

The Style of English Abstracts on the Scientific Papers

Takafumi UENO

学術論文における英文要約の文体

上野 貴史

1. Introduction

The English abstract, which sometimes troubles the Japanese, is obligatory on most scientific and technical papers. While it basically helps the reader to understand the contents of the paper, English abstracts on Japanese papers have little possibility of being read and are written not for the Japanese, but for persons who cannot read Japanese. Yet, the guidelines for papers emphasize the need of writing the abstracts.

Many textbooks about the style of English abstracts are published, and Keogh (1994) summarizes them as follows: 1) eliminate unnecessary words, 2) however, avoid “telegraphic” style by not eliminating articles and cohesive elements, 3) use conjunctions for proper subordination, 4) write in complete sentences, 5) make sentences active, 6) use parallel structure, 7) spell out acronyms, 8) use only the most common abbreviations, 9) avoid jargon except when it is common to the field of study, 10) avoid the first person, 11) make sure the abstract makes sense without reference to the report.

Taking these restrictions into consideration, I suppose that there should be an original style in the English abstracts. In this paper, I will discuss, from the standpoints of the text linguistics and the quantitative linguistics, five points: the length of sentences, the tense, passive sentences, the subject of the first person, and the used verbs. These items are often utilized as the investigation of writer’s style and authenticity. By investigating these features in English abstracts, to elucidate the typical style of the English abstracts is the purpose of this paper.

2. Data

I gathered 120 abstracts from 3 scientific journals published in Japan (JA) and 3 scientific journals published in America (EN). The journals are:

Ja1: Journal of the Linguistic Society of Japan. 1990-1993.

Ja2: Journal of Home Economics of Japan. 1991.

Ja3: Japanese Journal of Educational Research. 1990/1992.

En1: Journal of Experimental

Psychology. 1993.

En2: Journal of Food Science. 1993.

En3: Journal of Teacher Education. 1992.

The range of length in English abstracts is about 150 to 500 words, varying from journal to journal. I arbitrarily singled out 20 abstracts from each journal and made use of 120 abstracts in total.

Note that in these abstracts I will not treat as data (a) the parenthesized words and (b) the itemized sentences:

(a) The effect of long-term feeding (252 days) of ...

(En2-5)

(b) Contents of this paper are as follows:

I Introduction

II Diffusion of New Education in the City Elementary Schools

III

(Ja3-3)

The parenthesized words in (a), *252 days*, and the itemized sentences in (b), *I Introdu-*

ction/II Diffusion of New Education ..., will be disregarded; for I discuss the style of abstracts in the sentence level. a) and b) are incomplete sentences, and they are outside the range of this paper.

3. The Style of English Abstracts

While the style of English abstracts is suggested by the textbooks which Keogh (1994) treats, the guidelines for Japanese scientific journals only prescribe the number of words and the contents¹⁾. Making a comparison between the abstracts on the English journals and those on the Japanese journals, I discuss the style of English abstracts.

3.1 The Length of Sentences

The number of words in one sentence is available in order to investigate the normal length of sentences. <Table 1>²⁾ shows the average of the sentence length,

Table 1 The Length of Sentences

	Ja1	Ja2	Ja3	JA	En1	En2	En3	EN	Avg. ³⁾
Number of Words(%)									
1- 5	1.1%	1.8%		1.1%					0.8%
6-10	10.4%	11.2%	6.4%	9.6%	8.5%	8.1%	9.3%	8.6%	9.3%
11-15	18.7%	26.5%	18.6%	22.1%	17.0%	27.3%	18.7%	20.9%	21.6%
16-20	22.0%	23.9%	19.2%	22.0%	20.3%	25.3%	22.7%	22.6%	22.2%
21-25	18.7%	15.2%	19.7%	17.5%	28.8%	17.2%	16.0%	21.6%	18.8%
26-30	12.6%	9.4%	15.4%	12.1%	11.9%	14.1%	12.0%	12.7%	12.3%
31-35	7.7%	7.2%	8.5%	7.7%	8.5%	7.1%	12.0%	8.9%	8.1%
36-40	4.4%	2.9%	3.7%	3.6%	5.1%	1.0%	4.0%	3.4%	3.5%
41-45	2.8%	0.7%	3.7%	2.2%			1.3%	0.3%	1.6%
46-50	1.1%		2.7%	1.1%					0.8%
51-55	0.6%	0.7%	0.5%	0.6%			2.7%	0.7%	0.6%
56-60			1.1%	0.3%					0.2%
61-65							1.3%	0.3%	0.1%
65-70		0.4%		0.2%			1.3%	0.3%	0.2%
71-75			1.1%	0.2%					0.1%
Avg. (Words)	21.3	19.4	24.3	21.2	21.1	18.7	23.1	20.8	21.1
Median (Words)	20	18	22	20	21	18	21	20	20

the maximum frequency (mode), and the median. The average of the sentence length is 21.1 words and the JA is not differentiated from the EN. The average stands for the measure of central tendency, but the dispersion must depend on the frequency. The maximum frequency spells that the writer is the most favorite sentence length and the reader feels easier to understand. The 22.1% of the JA is from 11 to 15 words, while the 22.6% of the EN is 16 to 20 words. Though the maximum frequency presents that the EN prefers long sentences to the JA, sentences more than 60% in JA and EN have 11 to 25 words, and it is found that this range is the normal sentence length of English-abstracts. The median of the sentence length indicates not-long and not short sentences. The result is that the median of the sentence length in both JA and EN is 20 words. From these quantitative investigation, we understand that the normal sentence length in English abstracts is between 11 and 25 words by the fact that sentences more than 60% have 11 to 25 words. Moreover, we can regard around 20 words as the best sentence length; for the average and the median of the sentence length indicates about 20 words.

3.2 Tense

The tense-form is the recurrent element, and demonstrates many features in the limited texts like an abstract. Weinrich (1964) classified the tense into the Tense I and the Tense II on the theory of the text linguistics. The Tense I is interpreted as "review tense" and is mainly used in editorial articles, scientific reports, philosophic essays, etc. The Tense I group in English is represented as the present

tense, the future tense, the present perfect tense, and the future perfect tense. On the contrary, the Tense II is "narrative tense", appearing in historical descriptions, articles on the political conference, etc. The past tense and the past perfect tense belong to Tense II group. Following this classification, I investigated the tense used in abstracts (<<Table 2>>). It indicates that the tense in the abstracts has no fixed tense group. The decision of using narrative or review seems to depend on the field of study and the convention of scientific journals. This can be inferred from the deviation of tense used in scientific journals. For example, the review tense is practiced in many abstracts of Ja1 and En3, and the narrative is used in Ja2, Ja3, En1, and En2.

Table 2 The Tense Group

	Rate of Tense Group		Number of Abstracts	
	Tense I	Tense II	Tense I	Tense II
Ja1	84.1%	15.9%	18 abstracts	1 abstract
Ja2	35.6%	64.4%	7 abstracts	13 abstracts
Ja3	38.6%	61.4%	5 abstracts	15 abstracts
JA	53.7%	46.3%	30 abstracts	29 abstracts
En1	38.4%	61.6%	5 abstracts	15 abstracts
En2	22.4%	77.6%	4 abstracts	14 abstracts
En3	71.9%	28.1%	15 abstracts	4 abstracts
EN	43.7%	56.3%	24 abstracts	32 abstracts
Total	50.5%	49.5%	54 abstracts	61 abstracts

Secondly, I discuss the transition of tense. The tense transition can be divided into two types: homogeneous and heterogeneous. While the homogeneous transition indicates that a tense group shifts to the same tense group (the present tense → the present perfect tense, the past tense → the past perfect tense, etc.), the heterogeneous transition means that a tense group shifts to the different tense group (the present tense → the past perfect tense, the past tense → the present tense). The homo-

Table 3-a The Tense Transition of Ja1

	Present	Futer	Pre. Perfect ^o	Past	Past Perfect
Present	64.9%	2.2%	3.4%	5.9%	0.3%
Future	1.9%		0.6%	0.3%	
Pre. Perfect	1.9%		0.6%	1.9%	
Past	7.1%		0.6%	8.1%	0.3%
Past Perfect				0.6%	

Table 3-b The Tense Transition of Ja2

	Present	Futer	Pre. Perfect	Past	Past Perfect
Present	23.6%		0.9%	8.7%	
Future					
Pre. Perfect	1.3%		0.4%	0.9%	
Past	8.3%		0.4%	55.5%	
Past Perfect					

Table 3-c The Tense Transition of Ja3

	Present	Futer	Pre. Perfect	Past	Past Perfect
Present	24.4%	0.8%	0.8%	7.3%	
Future	0.5%	0.8%		0.3%	
Pre. Perfect	1.9%	0.3%	0.3%	1.3%	
Past	8.1%		1.6%	52.0%	0.3%
Past Perfect				0.3%	

Table 3-d The Tense Transition of En1

	Present	Futer	Pre. Perfect	Past	Past Perfect
Present	22.5%			8.4%	0.5%
Future					
Pre. Perfect	0.5%				
Past	16.8%			51.3%	
Past Perfect				0.5%	

Table 3-e The Tense Transition of En2

	Present	Futer	Pre. Perfect	Past	Past Perfect
Present	3.8%			1.0%	
Future					
Pre. Perfect				1.0%	
Past	6.7%			85.7%	1.0%
Past Perfect				0.5%	

Table 3-f The Tense Transition of En3

	Present	Futer	Pre. Perfect	Past	Past Perfect
Present	52.5%		4.2%	7.6%	
Future				0.8%	
Pre. Perfect	0.8%		0.8%		
Past	13.6%			13.6%	1.7%
Past Perfect				1.7%	

geneous transition presents the consistency in the text(textuality). When the textuality increases, the amounts of information decrease. Thus the heterogeneous transition produces the evocation to reader. <Table 3> shows the rate of tense transition from vertical tenses to horizontal tenses. The ratios of the homogeneous transition to the heterogeneous transition are 84.5% to 15.5%(Ja1), 81.7% to 18.3%(Ja2), 81.5% to 18.5%(Ja3), 74.8% to 25.2%(En1), 91.5% to 8.5%(En2), and 75.3% to 24.7%(En3). The homogeneous transition occupies more than 80%, and this number presents the high textuality of English abstracts.

3.3 The Passive Sentence

There is the description, "make sentence active", in the textbooks which Keogh (1994) treats. This means that passive sentences should be avoided when we write abstracts. But passive sentences, in fact, are often used in English abstracts. The ratios of passive voices to the verb number are 31.7%(Ja1), 37.2%(Ja2), 24.9% (Ja3), 32.7%(En1), 38.4%(En2), and 21.5%(En3). It is shown that about 1/3 verbs in English abstracts are used as passive voice. Although this number should be compared with other genres, it would be regarded as the high ratio.

3.4 The Subject of the First Person

I investigate the number of the first person appearing in the subject position (<Table 4>). In English abstracts the use of the first person is only 3.1% to the number of verbs. In particular the single first person is fewer(1.0%). It is frequently found that the plural first person is used in abstracts written by one author.

Table 4 The Use of the First Person

	I	We	First Person
Ja1	2.0%	3.2%	5.2%
Ja2	0 %	1.6%	1.6%
Ja3	1.2%	1.7%	3.0%
JA	1.2%	2.2%	3.4%
En1	0 %	2.8%	2.8%
En2	0 %	0 %	0 %
En3	1.5%	3.0%	4.4%
EN	0.4%	2.1%	2.6%
Total	1.0%	2.2%	3.1%

3.5 Verbs

Investigating high-frequency verbs, we note the difference of the used vocabulary between the JA and the EN. I can point out three facts: a) while the verbs which mean "affect" or "influence" are much used in the EN, they aren't used in the JA. b) while the EN has many kinds of verbs ("show", "indicate", "reflect", "demonstrate", "present", etc.) in order to express the meaning, "to show something", the JA only uses "show". c) There are verbs of "regard" and "consider" which mean "to think of something in a stated way" in the high-frequency verbs of the JA. In <Appendix 1 and 2> the lists are indicated.

4. Conclusion

This paper examined the style of English abstracts from the view points of the quantitative and text linguistics. The results are as follows; a) the average of the sentence length is 21 words, the maximum frequency is 11 to 25 words, and the medians 20 words. About 20 words in a sentence are the normal sentence length in English abstracts. b) The decision of the tense group depends on the field of study and the convention of the journals. c) The ratio of the homogeneous transition

to the heterogeneous transition is 81.5% to 18.5%. d) The ratio of the passive voice to the active voice is 30.5% to 59.5%. e) The use of the first person in the subject position is 3.1% to the total verbs. f) The verbs of "affect" and "influence" are the high-frequency verbs in the EN. g) Many kinds of synonyms which mean "show something" are used in the EN. h) The verbs of "regard" and "consider" are the high-frequency verbs in the JA. This could be inferred from the fact that these expressions much exist in Japanese (*kangaerareru*, *Omowareru*, etc.).

<Notes>

- 1) The guidelines of Ja2 provide as follows:
 - a) English abstracts should be briefly written by using simple sentences in order to understand the contents of paper, including the purpose, the method, the result, and the consideration.
 - b) The number of words should be about 150 words.
- 2) The hyphenated words and the numerical formulas are counted as one word.
- 3) Avg. is the abbreviation for "average".
- 4) Following Weinrich(1964), I don't count the infinitive, the imperative mode, the participle and the gerund.
- 5) Pre. Perfect is the abbreviation for "present perfect".

<Reference>

- de Beaugrande, R. & W.U. Dressler. 1981. *Introduction to Text linguistics*. Longman.
- Hatano, K. 1966. *Gendai Bunsyo Shinrigaku*. Dainippon Tosho.
- Keogh, T. 1994. *The Structure of Abstracts: Stylistic and Structural Elements in 48 Scientific and Technical Abstracts*. Doctoral Dissertation. Louisiana State University.

Peterfalvi, J. 1970. *Introduction à la Psycholinguistique*. Presses Universitaires de France.

Weinrich, H. 1964. *Tempus*. Besprochene und erzählte Welt.

<Appendix 1>
The List of High-frequency Verbs in JA

*The list is produced by extracting verbs of more than 3 appearances from 988 words in total.

Word	Frequency	Word	Frequency
be	192	add	4
have	30	attempt	4
show	30	bring	4
use	20	change	4
examine	12	concern	4
decrease	11	differ	4
find	11	express	4
increase	11	function	4
study	11	propose	4
include	10	reduce	4
become	9	represent	4
develop	9	result	4
base	8	see	4
describe	8	seem	4
explain	8	serve	4
investigate	8	support	4
regard	8	teach	4
suggest	8	try	4
appear	7	account	3
come	7	allow	3
conclude	7	analyze	3
consider	7	apply	3
consist	7	argue	3
discuss	7	assume	3
divide	7	believe	3
make	7	belong	3
observe	7	carry	3
obtain	7	clarify	3
take	7	correspond	3
aim	6	follow	3
know	6	insist	3
lead	6	publish	3
mean	6	realize	3
point	6	recognize	3
present	6	remain	3
summarize	6	require	3
classify	5	undertake	3
contain	5		
give	5		
introduce	5		
lack	5		
produce	5		
provide	5		
reveal	5		
say	5		
stand	5		

<Appendix 2>
The List of High-frequency Verbs in EN

*The list is produced by extracting verbs of more than 3 appearances from 468 words in total.

Word	Frequency	Word	Frequency
be	48	appear	4
have	16	demonstrate	4
increase	13	examine	4
suggest	13	give	4
show	12	occur	4
use	11	reflect	4
provide	10	report	4
affect	9	support	4
describe	9	tend	4
discuss	9	conduct	3
investigate	8	convert	3
indicate	7	decrease	3
reduce	7	detect	3
make	6	evaluate	3
obtain	6	gain	3
produce	6	identify	3
find	5	influence	3
focus	5	learn	3
include	5	mediate	3
observe	5	present	3
require	5	preserve	3
seek	5	remain	3
		result	3
		study	3
		test	3