



**AntConc (Windows, Macintosh OSX and Linux)**

Build 3.2.4

Laurence Anthony, Ph.D.

Center for English Language Education in Science  
and Engineering, School of Science and  
Engineering, Waseda University, 3-4-1 Okubo,  
Shinjuku-ku, Tokyo 169-8555, Japan

May 1, 2011

#####

**안트콘크 3.2.4 (윈도우즈, 매킨토시 OSX, 리눅스) 한국어 리드미**

정채관 박사 BEng(Hons) Bham MSc Warwick EdD Warwick Cert Oxford

주소: 서울시 서대문구 연세로 50 연세대학교 위당관 306-1

영어코퍼스연구실(우: 120-749)

전화: 0221237516

팩스: 023622381

이메일: corpuslab@yonsei.ac.kr

2011년 10월 7일

본 한국어 리드미는 지적재산권 법의 보호를 받는 지적재산입니다. 본  
한국어 리드미는 상업적 용도로 사용될 수 없으며, 이를 연구목적 혹은  
수업에 활용하고자 하는 분은 정채관 박사에게 사전에 허락을 받고  
사용하기 바랍니다.

#####

## 서론(Introduction)

안트콘크(AntConc)는 코퍼스 언어학 연구와 데이터를 이용한 학습에 유용한 프리웨어이자 멀티 플랫폼 도구이다. 안트콘크는 윈도우 98/ME/200/NT, XP, Vista, Win 7 등과 같은 마이크로 소프트 윈도우 기반과 OS X 10.4.x, 10.5.x, 10.6.x 등과 같은 매킨토시, 우분투(Ubuntu) 10과 같은 리눅스 환경에서 작동된다. 안트콘크는 이와 같이 다양한 컴퓨터 운영체제에서 작동되기 위해 액티브스태이트에서 제공하는 펄어플(PerlApp) 컴파일러를 사용해서 개발됐다.

## 설치(Installation)

### 윈도우(Windows)

윈도우 시스템에서는 단순히 안트콘크 아이콘을 두 번 누르면 프로그램이 시작한다. 별 다른 설치과정이 필요 없다.

### 매킨토시 OS X(Macintosh OS X)

매킨토시 환경에서는 우선 X11을 설치한 뒤 X11을 실행시킨다. X11은 그래픽적인 도구로 컴퓨터나 애플 웹 사이트에 있는 것을 사용할 수 있다. X11을 작동 시킨 후 안트콘크 아이콘을 더블 클릭하면 프로그램이 시작된다. 이것 이외의 별다른 설치과정은 필요 없다.

### 리눅스(Linux)

리눅스 환경에서는 시스템적인 문제 때문에 안트콘크가 실행파일로 사용될 수 있도록 사용자가 허락 해줘야 한다. 이 과정을 마친 후 안트콘크 아이콘을 더블 클릭하면 프로그램이 바로 실행된다. 이외의 별다른 설치는 필요 없다.

### 안트콘크 기능에 대한 간략한 소개(Overview of Tools)

안트콘크는 실행 후 메인 화면에서 볼 수 있는 탭을 클릭하거나, F1에서 F7까지 미리 지정된 기능키를 눌러 각각의 기능을 실행시킬 수 있다.

### 콘코던스(Concordance Tool)

검색결과를 문맥속의 키워드(KeyWord in Context, KWIC) 형식으로 볼 수 있다. 이 기능을 사용하면 단어와 구가 일반적으로 어떻게 코퍼스 안에서 사용되는지 볼 수 있다.

### 콘코던스 플롯(Concordance Plot Tool)

검색결과를 바코드 형식으로 보여주는 기능이다. 이 기능을 사용하면 검색어가 타겟 텍스트의 어느 지점에 나타나는지 알 수 있다.

### 파일 뷰(File View Tool)

개별적인 텍스트 파일을 보여주는 기능이다. 이 기능을 사용하면 안트콘크에서 제공하는 다른 기능들에 비해 보다 자세한 조사를 할 수 있다.

### **클러스터(Clusters, N-Grams)**

사용자가 설정한 조건에 따라 검색된 클러스터들을 보여준다. 이 기능은 실질적으로 콘코던스 기능이나 콘코던스 플롯 기능을 이용해서 생성된 결과를 요약하는 역할을 한다. 반면 N-그램은 타겟 코퍼스 전체에서 'N'(예를 들어, 1개 단어, 2개 단어 등)과 같이 클러스터 길이에 대한 검색을 한다. 이 같은 기능은 어떤 코퍼스에서 일반적인 표현을 찾는 데 유용하다.

### **콜로케이트(Collocates)**

사용자가 지정한 검색어의 콜로케이트를 보여준다. 이 기능을 이용해서 언어의 비순차적 패턴도 조사할 수도 있다.

### **워드 리스트(Word List)**

타겟 코퍼스에 있는 모든 단어를 세고, 순서대로 보여주는 기능이다. 이 기능을 사용하면 코퍼스에서 가장 많이 사용된 단어를 빨리 찾을 수 있다.

### **키워드 리스트(Keyword List)**

타겟 코퍼스와 레퍼런스 코퍼스를 비교해서, 타겟 코퍼스에서 눈에 띄게 빈도수가 높거나 또는 그와 반대로 현저하게 빈도수가 낮은 단어를 찾을 수 있다. 이 기능을 이용하면 타겟 코퍼스에서 볼 수 있는 특유의 단어들을 밝혀낼 수 있다. 이는 장르나 특수목적영어(English for Specific Purposes, ESP) 연구를 하는 사용자에게 특히 유용한 기능이다.

## 콘코던스(Concordance Tool)

이 기능은 문맥 속에서의 키워드(KeyWord In Context, KWIC) 형식으로 검색결과를 보여준다. 이 도구를 사용해서 어떤 코퍼스에서 단어나 구가 어떤 식으로 사용되었는지를 조사할 수 있다.

콘코던스 라인을 만들기 위해서는 아래와 같은 과정을 따라 하면 된다.

- 1) 파일 메뉴에서 '파일 열기(Open File(s)...)'나 '디렉터리 열기(Open Dir...)' 옵션 중 하나를 택한 뒤, 타겟 텍스트 파일을 불러온다. 참고로 이렇게 불러온 파일은 안트콘크 메인 화면 좌측 상단에 나타난다.
- 2) 좌측 하단에 있는 검색창에 본인이 원하는 검색어를 입력한다.
- 3) 우측에 있는 '윈도우 크기 검색(Search Window Size)'에서 검색어를 중심으로 양쪽에 몇 개의 문자가 나오기를 원하는지 선택한다(참고로 내장 값은 50이다).
- 4) '시작(Start)' 버튼을 클릭하면 콘코던스 라인이 만들어진다. '멈춤(Stop)' 버튼을 누르면 콘코던스 라인이 만들어지는 과정을 중단 시킬 수 있다(참고로 이 기능은 나중에 나올 다른 기능에서도 계속 반복된다).
- 5) 오른쪽 버튼 방에 있는 버튼을 사용해서 콘코던스 라인의 단어를 재정렬할 수 있다. 이때 0은 검색어이고, 1L, 2L...은 검색어의 왼쪽, 1R, 2R...은 검색어의 오른쪽을 의미한다. 안트콘크에서는 이와 같이 3레벨별로 설정을 할 수 있는데, 처음 콘코던스 라인이 처음 만들어진 후 바로 2번째와 3번째 레벨이 적용되지 않을 때가 있다. 이럴 경우에는 한번 더 시도해본다.
- 6) 결과 정렬에 대한 세팅이 끝났으면, '정렬(Sort)' 버튼을 누른다.
- 7) 콘코던스 라인 중 하이라이트가 된 단어 위에 컴퓨터 마우스 커서를 올리면 마우스 커서의 모양이 화살표에서 손 모양으로 바뀌는 것을 볼 수 있다. 이때 그 검색어를 클릭하면, 화면이 원본 텍스트로 이동한다. 이 과정에 대한 자세한 설명은 나중에 나오는 **\*\*파일 뷰(File View)\*\*** 부분을 참고하기 바란다.

총 콘코던스 라인의 수는 안트콘크 화면 중앙에 있는 '콘코던스 히트(Concordance Hits)'에서 확인 할 수 있다. 참고로 콘코던스 라인을 만드는 과정에서는 모든 처리과정이 끝날 때까지 숫자가 계속 올라간다.

만일 검색어에 해당하는 콘코던스 라인이 만들어지지 못할 경우에는 '노 히트(NO HITS)'라는 메시지가 뜨게 되고, 이럴 경우에는 콘코던스 라인은 업데이트 되지 못한다. 이때는 이전의 콘코던스 라인이 그대로 보이게 된다.

검색어는 초기 세팅으로 '단어(Word)'가 지정되어 있지만, 사용자가 원하면 옵션에서 '단어 조각(Word Fragment)'로 변경할 수 있다. 또한 콘코던스 검색을 할 때 대문자와 소문자를 구별해서 할 수 있고, 그렇게 하지 않을 수도 있다. 참고로 안트콘크는 초기 세팅에서는 대·소문자를 구별하지 않도록 되어있다. 한편 안트콘크에서 '레직스(Regex)' 옵션을 선택하면 정규 표현(Regular Expression)을 이용한 검색을 할 수도 있다.

정규 표현에 대한 보다 자세한 내용은 '정규 표현 마스터하기(Mastering Regular Expressions)' - O'Reilly & Associates Press나 인터넷에서 "regular expressions"를 검색해보기 바란다.

<http://www.regular-expressions.info/quickstart.html>과 같은 웹 페이지도 직접 방문해 보자.

안트콘크는 일반검색과 별도로 고급검색 기능도 갖고 있다. 고급검색은 '고급검색(Advanced Search)' 버튼을 누른 후 옵션을 지정하면 사용할 수 있는데, 첫 번째 고급검색 옵션은 검색어 그룹을 설정하는 것이다. 이때 한 개의 검색어를 지정할 수도 있고, 본인이 원하면 새로운 검색어를 추가로 지정할 수도 있다. 새롭게 검색어들을 추가할 때는 한 줄에 한 개씩만 입력해야 안트콘크가 인식을 할 수 있으니 주의하자. 검색어가 여러 개 있을 경우 이런 여러 개의 검색어가 저장되어 있는 파일을 불러 올 수도 있다. 안트콘크에서 이런 기능을 추가로 제공하는 이유는 만일 검색어가 많을 경우 사용자가 여러 번 다시 입력하는 수고를 덜 수 있기 때문이다. 두 번째 고급검색 옵션은 문맥에 등장하는 문맥어를 지정하는 옵션이다. 예를 들어 'student'라는 단어의 좌측이나 우측 3 번째 안에 'university'가 함께 나오는 경우를 찾고 싶다면, 'student'를 검색어, 'university'를 문맥어로 놓고, 문맥창의 범위를 3L부터 3R로 지정하면 된다.

이 기능과 관련해서 메뉴 옵션에서 추가적인 옵션 지정이 가능하다.

## 콘코던스 플롯(Concordance Plot Tool)

이 기능은 검색결과를 바코드 형식으로 보여준다. 이 기능을 사용하면 검색어가 타겟 텍스트의 어느 지점에 나타나는지 알 수 있다.

콘코던스 플롯은 기본적으로 위에서 설명한 **\*\*콘코던스(Concordance)\*\***를 이용해서 콘코던스 라인을 추출하는 과정과 거의 비슷하다. 하지만 **\*\*콘코던스 플롯(Concordance Plot)\*\***은 콘코던스 라인을 또 다른 시각에서 보여준다. 다시 말해서 **\*\*콘코던스 플롯(Concordance Plot)\*\***은 콘코던스를 통해 검색된 단어가 '바코드'에 나타나는 줄무늬 모양으로 보인다. 이 줄무늬(검색어)의 위치를 보면 한 눈에 검색어가 타겟 텍스트의 어느 부분에 오는지를 시각적으로 볼 수 있다. 콘코던스 플롯 기능을 사용해서 검색어가 어떤 클러스터와 함께 나타나는지도 확인할 수 있다. 이 기능은 기술문서에서 특정 문맥어가 어디에 나타나는지 확인하거나, 소설이나 희곡에서 특정 인물이 언제 등장하는지 알고 싶을 때 꽤 유용하게 사용될 수 있다.

한편 콘코던스 플롯 오른쪽에는 검색어가 검색된 숫자와 길이가 나타난다. 안트콘크 메인 화면 좌측 하단에 있는 확대·축소 버튼을 이용하면 사용자가 원하는 크기로 화면을 크게 하거나 작게 할 수도 있다.

이전 **\*\*콘코던스(Concordance)\*\***에서와 마찬가지로, 만일 마우스 커서를 하이라이트 된 줄무늬 위로 옮기면 마우스 커서가 작은 손 모양으로 바뀐다. 이때 그 특정 줄무늬를 클릭하면 **\*\*파일 뷰(File View)\*\***로 화면이 바뀌고, 텍스트 원문에서 해당 검색어를 바로 찾아 볼 수 있다.

## 파일 뷰(File View Tool)

원문 텍스트를 보여주는 기능이다. 이 기능을 사용하면 안트콘크에서 제공하는 다른 기능들에 비해 보다 세부적인 조사를 할 수 있다.

원문 텍스트를 보고 싶다면, 아래와 같은 과정을 따라한다.

- 1) 안트콘크 화면 왼쪽에 있는 텍스트 파일 리스트 중 하나를 선택한다.
- 2) **\*\*콘코던스(Concordance)\*\***를 이용해서 콘코던스 라인을 만든 뒤 특정 검색어를 클릭하면 텍스트 원문 전체가 화면에 나온다. 이때 해당 검색어가 하이라이트 되어있는 것을 볼 수 있는데, 검색옵션은 **\*\*콘코던스(Concordance)\*\***이나 **\*\*콘코던스 플롯(Concordance Plot)\*\***이나 비슷하다.
- 3) '히트 로케이션(Hit Location)' 버튼을 사용하면 파일 원문에서 검색어가 등장하는 곳으로 점프를 할 수 있다.
- 4) 원문에서 다른 검색어를 보고 싶다면, **\*\*콘코던스(Concordance)\*\***로 돌아와서 새로운 검색어를 입력한 뒤 '시작(Start)' 버튼을 누른다.
- 5) 참고로 **\*\*파일 뷰(File View)\*\*** 화면에서 하이라이트가 되어있는 검색어를 클릭하면, 이것이 새로운 검색어로 인식되어 새로운 KWIC 라인이 만들어진다.

참고로 아래는 **\*\*파일 뷰(File View)\*\***에서 사용할 수 있는 단축키(Shortcuts)이다.

- CTRL-Click = 화면에서 가장 가까운 곳으로 점프



## 클러스터(Clusters Tool)

사용자가 설정한 조건에 따라 검색된 클러스터들을 보여주는 기능이다. 이 기능은 실질적으로 콘코던스나 콘코던스 플롯 도구를 이용해서 생성된 결과를 요약한다.

클러스터는 사용자의 선택에 따라 그 결과를 빈도수, 검색어의 첫 번째, 또는 맨 마지막에 나오는 문자의 알파벳 순서 등에 따라 나열시킬 수 있다. 또한 클러스터를 구성하고 있는 단어들 중 첫 번째 단어와 바로 뒤 따라 나오는 단어와의 확률관계에 따라 결과를 나열시킬 수도 있다. 물론 사용자가 원하면 이런 결과를 반대 순서로 나열 시킬 수도 있다. 이밖에 사용자는 각 클러스터에서 본인이 원하는 최소길이와 최대길이를 정할 수 있고, 클러스터의 최소 빈도수를 정해서 그 범위 안에서의 클러스터만 따로 볼 수도 있다. 이와 별도로 검색어가 항상 클러스터의 좌측이나 우측에 나오도록 지정할 수도 있다. 참고로 현재 버전에서 검색어가 한 개 이상일 경우, '검색어 오른쪽(Search Term on Right)' 옵션을 선택했을 경우 첫 번째 단어가 항상 오른쪽에 나온다.

클러스터 리스트를 만들기 위해서는 아래 과정을 순차적으로 따라한다.

- 1) 클러스터 결과의 정렬 옵션을 정한다.
- 2) '시작(Start)' 버튼을 누른다(다른 기능들과 마찬가지로 클러스터 리스트를 만드는 과정에서 중간에 '멈춤(Stop)' 버튼을 누르면 이 과정은 중단된다).
- 3) 클러스터 리스트가 만들어진 후 특정 클러스터를 클릭하면 그와 관련된 KWIC 라인이 만들어진다.

이전에 소개한 다른 기능들과 마찬가지로 클러스터도 다양한 선택사항이 있다. 이에 대한 더 자세한 내용은 뒤에 나오는 메뉴 선택을 참고하기 바란다.

## N-그램(N-Grams Tool - part of the Cluster Tool)

이 기능은 말 그대로 사용자가 선택한 타겟 텍스트 파일에서 N-그램을 추출하기 위해 사용된다. N-그램은 워드 N-그램이기 때문에 텍스트 파일이 크면 많은 수의 N-그램이 만들어진다. 예를 들어 N-그램 사이즈를 2라고 했을 때, 'this is a pen'이라는 문장에서 'this is', 'is a', 'a pen'이라는 3개의 N-그램이 만들어진다.

앞서 논의한 **\*\*클러스터(Clusters Tools)\*\***와 마찬가지로, N-그램의 결과 역시 빈도수의 높고 낮은 순서나 단어를 구성하는 문자 중 첫 번째 나오거나 제일 마지막에 나오는 문자의 알파벳 순서대로 나열시킬 수 있다. 반복되는 얘기지만 사용자가 옵션을 어떻게 지정하느냐에 따라 N-그램의 첫 번째 나오는 단어와 그 뒤에 나오는 단어들과의 확률을 바탕으로 결과를 나열시킬 수 있고, 그 반대로 나타나게 할 수도 있다. 또한 각 N-그램에 대한 최소값이나 최대값을 지정하거나, 최소 빈도수를 사용자가 원하는 대로 사전에 지정할 수도 있다.

N-그램 리스트를 만들기 위해 사용자가 해야 할 일은 아래와 같다.

- 1) 검색창 위에 있는 'N-그램(N-Grams)' 옵션을 선택한다.
- 2) 결과 정렬 옵션을 정한다.
- 3) '시작(Start)' 버튼을 누른다. 이때 앞서 다른 기능들을 설명했을 때와 마찬가지로 '멈춤(Stop)' 버튼을 누르면 언제든지 이 과정을 멈출 수 있다.
- 4) 특정 '렉시컬 번들(Lexical Bundle)'을 누르면 이와 관련된 KWIC 라인이 만들어진다.

다른 도구들과 마찬가지로 N-그램과 관련해서 사용자가 원하는 대로 다양한 옵션을 지정할 수 있다. 더 자세한 내용은 나중에 나오는 메뉴 선택사항을 참고하기 바란다.

## 콜로케이트(Collocates Tool)

이 기능은 검색어의 콜로케이트를 보여준다. 또한 이 기능을 사용해서 언어의 비순차적 패턴을 조사할 수도 있다.

안트콘크는 사용자가 원하는 대로 콜로케이트의 빈도수나 검색어 좌측과 우측에 오는 단어의 빈도수, 또 검색어의 첫 번째 나오거나 제일 마지막에 나오는 문자의 알파벳 순서로 결과를 나열 시킬 수 있다. 또한 검색어와 콜로케이트의 통계적 측정값에 따라 결과를 나열시킬 수 있다. 이 값은 검색어와 콜로케이트가 얼마나 관련이 있는지에 대한 측정값이다. 현재 사용되는 측정방법은 아래와 같다. 물론 이 모든 결과를 반대순서로 정렬시킬 수도 있다. 또한 이와 별도로 사용자는 검색어를 중심으로 콜로케이트 좌측과 우측에서 볼 수 있는 단어의 범위를 정하거나 정렬 될 최소 빈도수를 지정할 수도 있다. 만일 검색어에서 바로 옆에 오는 단어만을 볼 생각이려면 '같음(Same)' 박스를 선택하면 최소범위와 최대범위가 같게 되고, 결과적으로 왼쪽과 오른쪽에서 볼 수 있는 단어들의 범위 수가 동일하게 된다.

안트콘크에서 콜로케이트를 계산하는 통계적 방식은 아래와 같다.

- 상호 정보, (MI)Mutual Information: 이 방식은 M. Stubbs가 1995년 Functions of Language에 출판한 'Collocations and Semantic Profiles'이라는 논문에서 사용한 계산방식이다. [M. Stubbs, Collocations and Semantic Profiles, Functions of Language 2, 1 (1995)]
- T-스코어, (T-Score)T-Score: 이 역시 Stubbs가 같은 논문에서 소개한 계산식이다. [M. Stubbs, Collocations and Semantic Profiles, Functions of Language 2, 1 (1995)]

콜로케이트 리스트를 만들기 위해서는 아래와 같은 과정을 따라한다.

- 1) 결과 정렬 방식을 정한다.
- 2) '시작(Start)' 버튼을 누른다. 다른 기능들과 마찬가지로 '멈춤(Stop)' 버튼을 누르면 중간에 이 프로세스를 멈출 수 있다.

3) 콜로케이트 리스트가 만들어진 후 특정 콜로케이트를 클릭하면 그에 대한 KWIC 라인이 만들어 진다.

뒷부분에 나오는 메뉴 선택을 보면 콜로케이트에 대한 다양한 옵션이 있으니 참고하기 바란다.

## 워드 리스트(Word List Tool)

이 기능은 타겟 코퍼스에 있는 모든 단어를 셀 뿐 아니라 이를 순서대로 보여준다. 이 기능을 사용하면 코퍼스에서 어떤 단어가 가장 자주 사용되는지를 빨리 알 수 있다.

안트콘크를 사용하면 단어를 빈도수나 단어의 시작 문자, 또는 맨 마지막 문자에 따라 보이게 할 수 있고, 정반대로 안 보이게 할 수도 있다. 또 사용자가 원하면 단어를 구성하는 문자들의 대문자와 소문자를 구분할 수도 있고, 구분하지 않을 수도 있다. 초기 설정은 대·소문자를 구분하지 않도록 되어있다.

워드 리스트를 만들기 위해서는 아래 과정을 따라 한다.

- 1) 결과를 어떻게 정렬할지 정한다.
- 2) '시작(Start)' 버튼을 누른다. 이전과 마찬가지로 '멈춤(Stop)' 버튼을 누르면 언제든지 이 과정을 멈출 수 있다.
- 3) 워드 리스트가 만들어진 후 특정 워드를 클릭하면 그에 대한 KWIC 리스트가 만들어진다.

워드 리스트와 관련해서도 메뉴 선택사항에서 다양한 옵션을 지정할 수 있다. 뒤에 나오는 메뉴 선택을 참조하기 바란다.

## 키워드 리스트(Keyword List)

이 기능을 사용하면 타겟 코퍼스와 레퍼런스 코퍼스를 비교해서 타겟 코퍼스에서 현저하게 빈도수가 높거나 낮은 단어를 찾을 수 있다. 한편 이 기능을 이용해서 타겟 코퍼스를 대변하는 특징적인 단어들도 조사할 수 있는데, 이는 장르나 특수목적영어(English for Specific Purposes, ESP) 연구에 유용하다.

안트콘크를 사용해서 키워드 리스트를 만들기 위해 해야 할 일은 아래와 같다.

- 1) 타겟 텍스트 파일을 선택한다.
- 2) '선택사항(Preferences)' 메뉴에서 '키워드 선택사항(Keyword Preferences)' 옵션을 선택한다.
- 3) '키니스(Keyness)'에 대한 통계적 계산방식을 정한다. 참고로 안트콘크는 초기 세팅은 로그-라이클리후드(Log-Likelyhood)로 되어있다.
- 4) 몇 개의 키워드가 디스플레이 되길 원하는지 정한다.
- 5) '부정적인 키워드(Negative Keywords)'(레퍼런스 코퍼스와 비교해서 타겟 텍스트에서 눈에 띄게 빈도수가 낮은 키워드)도 결과에서 디스플레이 할지 정한다.
- 6) 레퍼런스 코퍼스 옵션을 정한다. 'Use raw file(s)'를 선택하면 레퍼런스 코퍼스로 사용될 텍스트 파일들을 선택하는 것이다. 반면에 'Use word list(s)'를 선택하면 이미 만들어진 레퍼런스 코퍼스의 워드리스트를 이용할 수 있다. 참고로 이때 사용되는 워드 리스트의 형식은 아래와 같다.

RANK FREQUENCY WORD (스페이스나 탭으로 단어 사이를 분리한다)

예)     1       the       12838  
          2       a        11289  
          3       of        8583

- 7) 타겟 텍스트를 불러왔던 방식으로 레퍼런스 코퍼스의 텍스트(.txt) 파일도 불러온다.
- 8) 레퍼런스 코퍼스 불러오기를 마쳤으면, 불러온 파일들이 '키워드 선택사항(Keyword Preferences)' 옵션 메뉴에 나타난다.
- 9) '키워드 선택사항(Keyword Preference)' 메뉴에서 '확인(OK)'를 누른 뒤

다시 원래 키워드 화면으로 돌아온다.

- 10) 워드 리스트 결과를 정렬했던 비슷한 방식으로 본인이 원하는 결과 정렬 옵션을 선택한다.
- 11) '시작(Start)' 버튼을 누른다. 이때도 이전 워드 리스트를 만들 때처럼 '멈춤(Stop)' 버튼을 누르면 언제든지 이 과정을 중단 시킬 수 있다.
- 12) 생성된 키워드 리스트에서 본인이 흥미 있어 하는 키워드를 클릭하면 자동으로 KWIC 리스트가 만들어진다.

키워드 리스트와 관련된 다른 옵션들은 뒤에 나오는 메뉴 선택을 참조하기 바란다.

## 메뉴 옵션(MENU OPTIONS)

메뉴옵션은 '파일(File)', '글로벌 세팅(Global Settings)'과 '도구 선택사항(Tool Preferences)' 3가지로 구성되어 있다. 각 옵션에 대한 설명은 아래와 같다.

### <파일(FILE)>

'파일(File)'과 관련된 옵션은 주로 안트콘크로 파일을 불러오거나, 안트콘크를 이용해서 만들어진 다양한 형태의 데이터를 하드 디스크로 보내는 것과 관련이 있다. 파일 옵션에는 현재 세팅을 파일로 보내거나 파일로 저장되어 있는 사용자가 정한 세팅을 불러오는 기능도 있다. 만일 사용자가 안트콘크를 사용하다가 어떤 이유에서인지 프로그램에 문제가 생기면 프로그램을 다시 시작하거나, '기본세팅 복원(Restore Default Settings)' 옵션을 사용해서 안트콘크 초기 세팅으로 돌아가 프로그램을 다시 시작하는 방법도 있다.

### <글로벌 세팅(GLOBAL SETTINGS)>

이 옵션에서 제공하는 카테고리는 안트콘크의 다양한 기능에 영향을 끼친다.

#### <파일 세팅(File Settings)>

파일 세팅에서 사용자는 파일의 이름만 보이게 할지, 아니면 파일의 경로를 모두 표시할지 선택할 수 있다. 또한 이 옵션에서 사용자는 타겟 텍스트 파일에 들어있는 태그를 보이게 할지 숨길지를 선택할 수 있고, 태그에 관한 범위도 정할 수 있다.

#### <태그 세팅(Tag Settings)>

안트콘크 사용자는 태그 세팅 카테고리에서 코퍼스 파일에 담겨있는 태그를 보이게 할지, 아니면 그렇게 하지 않을지를 선택할 수 있다. 이때 만일 사용자가 태그를 숨기면 태그 표시의 시작과 끝을 명시해야 하는데, 초기세팅으로는 '< >'로 되어있다.

#### <와일드카드 세팅(Wildcard Settings)>

사용자는 와일드카드 세팅에서 초기에 설정되어 있는 와일드카드 문자나 기호를 변경할 수 있다. 예를 들어 'or'로 설정되어 있는



기호는 '|' (파이프 기호)이다. 하지만 키보드에 따라 이 기호가 없을 수도 있는데, 이럴 때는 사용자가 이 옵션에서 이 기호대신 '/'로 변경시킬 수 있다.

### <토큰(단어) 정의(Token(Word) Definition)>

이 옵션에서 사용자는 어떤 문자나 숫자, 또는 그 밖의 것들을 '단어(Word)'로 취급할 것인지 선택할 수 있다. 예를 들어 어떤 사용자는 글자만 단어를 구성하는 요소로 생각할 수 있지만, 어떤 사용자는 숫자나 대시(Dash) 같은 기호도 단어를 구성하는 범주에 포함시키고 싶을 수도 있다. 안트콘크는 유니코드가 호환될 뿐 아니라 유럽어와 아시아어 등 다양한 형태의 언어를 대부분 처리할 수 있다. 이런 이유 때문에 안트콘크의 초기세팅은 아래 웹 페이지 리스트에서 언급되는 입력코드를 사용한 문자나 기호(예를 들면 일본어 문자 등), 숫자도 광범위하게 포함한다. 한편 이와 별도로 사용자가 원한다면 본인이 원하는 새로운 '단어(Word)'도 정의할 수도 있다.

유니코드(Unicode) 기준에 대한 더 자세한 내용은 아래 웹 페이지들을 참고하기 바란다.

<http://www.cs.tut.fi/~jkorpela/unicode/guide.html>  
<http://www.cs.tut.fi/~jkorpela/unicode/guide.html>  
<http://www.unicode.org/>  
<http://www.unicode.org/Public/5.0.0/ucd/UCD.html>  
<http://www.unicode.org/Public/UNIDATA/PropList.txt>  
<http://www.unicode.org/charts/>

### <컬러 세팅(Color Settings)>

안트콘크 사용자는 결과나 다른 정보를 어떤 색으로 보이게 할지 컬러 세팅 카테고리에서 지정할 수 있다.

### <폰트 세팅(Font Settings)>

안트콘크 사용자는 폰트 세팅 카테고리에서 폰트 타입, 사이즈, 스타일 등 결과나 그 밖의 정보를 어떻게 표시할지 고를 수 있다.

### <언어 인코딩(Language Encodings)>

안트콘크는 유니코드와 호환되기 때문에 다양한 유럽어뿐 아니라

수많은 아시아어까지 처리할 수 있다. 따라서 다르게 인코딩이 되어있는 언어는 이 세팅에서 해당 인코딩에 맞게 지정할 필요가 있다. 참고로 사용자가 서양 언어를 사용하는 경우 대부분 ISO-8859-1로 코딩이 되어있다. 안트콘크는 이 코딩방식을 초기 설정으로 잡고 있다. 일본어 텍스트의 경우는 시프트지스(Shiftjis)로 코딩되어있는 경우가 많다. 사용자가 안트콘크를 사용하기 전에 사용할 텍스트가 어떤 코딩을 사용했는지 확인하고 이를 사전에 지정하면 안트콘크가 제대로 작동한다.

### <기능 선택사항(TOOL PREFERENCES)>

**\*\*콘코던스 플롯(Concordance Plot)\*\***과 **\*\*파일 뷰(File View)\*\***를 제외하고, 각 기능은 선택사항 카테고리에서 보다 정밀하게 세팅되어 질 수 있다. 모든 기능 선택사항 카테고리는 사용자가 결과를 어떤 프레임에서 특정 요소를 보이게 하거나 또는 감출 수도 있다. 예를 들어 사용자가 원하면 **\*\*콘코던스(Concordance)\*\*** 결과를 보여주는 화면에서 파일 이름을 보여주는 섹션을 감출 수 있다. 또 안트콘크는 데이터 처리과정에서 대·소문자를 구분하지 않게 할 수 있고, 결과를 정렬할 때 문자의 구분을 둘 수도 있다. 만일 결과를 대문자와 소문자를 구별하는 방식으로 나타내면, 대문자를 포함한 단어들이 리스트의 상위를 차지하게 된다.

#### <콘코던스 선택사항(Concordance Preferences)>

이 선택사항에서는 결과를 나열할 때 검색어를 기준으로 왼쪽과 오른쪽에 나오는 단어에 따라 정렬하는 방식 이외에 검색어의 첫 문자나 제일 마지막 문자를 기준으로 결과가 나열되도록 해줄 수 있다. 이런 방식을 통해 사용자는 스펠링이 다른 것을 찾을 수 있을 것이다. 또한 이 선택사항에서는 KWIC 라인에서 검색어를 감출 수 있는데, 이를 이용해서 교사는 수업시간에 학생들에게 괄호 넣기와 같은 퀴즈를 바로 할 수도 있다.

#### <클러스터 선택사항(Clusters Preferences)>

안트콘크에서 제공하는 클러스터 선택사항은 이미 이전에 충분히 설명했기 때문에 이 섹션에서는 생략하기로 한다.

#### <콜로케이트 선택사항(Collocates Preferences)>

콜로케이트 기능 관련, 이 선택사항에서 사용자는 텍스트를 분석하는 통계적 측정방법에 대한 방식을 지정할 수 있다. 참고로 현재 버전에서는 상호정보(Mutual Information(MI))와 T-스코어(T-Score) 방식을 지원한다.

#### <워드 리스트 선택사항(Word List Preferences)>

사용자는 이 선택사항에서 파일로 되어있는 리마 리스트(Lemma List)를 불러 올 수 있다. 또한 기존의 리마 리스트를 사용하면 타겟 텍스트에서 워드 리스트 대신 리마 리스트를 추출해낼 수 있다.

따라서 이 과정을 거치면 결과화면에는 리마 워드 폼(Lemma Word Form(s)) 칼럼이 생기고, 그 밑에 각 단어에 대한 리마와 빈도수가 보이게 된다.

한편 사용자가 원하면 '리마 엔트리->단어'의 형식을 사용해서 안트콘크 사용자가 직접 리마를 리마 리스트에 더하거나 리스트 자체를 새로 만들 수도 있다. 예를 들어 이와 같은 형식으로 리마 리스트를 만들면 아래와 같다.

be->is, are

play->play, plays, playing, played

참고로 위와 같은 예는 콤마와 스페이스(Space)는 토큰의 범주에 들어가지 않는다는 것을 전제 했다. 만일 안트콘크 웹 사이트에서 다운받은 리마 리스트를 사용하는 사용자가 있다면, 이때 대시(Dash)가 토큰(Word) 정의 범주에 포함되어 있다는 점을 알고 있을 필요가 있다. 왜냐하면 그 리마 리스트에는 하이픈으로 연결된 단어도 나오기 때문이다.

워드 리스트는 타겟 텍스트에 나오는 단어가 모두 추출 되도록 할 수도 있고, 특정 그룹의 단어만 나오게 할 수도 있다. 또 그와 반대로 스톱 리스트(Stop List)를 만들어 어떤 특정 단어 그룹만 나오지 않게 지정할 수도 있다. 이 같은 방식을 워드 리스트 레인지(Wordlist Range)라고 부르는데, 사용자가 직접 입력한 단어나 파일로 저장된 단어들, 또는 이를 혼합한 방식을 섞어서 특정 단어만 나오게 하거나, 혹은 그 반대로도 지정할 수 있는 것이다.

### <키워드 리스트 선택사항(Keyword List Preferences)>

키워드 리스트와 관련해서 앞서 **\*\*키워드 리스트(Keyword List)\*\***

섹션에서 언급한 것 이외에 몇 가지 추가적으로 선택할 수 있는 사항들이 있다. 안트콘크에서 키워드 리스트를 만들기 위해서는 레퍼런스 코퍼스를 지정해야 한다. 본격적으로 키워드 리스트를 만드는 과정에서 타겟 텍스트의 레퍼런스 코퍼스의

'키니스(Keyness)'를 통계적인 방법으로 측정한다. 이때 안트콘크의

초기 세팅에는 '키니스'와 임계값이 지정되어 있다. 이 값들을 그대로 쓸 수도 있지만, 사용자가 원하면 이 선택사항 메뉴에서 이 값들을 변경할 수도 있다.

한편 이때 '부정적인 키워드 보기(Show Negative Keywords)' 옵션을 선택하면, 레퍼런스 코퍼스와 비교해서 타겟 텍스트에서 눈에 띄게 빈도수가 낮은 단어들도 디스플레이 되게 된다. 물론 이때도 사용자가 원하면 옵션을 반대로 지정하는 것도 가능하다.

## <단축키(SHORTCUTS)>

아래는 안드로이드에서 사용할 수 있는 몇 가지 단축키 모음이다.

- CTRL-C= 현재 선택된 텍스트를 복사
- CTRL-A= 화면에 있는 모든 텍스트를 선택
- ALT-A= 보이는 모든 화면에 있는 모든 텍스트 선택
  
- 두 번 클릭= 현재 단어 선택
- 세 번 클릭= 화면의 현재 라인 선택
- SHIFT-클릭= 화면에 보이는 모든 라인을 연속으로 선택
- CTRL-클릭= 화면에 보이는 모든 라인을 선별적으로 선택
- DELETE= 화면에 보이는 라인 중 선택한 라인을 지움
- INSERT= 화면에서 선택한 라인을 제외한 다른 것들을 지움

참고로 검색 창에서 '위(UP)', 또는 '아래(DOWN)' 화살표 버튼을 이용해서 이전에 입력했던 단어들을 살펴볼 수 있다.

## 노트(NOTES)

### <결과 저장하기(Saving Results)>

안트콘크를 이용해서 작업한 내용을 저장하는 방법은 여러 가지가 있다. 우선 텍스트 파일(.txt)이나 파일 메뉴에서 '화면 저장(Save Window)'를 선택해서 저장할 수 있고, 또 안트콘크 실행파일을 다시 클릭해서 또 하나의 안트콘크를 시작해서 여러 개의 결과를 다르게 저장할 수도 있다.

### <코멘트/제안/고친 점(Comments/Suggestions/Bug Fixes)>

과거 이력을 보면 그동안 모든 에디션에 대한 내용과 수정된 버그들에 대한 기록이 나온다. 그럼에도 불구하고 만일 이 프로그램을 사용하다가 버그를 발견한다거나, 또는 개선을 위한 제안을 알려주면 다음 버전에 반영을 하도록 노력하겠다. 따지고 보면 그동안 발견된 버그나 업데이트 된 내용들도 대부분 세계 곳곳에서 안트콘크를 사용하는 사용자들이 알려준 것에 기인한다. 그 점에 대해 무척 감사하게 생각한다.

이후 법적고지에서 다시 언급하겠지만, 이 프로그램은 무료로 사용할 수 있다. 하지만 이 프로그램을 개발할 때 지원 받은 단체에, 이 프로그램이 사람들에게 정말 유용하게 사용되고 있다는 소식을 전할 필요가 있기 때문에 사용소감을 이메일로 보내주면 고맙겠다.

### <감사의 글(ACKNOWLEDGEMENTS)>

그동안 안트콘크의 기능향상을 위해 많은 사용자들이 다양한 제안을 해줬다. 우선 그들에게 감사하다는 말을 전하고 싶다. 참고로 안트콘크 개발은 일본의 교육, 과학, 체육, 문화의 촉진을 위한 모임(No. 16700573)과 와세다 대학교 특별연구과제(No. 2004B-861)의 지원 아래 이뤄졌다.

### <법적 고지(LLEGAL MATTER)>

안트콘크3.1.302는 상업적 연구목적은 제외한 개인적인 용도로 자유롭게 사용될 수 있고, 개발자의 리드미 파일이 변형되지 않은 조건에서 자유롭게 배포될 수 있다. 만일 안트콘크가 단체로 사용될 계획이라면 개발자에게

사전에 허락을 받아야 한다. 안트콘크는 현재 모습과 기능 그대로 제공되며, 이 소프트웨어를 사용함으로써 어떤 결과를 초래할 것인지에 대한 법적 책임이 없다.

## <기타 알려진 문제들(KNOWN ISSUES)>

아래는 그동안 안트콘크 사용자들이 전해온 약간 특이한 문제들이다. 본인에게 해당되는 사항도 있고, 그렇지 않은 경우도 있으니 선별적으로 참고만 하자.

### 1. 파일 열기(File Opening)

'파일열기(Open File(s))'를 클릭해서 파일 불러오기를 할 때, 리눅스 운영 시스템에서 한 번에 불러 올 수 있는 수가 950개로 정해져 있다는 보고를 받았다. 만일 이 숫자를 넘어가는 파일을 한 번에 불러올 경우 하나도 읽혀지지 않는 문제가 발생한다. 이 문제를 피하는 방법은 간단하다. 만일 많은 수의 파일을 불러와야 할 경우에는 '디렉터리 열기(Open Directory)' 옵션을 선택하면 된다. 이 옵션에서는 불러올 수 있는 파일 수가 제한이 없기 때문이다. 윈도우 기반에서는 간혹 알 수 없는 대화상자가 만들어지는 경우도 있는데, 이 부분은 개발자도 이해가 잘 가지 않는 부분이다.

### 2. 스크롤바(Scrollbar)

많은 수의 콘코던스 라인(또는 워드나 키워드)가 만들어질 경우 스크롤바가 잘 작동이 안 된다는 보고를 받은 적이 있다. 특히 밑에 있는 것들을 보려 할 때, 마우스 커서를 움직이지 않으면 제일 밑에 있는 부분을 볼 수 없다는 얘기도 들었다. 상당히 화가 나는 부분이지만, 나도 이 점은 어쩔 수 없는 버그라고 할 수 밖에 없다.

### 3. 워드 클러스터(Word Clusters)

**\*\*워드 클러스터(Word Clusters)\*\*** 기능을 사용할 때 만일 2개 이상의 검색어를 동시에 입력하면, '검색어 오른쪽(Search Term on Right)' 옵션이 선택된 경우에는 첫 번째 검색어만 나타난다.



#### 4. 태그 감추기 옵션(Hide Tags Option)

이상하게 들릴지 모르지만, '태그 감추기 옵션(Hide Tags Option)'이나 다른 기능을 사용하기 전에 **\*\*파일 뷰(File View)\*\*** 기능을 쓸 때 안트콘크가 갑자기 작동하는 것을 멈춘다는 보고를 받았다. 이 문제는 개발자도 잘 이해가 안 가는 부분인데, 조사 결과 일본어 버전 윈도우에서 파일을 바탕화면에 저장했거나 파일 이름이 영어가 아닌 경우에 자주 나타나는 것 같았다. 따라서 가급적 파일이 담겨 있는 디렉터리를 영어로 해놓고, 파일 이름 역시 영어로 해 놓는 것이 이런 문제를 막을 수 있는 대안이라고 생각한다.

#### 5. 중국어 문자(Chinese Characters)

앞서 언급한 대로 영어가 아닌 언어에서 주로 여러 가지 문제점들이 발생한다는 보고를 받았다. 많은 사용자들이 중국어 글자를 제대로 보기 어렵다고 하는 얘기도 들었다. 이는 중국어에도 여러 가지 다른 종류의 중국어가 있고, 이에 동작하는 운영시스템 또한 다를 수밖에 없다. 그렇기 때문에 현재로서는 이 문제를 해결하기 위한 최선의 방법은 운영시스템에 대한 적합한 폰트 입력을 해 본 뒤 안트콘크를 실행시키는 방법뿐이다.

#### 6. 워드 앤드(Word End)

워드 리스트를 만든 후 '워드 앤드(Word End)' 옵션을 선택하면 간혹 안트콘크가 작동불능 상태에 빠진다는 보고가 있으니 참고하기 바란다.

#### 7. 안트콘크와 매킨토시(AntConc on Macintosh OS X)

매킨토시 OS X에서 안트콘크를 사용할 때 영어가 아닌 언어 중 폰트가 제대로 나타나지 않는 문제가 발생한다는 리포트를 본 적이 있다. 이것은 안트콘크 자체의 문제라기보다는 X11의 문제로 보인다. 그렇기 때문에 최신 X11 버전을 설치한 후에 안트콘크를 사용하길 권한다.

#### 8. 마이크로소프트 윈도우 버그

간혹 윈도우 시스템에서 만일 코퍼스 파일이 데스크 탑에 있는 경우 파일을 두 번째 열 때 파일 이름이 쓰인 팝업 창이 뜬 뒤 안트콘크가 갑자기 사라지는 증상이 나타난다. 이는 안트콘크의 버그가 아니라 마이크로소프트 윈도우의 버그이다. 이와 같은 증상은 마이크로소프트 노트북을 사용할

때도 일어난다. 이 문제를 피하기 위해서는 가급적 파일을 데스크 탑에 저장하지 않는 것이 좋다. 또 팝업 창이 뜨기 전에 빨리 파일을 선택하는 것도 이 문제를 피해가는 방법이기도 하다.

- 끝(The END) -