to shift to the *better-looking* adjectivization; rather, they are of the same group of data.

There are only a very small number of *more good-looking* constructions as in (8). However, of the eight *more good-looking* constructions, there is still one in which *more* functions as Numerative in the nominal group as in (9).

- (8) a. His seven-year-old son was even *more good-looking*. (COHA_1985)
 b. He was even *more good-looking* now than when she'd met him five years ago. (COHA_2009)

- (9) You get *more good-looking* envelopes than anybody here. (COHA_1993)

It can be concluded that the *better-looking* adjectivization does not occur later than the *good-looking* adjectivization; rather it occurs while the *good looking* adjectivization is shifting to the *good-looking* adjectivization. It can also be seen that, in recent years, the *good-looking* adjectivization has begun to accept degrees of comparison, indicating the full lexicalization.

3.2 Comparing modifier and predicative uses of ADJ-looking adjectivizations

In this section, we compared the diachronic distribution of the modifier ADJ-looking adjectivizations and that of the predicative ADJ-looking adjectivizations. To facilitate comparison, we converted the normalized total frequencies of

the modifier *ADJ looking* adjectivizations (108.2), the predicative *ADJ looking* adjectivizations (3.61) and the predicative *ADJ-looking* adjectivizations (22.97) to that of the modifier *ADJ-looking* adjectivizations (428.68) (see Figure 3).

Figure 3: Diachronic distributions of the modifier and predicative uses of the *ADJ looking* and *ADJ-looking* adjectivizations.

Figure 3 shows that, in general, the modifier use of the *ADJ looking* adjectivizations is decreasing while the predicative use is fluctuating. Bivariate correlation test shows that the two groups of data are negatively but not significantly correlated ($r = -0.399$; $p = 0.08 > 0.05$), indicating that some of the modifier *ADJ looking* adjectivizations do not shift to the predicative use.

As shown in Table 1, the raw frequencies of the modifier and the predicative *ADJ looking* adjectivizations are 1,796 and 86 respectively, the predicative use accounting for 4.57% of the total, whereas the raw frequencies of the modifier and the predicative *ADJ-looking* adjectivizations are 9,253 and 534 respectively, the predicative use accounting for 5.46% of the total. This also indicates that some modifier *ADJ looking* adjectivizations have not yet been ready to shift to predicative use.

The diachronic distribution of the modifier *ADJ-looking* adjectivizations increases quickly over the first five decades, then begins to decrease slowly from the 1850s to the 1900s, and levels off from the 1900s to the 2000s, whereas that of the predicative use, though sometimes fluctuates, maintains an overall trend of increase. The two groups of data are significantly correlated at the 0.05 level ($r = 0.492$; $p = 0.028 < 0.05$), indicating that the *ADJ-looking* adjectivizations have both the modifier use and the predicative use.

The diachronic distribution of the modifier *ADJ looking* adjectivizations and that of the modifier *ADJ-looking* adjectivizations are negatively correlated

($r = -0.426$; $p = 0.061 > 0.05$), although not significantly. The distributions of the predicative *ADJ looking* adjectivizations and the predicative *ADJ-looking* adjectivizations are even less correlated ($r = 0.379$; $p = 0.133 > 0.05$). This indicates that the modifier *ADJ looking* adjectivizations have the potential to shift to the modifier *ADJ-looking* adjectivizations. However, the distribution of the predicative *ADJ-looking* adjectivizations is significantly negatively correlated with that of the modifier *ADJ looking* adjectivizations at the 0.01 level ($r = -0.772$; $p = 0.000 < 0.01$), indicating that the predicative *ADJ-looking* adjectivizations are shifted from the modifier *ADJ looking* adjectivizations. The predicative use of the *ADJ-looking* adjectivizations is significantly correlated with the predicative use of the *ADJ looking* adjectivizations at the 0.01 level ($r = 0.719$; $p = 0.000 < 0.01$), indicating that they are of the same group of data, both being shifted from the modifier *ADJ looking* adjectivizations.

From the above analysis, we can conclude that the *ADJ looking* adjectivizations tend less and less to occur at the modifier place in nominal groups. A small number of modifier *ADJ looking* adjectivizations have been shifted to the predicative use and a larger number to either the modifier use or the predicative use of the *ADJ-looking* adjectivizations.

4 Synchronic distribution of *ADJ-looking* adjectivizations

In this section, we investigated the synchronic distributions of the *ADJ-looking* adjectivizations across modes, registers, and disciplines.

4.1 Mode distribution of *ADJ-looking* adjectivizations

As shown in Table 2, we collected 12,823 and 56,138 occurrences of the total three types of constructions in the spoken sub-corpus and the written sub-corpus. To facilitate comparison, we converted the total normalized frequency of the *ADJ looking* adjectivizations (4.69) and that of the *ADJ-looking* adjectivizations (23.83) to that of the *look ADJ* constructions (186.52) in the COCA (see Figure 4).

Figure 4 shows that under equal total frequencies, the distributions of the three groups of data are significantly different in spoken texts and written texts ($\chi^2 = 45.051$, $df = 2$, $p = 0.000 < 0.01$) and the correlation strength is modest ($F = 0.284$). The *look ADJ* constructions are not obviously mode sensitive. The *ADJ looking* adjectivizations are more prevalent in spoken texts than in written texts,

Figure 4: Mode distribution of the *ADJ-looking* adjectivizations.

whereas the *ADJ-looking* adjectivizations are more prevalent in written texts. This is in agreement with our hypothesis that the *ADJ-looking* adjectivizations tend to occur more in spoken texts than in written texts with regard to mode distribution.

In the following, we compared the mode distributions of the modifier and predicative *ADJ-looking* adjectivizations. To facilitate comparison, we converted the normalized total frequencies of the modifier *ADJ looking* adjectivizations (4.04), the predicative *ADJ looking* adjectivizations (0.65) and the predicative *ADJ-looking* adjectivizations (1.84) to that of the modifier *ADJ-looking* adjectivizations (21.99) (see Figure 5).

Figure 5: Mode distributions of the modifier and predicative uses of the *ADJ looking* adjectivizations and the *ADJ-looking* adjectivizations.

Figure 5 shows that both the modifier use and the predicative use of the two types of adjectivizations are mode sensitive. The distributions of the four groups of data are significantly different in spoken texts and in written texts ($\chi^2 = 12.416$, $df = 3$, $p = 0.008 < 0.01$) and the correlation strength is moderate ($F = 0.371$). Generally, both

the modifier and the predicative *ADJ looking* adjectivizations tend to occur more in spoken texts, whereas both the modifier and the predicative *ADJ-looking* adjectivizations are more prevalent in written texts. Both the modifier *ADJ looking* adjectivizations and the modifier *ADJ-looking* adjectivizations prefer to occur more in written texts, whereas both the predicative uses are more prevalent in spoken texts. The reason is that noun modifiers increase the complexity of nominal groups and hence contribute to the lexical density of written texts (Halliday 1989).

4.2 Register distribution of *ADJ-looking* adjectivizations

Of the 56,138 occurrences of the three types of constructions in the written sub-corpus, there are 31,889, 14,241, 7,920, and 2,088 occurrences in the Fiction, Magazine, Newspaper, and Academic sub-corpora of the COCA respectively. In this section, we compared the register distribution of the *ADJ-looking* adjectivizations. To facilitate comparison, we converted the normalized frequencies of the *ADJ looking* adjectivizations (6.95) and the *ADJ-looking* adjectivizations (68.02) to that of the *look ADJ* constructions (385.37) in written texts of the COCA (see Figure 6).

Figure 6: Register distributions of the *ADJ-looking* adjectivizations.

Figure 6 shows that all three types of constructions tend to occur mostly in non-academic texts, especially in fiction texts, but tend least to occur in academic texts. This is in agreement with our hypothesis that the *ADJ-looking* adjectivizations are more prevalent in fiction texts with regard to register distribution. The reason is that the *ADJ looking* adjectivizations and the *ADJ-looking* adjectivizations are evaluative and the speaker’s involvement is not a characteristic feature of academic writing (He and Guo 2021). According to Halliday (2004), academic language is characteristic of objectivity and impersonality.

The *ADJ looking* and the *ADJ-looking* adjectivizations are similarly distributed in the fiction and the academic registers ($\chi^2 = 0.061$, $df = 1$, $p = 0.805 > 0.05$) but the correlation strength is very weak ($F = 0.003$). However, the distribution of the *look ADJ* constructions and that of the *ADJ looking* and the *ADJ-looking* adjectivizations are significantly different in the two registers ($\chi^2 = 238.920$, $df = 1$, $p = 0.000 < 0.01$), but the correlation strength is also weak ($F = 0.084$). Comparatively, the *ADJ looking* and *ADJ-looking* adjectivizations tend less to occur in fiction texts.

It can also be seen from the data shown in Table 2 that the raw frequency of the *look ADJ* constructions accounts for 86.03% of the total frequency of the three types of constructions in fiction texts, ranking the highest among all the four written registers, 80.56, 82.27 and 73.66% in the magazine, newspaper, and academic texts respectively. According to Copeck et al. (1997), fiction is the least technical, and academic is the most technical. The reason for the highest frequency of the *look ADJ* constructions in fiction texts is that relational clauses of the attributive type are mostly evaluative, which is not a characteristic feature of academic writing.

Adjectivizations can package information to increase the semantic density in academic texts, but our data also show a lower frequency of adjectivizations in academic texts. The reason is that the *ADJ-looking* adjectivizations are all evaluative. They function as the interpersonal Epithet in nominal groups. However, it can also be seen that in the academic register the *ADJ looking* adjectivizations count the most in the three groups of data, and their frequency is much higher than that of the *look ADJ* constructions.

In the following, we compared the register distributions of the modifier and the predicative *ADJ-looking* adjectivizations. We converted the nominalized frequencies of the modifier *ADJ looking* adjectivizations (6.31), the predicative *ADJ looking* adjectivizations (0.63) and the predicative *ADJ-looking* adjectivizations (4.21) to that of the modifier *ADJ-looking* adjectivizations (63.82) (see Figure 7).

Figure 7: Register distributions of the modifier and predicative *ADJ looking* adjectivizations and *ADJ-looking* adjectivizations.

Figure 7 shows that the modifier and the predicative *ADJ looking* adjectivizations and the modifier and the predicative *ADJ-looking* adjectivizations, on the whole, tend to occur more in non-academic texts, especially in fiction texts ($\chi^2 = 84.665$, $df = 9$, $p = 0.000 < 0.01$). In fiction texts, both the predicative *ADJ looking* adjectivizations and the predicative *ADJ-looking* adjectivizations are more prevalent than the modifier uses. This distribution pattern is the opposite of that in academic texts. This is in agreement with our hypothesis. The reason is that fiction texts are the least technical and tend to represent experience as processes (Halliday 1989). The predicative use can explicitly express the speaker's judgement and tends to appear in relational clauses to construe quality. This fits into the feature of non-academic language.

One reason for the higher frequency of the modifier use than the predicative use in academic texts is that modifier adjectivizations can increase the complexity of nominal groups, which is a characteristic feature of academic texts. The other reason is that the modifier use can avoid the author's involvement and allows little negotiation between speaker and hearer in the text, hence increasing the objectivity of academic texts. The predicative use of the *ADJ-looking* adjectivizations will again increase the non-technicality of fiction texts. This explains the relatively lower frequency of the predicative use of the *ADJ-looking* adjectivizations than that of the modifier used in academic texts.

4.3 Discipline distribution of *ADJ-looking* adjectivizations

As shown in Table 2, we collected only a very small number of *ADJ-looking* adjectivizations from the academic sub-corpus of the COCA. Of the eight disciplines in the academic sub-corpus of the COCA, Science/Technology and Medicine are categorized as hard sciences and History, Education, Geography/Social Science, Law/Political Science, Humanities, and Philosophy/Religion as soft sciences. The data we retrieved from the eight disciplines of the academic sub-corpus are shown in Table 4.

For comparison, we converted the normalized total frequencies of the *ADJ looking* adjectivizations (0.97) and the *ADJ-looking* adjectivizations (6.78) to that of the *look ADJ* constructions (20.59) (see Figure 8).

Figure 8 shows that the *look ADJ* constructions are more prevalent in soft science texts than in hard science texts, but the distribution of the *look ADJ* constructions is not significantly different from that of the *ADJ-looking* and the *ADJ-looking* adjectivizations ($\chi^2 = 2.529$, $df = 1$, $p = 0.112 > 0.05$) and the correlation

Table 4: Discipline distributions of the ADJ/-looking adjectivizations.

Section	Look ADJ	ADJ/-looking		ADJ/-looking		Total
		Modifier	Predicative	Modifier	Predicative	
Soft science	History	4	0	54	6	195
	Education	2	0	19	2	175
	Geography/social science	4	1	56	4	300
	Law/political science	6	2	46	6	184
	Humanities	19	1	99	2	390
	Philosophy/religion	2	0	22	0	117
	Sub-total	37	4	296	20	1,361
Hard science	Science/technology	14	0	94	1	319
	Medicine	2	0	7	1	74
	Sub-total	16	0	101	2	393
Total	1,278	53	4	397	22	1,754

Figure 8: ADJ-looking adjectivizations across soft and hard sciences.

strength is very weak ($F = 0.038$). However, it is too early to conclude that the ADJ-looking adjectivizations are not discipline sensitive because the ADJ-looking adjectivizations have modifier and predicative uses.

As shown in Table 4, we retrieved 476 occurrences of the ADJ looking and the ADJ-looking adjectivizations from the soft science texts and the hard science texts in the COCA, including 53 modifier ADJ looking adjectivizations, 4 predicative ADJ looking adjectivizations, 397 modifier ADJ-looking adjectivizations, and 22 predicative ADJ-looking adjectivizations. For comparison, we converted the nominalized total frequencies of the modifier ADJ looking adjectivizations (0.92), the predicative ADJ looking adjectivizations (0.045) and the predicative ADJ-looking adjectivizations (0.307) to that of the modifier ADJ-looking adjectivizations (6.47) (see Figure 9).

Figure 9: Discipline distributions of the modifier and predicative ADJ looking adjectivizations and ADJ-looking adjectivizations.

Figure 9 shows that neither the modifier nor the predicative *ADJ-looking* adjectivizations tend to occur in hard science texts. Comparatively, under equal total frequencies, the predicative uses are more prevalent than the modifier uses in soft sciences texts, whereas the modifier uses are more prevalent than the predicative uses in hard science texts. This is in agreement with our hypothesis. The reason for the prevalence of modifier uses in hard science texts is that hard science text writers seek to downplay the personal role and distance from interpretations while soft science text writers value personal credibility and tend to express their conviction and emphasize their contribution (Hyland 2008). The modifier use enables speakers to package the quality of nouns without revealing who is responsible for the evaluation. Conversely, the predicative use presents a straightforward comment or judgement (Kress and Hodge 1979).

5 Discussion

The COHA-based study finds that the *look ADJ* constructions have the potential to shift to the *ADJ looking* adjectivizations and the transcategorization from the *ADJ looking* adjectivizations to the *ADJ-looking* adjectivizations increases the degree of lexicalization. The COCA-based study finds that the *ADJ-looking* adjectivizations are mode and register sensitive but not discipline sensitive and that the modifier use prefers to occur in hard science texts and the predicative use prefers to occur in soft science texts. This section is devoted to the discussion of these research findings.

5.1 On the diachronic distribution of *ADJ-looking* adjectivizations

The process of adjectivization can be considered as a process of metaphorization. Metaphorization in the Hallidayan sense (Halliday and Matthiessen 1999) can be interpreted as the semantic blending or semantic junction. The semantic categories of sequence, figure, and element are realized congruently as clause complex, clause, and group/phrase respectively at the lexicogrammatical stratum. The logico-semantic relation between the two clauses in a clause complex is construed by the functional element relator, which “forms figures into sequences” (Halliday and Matthiessen 1999: 177). The relator is realized as a conjunction group. The elements of a figure include participant, process, and circumstance which are congruently realized as nominal group, verbal group, and adverbial group or prepositional phrase. The mismatch between semantics and lexicogrammar

creates ideational grammatical metaphor. According to Halliday and Matthiessen (1999), adjectivization joins process with quality and hence is a sub-type of grammatical metaphor (He 2019b; Taverniers 2004). For example:

- (10) a. One day he came upon a *funny looking* little cap. (COHA_1921)
 b. One day he came upon a little cap. The cap *looked funny*.

The second clause in (10b) is rank-shifted to a nominal group in (10a). In the process of nominalization, the *look ADJ* construction *looked funny* is adjectivized to *funny looking*, joining process with quality, and hence creating ideational grammatical metaphor.

The process of adjectivization can also be considered as a process of lexicalization. Lexicalization is the process by which new linguistic entities become conventionalized on the level of the lexicon (Blank 2001). It involves fusion concerning alternative ways of expressing the same concept. Lexicalization is a gradual historical change which results in new lexical forms of any complexity. It should be noted that lexicalization is not a word formation process such as clipping, conversion, compounding and borrowing (Brinton and Traugott 2005).

The shift from the *ADJ looking* adjectivization to the *ADJ-looking* adjectivization is a process of lexicalization. According to Vartiainen (2016), the *ADJ looking* adjectivization has to be lexicalized because when it is modified by a degree adverb the whole construction will be grammatically ambiguous. The hyphenated *ADJ-looking* adjectivization will reduce the ambiguity. For example:

- (11) a. In all, Mr. Tripp was a *most impressive-looking* man. (COCA_FIC)
 b. He spoke for 10 minutes to the ever *more skeptical-looking* official.
 (COCA_FIC)

Our research also shows that the *ADJ-looking* adjectivizations (e.g. *good-looking*) and their comparative forms (e.g. *better-looking*) are of the same group of data. This indicates that *better looking* does not necessarily have the potential to shift to *better-looking*; rather, they are also of the same group of data. In recent years, the fully lexicalized *good-looking* adjectivization has begun to accept degrees of comparison.

If an adjective is semantically non-gradable, it always functions as modifier instead of predicative (Bauer and Huddleston 2002). According to Biber et al. (1999), being predicative and accepting degrees of comparison are two defining characteristics of adjectives. The more characteristics an adjective has, the more central it is as an adjective. The *ADJ-looking* adjectivizations are semantically gradable, so they have both the modifier use and the predicative use since their formation.

5.2 On the synchronic distribution of *ADJ-looking* adjectivizations

The *ADJ-looking* adjectivization is a resource for creating ideational grammatical metaphor, it should have a similar distribution trend as nominalization and verbalization. However, our research shows that the *ADJ-looking* adjectivizations tend to occur more in spoken texts than in written texts with regard to mode distribution, more in fiction texts than in academic texts with regard to register distribution and more in soft science texts than in hard science texts with regard to discipline distribution. This is the polar opposite of the synchronic distribution of nominalization or that of verbalization. There might be two reasons: (1) Nominalization shifts clause to group and verbalization shifts clause complex to clause. The *ADJ-looking* adjectivization functions as interpersonal Epithet in nominal groups or as Attribute in relational clauses. Both uses highlight the author involvement, and the author involvement is not a characteristic feature of academic writing. (2) The target semantic domain of the *ADJ-looking* adjectivizations is Quality which is regarded to be less stable and represents values on a single dimension (Halliday and Matthiessen 1999). Academic language tends to move towards thingness because the Thing is regarded as the most stable domain to be typologized.

The modifier use of the *ADJ-looking* adjectivizations helps increase the lexical density of clauses or the complexity of nominal groups. This is a characteristic feature of written texts (Halliday 1994). The predicative use, however, helps increase the grammatical intricacy. Seen from this perspective, nominalization, verbalization, and adjectivization are not necessarily the characteristic feature of academic language. They are the characteristic feature of academic language only when they increase the lexical density of language. This can explain why the modifier *ADJ-looking* adjectivizations are more prevalent in academic texts than in fiction texts with regard to register distribution and more prevalent in hard science texts than in soft science texts with regard to discipline distribution despite the fact that they function as interpersonal Epithet.

6 Conclusion

In this study, we investigated the diachronic and synchronic distributions of the *ADJ-looking* adjectivizations. The COHA-based diachronic research finds that the process of *ADJ-looking* adjectivization includes two phases. The first phase is the shift from the *look ADJ* construction to the modifier *ADJ looking* adjectivization. This is a process of metaphorization. The second phase is the shift from the *ADJ*

looking adjectivization to the *ADJ-looking* adjectivization. This is a process of lexicalization. During the second phase, the shift from the modifier use to the predicative use reinforces the adjective features of the *ADJ-looking* adjectivization.

The COCA-based synchronic research finds that the *ADJ-looking* adjectivizations are mode and register sensitive but not discipline sensitive and that the modifier use prefers to occur more in hard science texts and the predicative use prefers to occur more in soft science texts. The modifier *ADJ-looking* adjectivizations are more prevalent in written texts with regard to mode distribution, more prevalent in academic texts with regard to register distribution, and more prevalent in hard science texts with regard to discipline distribution. The predicative *ADJ-looking* adjectivizations, however, are oppositely distributed. The reason is that the modifier use of the *ADJ-looking* adjectivizations increases the lexical density and the predicative uses of the *ADJ-looking* adjectivizations increases the grammatical intricacy.

Research funding: This work was supported by the Philosophy and Social Science Planning Project of Guangdong Province (grant number GD20WZX01-11) and the National Social Science Fund of China (grant number 21BYY043).

References

- Banks, David. 2005. On the historical origins of nominalized process in scientific text. *English for Specific Purposes* 24(3). 347–357.
- Bauer, Laurie & Rodney Huddleston. 2002. Lexical word-formation. In Rodney Huddleston & Geoffrey K. Pullum (eds.), *The Cambridge grammar of the English language*, 1621–1721. Cambridge: Cambridge University Press.
- Biber, Douglas. 2006. *University language: A corpus-based study of spoken and written registers*. Philadelphia & Amsterdam: John Benjamins.
- Biber, Douglas & Bethany Gray. 2013. Being specific about historical change: The influence of subregister. *Journal of English Linguistics* 41(2). 104–134.
- Biber, Douglas, Stig Johansson, Geoffrey Leech, Susan Conrad & Edward Finegan. 1999. *Longman grammar of spoken and written English*. Harlow: Pearson Education.
- Blank, Andreas. 2001. Pathways of lexicalization. In Martin Haspelmath, Ekkehard König, Wulf Oesterreicher & Wolfgang Raible (eds.), *Language typology and language universals: An international handbook*, vol. 2, 1596–1608. Berlin & New York: Walter de Gruyter.
- Brinton, Laurel J. & Elizabeth Closs Traugott. 2005. *Lexicalization and language change*. Cambridge: Cambridge University Press.
- Copeck, Terry, Ken Barker, Sylvain Delisle, Stan Szpakowicz & Delannoy Jean-François. 1997. What is technical text? *Language Sciences* 19(4). 391–423.
- Galve, Ignacio Guillén. 1998. The textual interplay of grammatical metaphor on the nominalizations occurring in written medical English. *Journal of Pragmatics* 30. 363–385.

- Halliday, Michael Alexander Kirkwood. 1985. *An introduction to functional grammar*. London: Arnold.
- Halliday, Michael Alexander Kirkwood. 1989. *Spoken and written language*. Oxford: Oxford University Press.
- Halliday, Michael Alexander Kirkwood. 1994. *An introduction to functional grammar*, 2nd edn. London: Arnold.
- Halliday, Michael Alexander Kirkwood. 1998. Things and relations: Regrammaticising experience as technical knowledge. In James R. Martin & Robert Veel (eds.), *Reading science, critical and functional perspectives on discourses of science*, 185–235. London: Routledge.
- Halliday, Michael Alexander Kirkwood. 2004. Language and knowledge: The unpacking of text. In Jonathan J. Webster (ed.), *The language of science: Vol. 5 in the collected works of M. A. K. Halliday*, 24–48. London: Continuum.
- Halliday, Michael Alexander Kirkwood & Christian Matthias Ingemar Martin Matthiessen. 1999. *Construing experience through meaning: A language-based approach to cognition*. London: Continuum.
- Halliday, Michael Alexander Kirkwood & Christian Matthias Ingemar Martin Matthiessen. 2004. *An introduction to functional grammar*, 3rd edn. London: Arnold.
- Halliday, Michael Alexander Kirkwood & Christian Matthias Ingemar Martin Matthiessen. 2014. *Halliday's introduction to functional grammar*, 4th edn. London & New York: Routledge.
- He, Qingshun. 2019a. A corpus-based study of transfers in English nominal groups. *Glottology* 10(1–2). 57–84.
- He, Qingshun. 2019b. Types of transfer in ideational metaphor: A corpus-based study. *Studia Neophilologica* 91(3). 249–272.
- He, Qingshun & Man Guo. 2021. A corpus-based study of adjectivizations in English. *Lingua* 260. 1–15.
- He, Qingshun & Bingjun Yang. 2018. A corpus-based study of the correlation between text technicality and ideational metaphor in English. *Lingua* 203. 51–65.
- Hyland, Ken. 2008. Genre and academic writing in the disciplines. *Language Teaching* 4. 543–562.
- Kress, Gunther & Robert Hodge. 1979. *Language as ideology*. London: Routledge & Kegan Paul.
- Sanchez-Stockhammer, Christina. 2018. *English compounds and their spelling*. Cambridge: Cambridge University Press.
- Sun, Kun & Harald Baayen. 2021. Hyphenation as a compounding technique in English. *Language Sciences* 83. 1–20.
- Taverniers, Miriam. 2004. Grammatical metaphors in English. *Moderna Språk* 98(1). 17–26.
- Vartiainen, Turo. 2016. A constructionist approach to category change: Constraining factors in the adjectivization of participles. *Journal of English Linguistics* 44(1). 34–60.
- Williams, Edwin. 1981. Argument structure and morphology. *Linguistic Review* 1(1). 81–114.
- Yang, Bingjun. 2020. Full realization principle for the identification of ideational grammatical metaphor: Nominalization as example. *Journal of World Languages* 6(3). 161–174.