

KrKwic

\*

Daum.net

<sup>1</sup>Loet Leydesdorff<sup>2</sup>

FullText

KrKwic

KrKwic

KrKwic

Daum.net

: FullText, KrKwic,

1.

(Krippendorff, 1980).

가

Danowski(1993)

가

(category)

(external)

가 ( , 2001).

(

, <http://academic.csuohio.edu/kueuendorf/content/cpuc/>

\* 2003

(KRP-2003-042-H00003).

<sup>1</sup>712-749

214-1

. E-mail : hanpark@yutmail.ac.kr

<sup>2</sup>

, E-mail : loet@leydesdorff.net

ccap.htm).

가  
 LIWC(Linguistic Inquiry and Word Count) K  
 (Korean)-LIWC ( , 2004).  
 ( , 2004). ( )  
 가  
 가

FullText  
 KrKwic (FullText  
 (<http://users.fmg.uva.nl/leydesdorff/software/fulltext>). KrKwic  
 가  
 가 KrKwic  
 가 가 , <http://www.banpark.net/KrKwic.zip>.

2. KrKwic

“ ”  
 (Wasserman & Faust, 1994).  
 가  
 가 (symbol)  
 (cognitive) (Leydesdorff, 2003).  
 가  
 (prominent) (semantic association)  
 “ ” “  
 “ ”  
 가, 가  
 (Park, 2001).  
 KrKwic 가  
 가 (frame)

가 ( , 2001). 가

### 3. KrKwic

KrKwic KrKwic, KrTitle, KrText 3

#### 3.1 KrKwic(Korean Key Words In Context)

KrKwic

KrKwic 가

(word frequency analysis) (default) ASCII

( ANSI ), 1,000

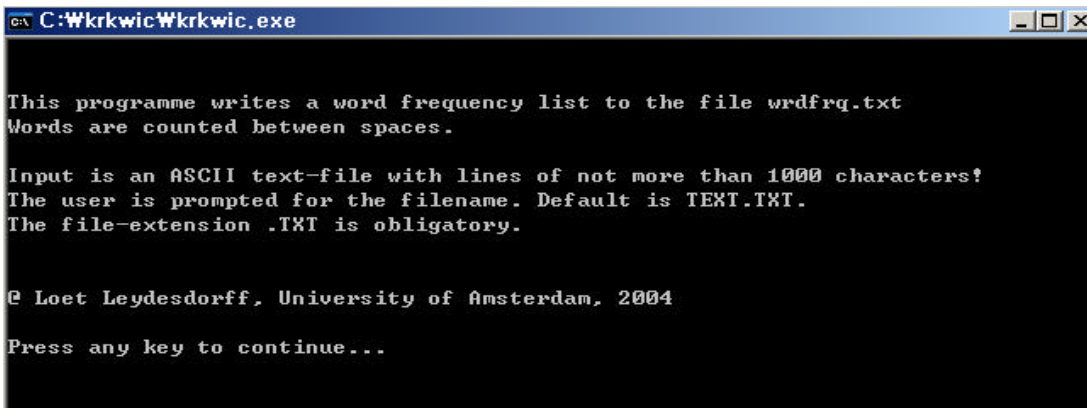
(default) text.txt

wrdfrq.txt wrdfrq.dbf 가

(salient)

I 가

가



## 3.2 KrTitle(Korean Title)

```

KrTitle
(mission statement),
KrTitle
ASCII 1
A 999
, A 가 999 가
(default) text.txt
가 ASCII
가 1024
(default) words.txt , KrKwic

matrix.dbf: (document) × (words) 2 가
(cases) × (variables)

```

```

C:\Wkrkwic\Wkrtitle.exe
© Loet Leydesdorff, 2004; <http://www.leydesdorff.net>

The maximum number of different words is 1024.
The words have to be unique.
This version also works with Korean characters.

The number of documents is only limited by the disk-size.
This version reads from a text-file text.txt into ti.dbf
Each title should be on a separate line.
The maximum length of lines is 999.

The programme creates a file MATRIX.dbf, containing a
matrix in which the documents are the cases and
the words the variables. COOCC.dbf contains the co-word
matrix and COSINE.dbf the normalized co-occurrence matrix.
Additionally, the files COOCC.DAT and COSINE.DAT are written
in the so-called DL format for Ucinet and Pajek.

The programme may overwrite files in the same directory!
Therefore, run the programme in a special directory.

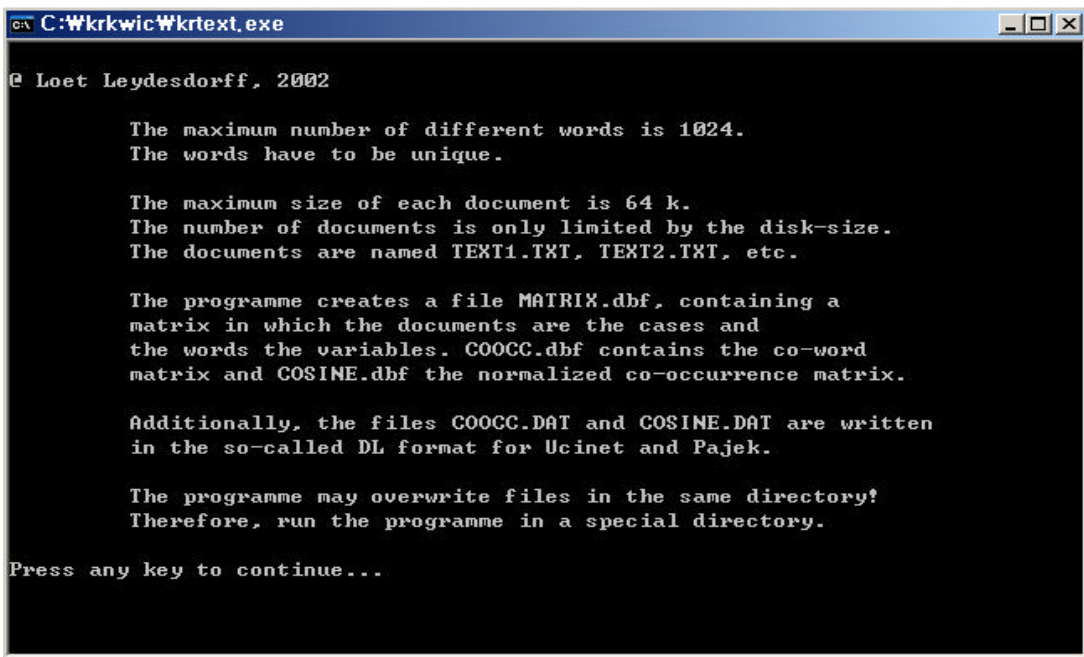
Press any key to continue...

```

coocc.dat coocc.dbf: (words) × (words) I (co-occurrence)  
 (symmetric) (cell)  
 cosine.dat cosine.dbf: (words) × (words) I (cosine) (symmetric)  
 가 .dat UciNet(  
 가 가 : http://www.analytictech.com/download/adc6.htm) Pajek(  
 가 가 : http://mrvar.fdv.uni-lj.si/sola/info4/programme.htm) (import)  
 (mapping) 가  
 가 .dbf Excel SPSS 가  
 ( , 2004).

### 3.3 KrText(Korean Text)

KrText KrTitle  
 A 64k  
 ASCII text1.txt, text2.txt, text3.txt 가  
 KrTitle



< 3> KrText

```

Krkwic
가
copy text.txt + text*.txt
text.txt

```

4.

```

"
Krkwic
"
"
, Daum.net
1,310
가
10
( 4 ).
2004 08 03 10 45

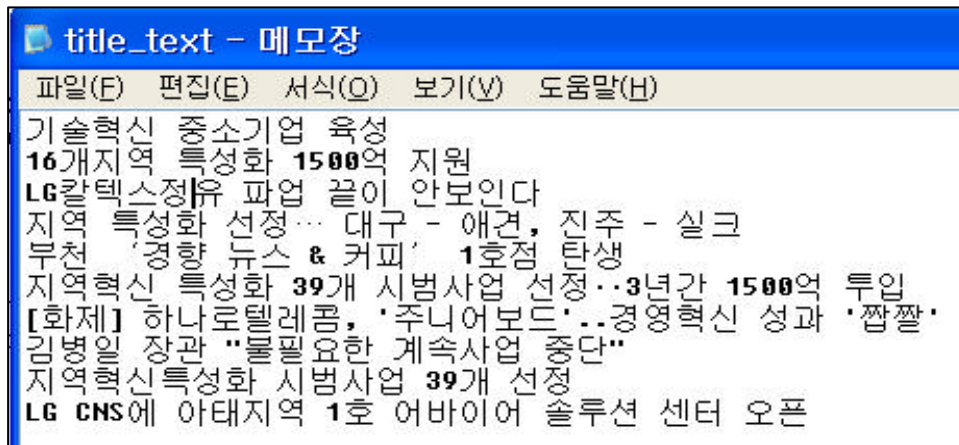
```



```

5> title_text.txt
가
가
" " " " " "
Krkwic
(cleaning)

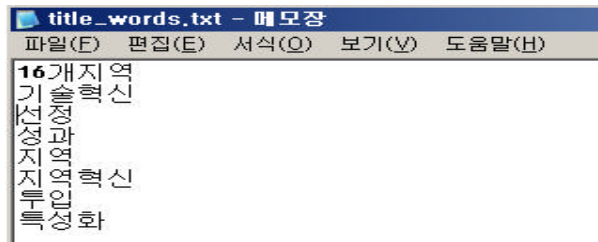
```



< 5> title\_text.txt

title\_text.txt KrKwic

title\_words.txt ( 6 ) 8



< 6> title\_words.txt

title\_text.txt title-words.txt KrTide 8

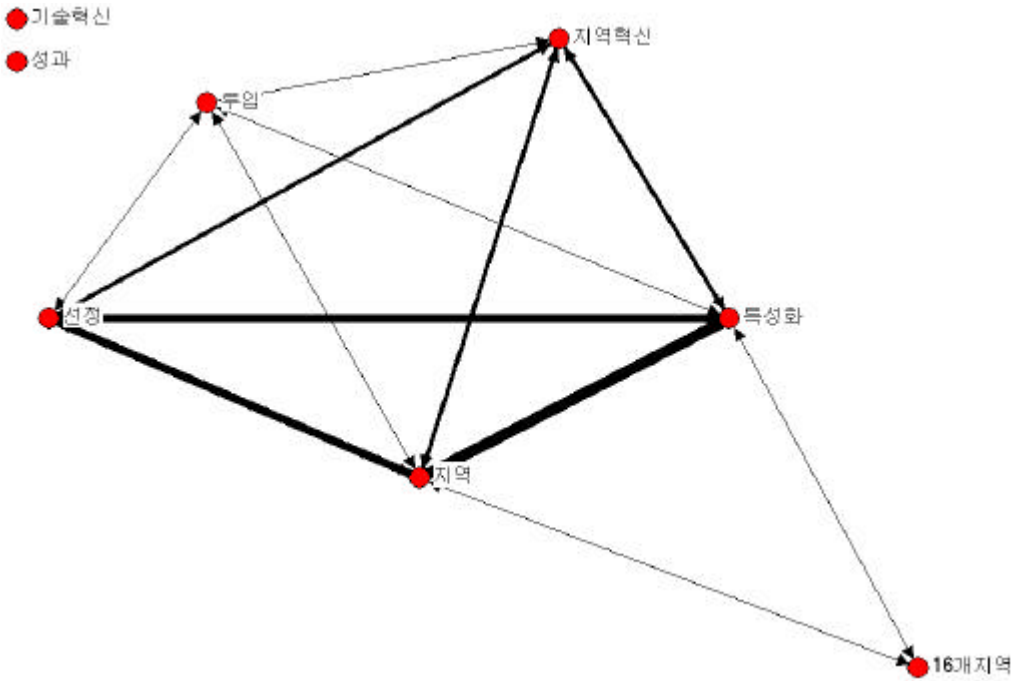
( 1 )

x x

< 1> x

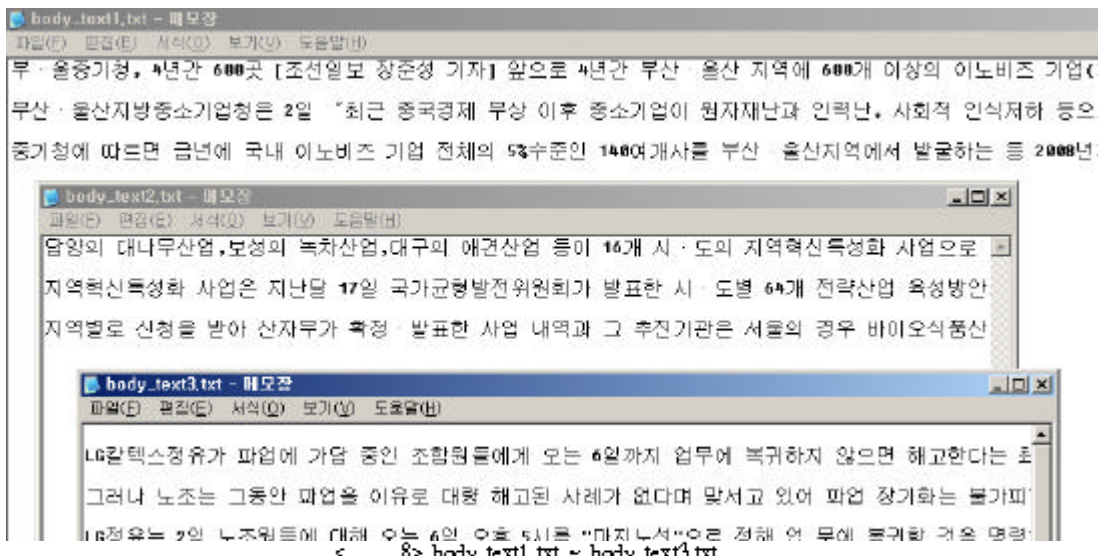
	16							
16		0	0	0	1	0	0	1
	0		0	0	0	0	0	0
	0	0		0	3	2	1	3
	0	0	0		0	0	0	0
	1	0	3	0		2	1	4
	0	0	2	0	2		1	2
	0	0	1	0	1	1		1
	1	0	3	0	4	2	1	

Ucinet (centrality),  
 (density), (clustering analysis), (multi-dimensional scaling)  
 ( : , 2004; , 2004).  
 , Ucinet NetDraw  
 < 7> 가 (line)  
 가 가



< 7>  
 “ - ” “ - ” 4 가 “ ” “ ”  
 “ - ” “ - ” 3 가 “ ” “ ”  
 가  
 (performance-oriented)  
 (horse-race journalism)  
 < 4>  
 I  
 ( 8 )  
 10 , 10 txt





5.

KrKwic

KrKwic

- [1] 김민준 (2004). *Journal of the Korean Data Analysis Society*, Vol.6, No. 1, 255-265.
- [2] 김민준 (2001). *Journal of the Korean Data Analysis Society*. Riffe, D., Lacy, S., Fico, F.G. (1998). *Analyzing media messages*, Lawrence Erlbaum Associates.
- [3] 김민준 (2004). *Journal of the Korean Data Analysis Society*, Vol. 6, No. 3, 901-915.
- [4] 김민준 (2004). *Journal of the Korean Data Analysis Society*, Vol. 6, No. 3, 887-900.
- [5] 김민준 (2004). *(K-LIWC)*, 2004

- [6] (2001). , Vol. 16, 108-133.
- [7] (2004). , *Journal of the Korean Data Analysis Society*, Vol. 6, No. 3, 861-872.
- [8] Danowski, J. (1993). Network analysis of message content, In: G. Barnett and W. Richards, (editors). *Progress in Communication Sciences*, Norwood, N.J.: Ablex, pp. 197-222.
- [9] Krippendorff, K. (1980). Content analysis: an introduction to its methodology, *The Sage Comtext Series*, Vol 5, London: Sage.
- [10] Leydesdorff, L. (2003). *A sociological theory of communication: The self-organization of the knowledge-based society*, Universal Publishers.
- [11] Park, H. W. (2001). Content Analysis of Community Newspaper on Educational Computing: A Critical Perspective, *Proceedings of the Working Boundaries Conference (Science, Technology, & Society Conference)*, Rensselaer Polytechnic Institute, Troy, NY.
- [12] Wasserman, S. and Faust, K. (1994). *Social network analysis: Methods and applications*, Cambridge, NY: Cambridge University Press.

[ 2004 8 , 2004 10 ]

## Understanding the KrKwic: A computer program for the analysis of Korean text<sup>\*</sup>

*Han Woo Park<sup>1</sup>, Loet Leydesdorff<sup>2</sup>*

### Abstract

This paper examines the KrKwic (Korean Key Word In Context) which is a computer-based content analysis software. The KrKwic is a social network computer program for analyzing a Korean text. The software is able to count how many times a word occur in the text and produces a list of words in the order of high frequency. The most frequently occurring words are defined as key words and a co-occurrence (or cosine) matrix among those words are produced. The matrix can be used as an input file to social network analysis softwares such as Ucinet or Pajek. for a wide variety of analytic techniques (e.g., centrality, density, cluster analysis, multi-dimensional scaling, network diagram). This paper illustrates how to use the KrKwic using the 10 news stories regarding "regional innovation." The sampled articles were collected at <http://www.daum.net> which is the most comprehensive search engine in S. Korea.

*Keywords* : KrKwic, computer-based content analysis software, Korean texts, social network analysis.

---

<sup>\*</sup>This research was partly supported by the Korea Research Foundation grants in 2003 (KRF-2003-042-H00003).

<sup>1</sup>Full-time lecturer, School of Social Science, YeungNam University, KyeongSan, South Korea.  
E-mail : parkhanwoo@hotmail.com

<sup>2</sup>Associate Professor, Amsterdam School of Communication Research, University of Amsterdam, Amsterdam, The Netherlands. E-mail : loet@leydesdorff.net