多言語検索ツール AntPConc の機能と使い方

(Ver. 1)

O AntPConc とは

*AntPConc*は、コンコーダンサーという種類の検索ツールであり、検索対象である多言語 コーパスをデータベースに読み込み、検索を行う。本マニュアルでは、英語と日本語の2 言語のコーパスを読み込んで利用する場合について述べる。

AntPConc で検索すると,指定した検索語を中心に据えた Key Word In Context (KWIC) と 呼ばれる表示形式のコンコーダンス画面 (コンコーダンスライン)が,英語文と日本語文 の二言語について同時に得られる。

検索したい語句,たとえば"named"を Start ボタンの左側のボックスに入れて,Start ボ タンをクリックするか ENTER キーを押すと,このような検索結果が得られる。画面上半分 の KWIC 画面に"named"を含む英語文が表示され,それらの英文に対応する日本語文が, 画面下半分の Reference 画面に表示される。

a AntPConc 1.0.2	x
File Help	
Search Term 🗹 Words 🔲 Case Show Nth Entry 🔟 💌 Search Corpus 💿 Target 🔘 Reference	
(named) Start Context Size 20 vords	
Sort 1: CEN ▼ Sort 2: 1R ▼ Sort 3: 2R ▼ Sort Page Size All ▼ entries	
Target: Target Hits: 39	
Line KWIC	
1 An explorer named Amundsen reached it first.	<u> </u>
2 a added, "Paul Bunyan had a huge blue ox named Babe."	
3 His pony is named Blaze.	_
4 A man named Charless Darrow was creative.	_
5 Alex owns a sheepdog named Charlie.	-
6 Do you know an animal doctor named Dr. Strong?	-
Line Reference	
1 Amundsen という探検 家 が 最初 に到達した 。	<u> </u>
2 セレナ は 付け加え た 。「 Paul Bunyan は Babe という大きな土気色の ウシを 飼って いた 。」	
3 彼の馬は Blaze という名前です。	
4 Charless Darrow という名の男はクリエイティブでした。	
5 AlexはCharlie という名前の牧羊犬を所有している。	
6 あなたは Dr. Strong という獣医 をご存じですか。	

本マニュアルでは, AntPConc に実装された機能とその使い方を説明する。さらに, AntPConc で検索に用いる自作の多言語パラレルコーパスの作成方法を紹介する。

○ 準備編

AntPConc のダウンロード

http://www.antlab.sci.waseda.ac.jp/software.html#antpconc にアクセスするか,あるいはイン ターネットで「Anthony Laurence」を検索して最初に得られる検索結果をダブルクリックし, 出てきた右上の Latest Software Releases の中から, AntPConc の最新版をみつけてダブルクリ ックする。次に,AntPConc の Download の文字の下のブルーの文字の部分 (AntPConc 1.0.2w (beta))を右クリックする。「対象をファイルに保存」を選び,保存先を「デスクトップ」に 選び「保存」を押す。



検索ファイル(コーパスデータ)の読み込みと保存:

AntPConc を使って検索する際には、コーパスデータを用意する必要がある。英語コーパス (ここでは 607-en.txt と呼ぶ) と日本語コーパス (607-ja.txt と呼ぶ) をデスクトップにおく。 AntPConc もデスクトップにおく。

あるいは、USB ドライブに AntPConc,英語コーパス、日本語コーパスを保存しておいて 使用することも可能である。ただし、上記のファイルを USB ドライブに直接入れること、 すなわちフォルダに入れないよう留意する。

① File

AntPConcのアイコンをダブルクリックすると、初期画面が得られる。

AntPConc 1.0.2		- - x
File Help		
Search Term 🗹 Words 🔲 Case	Show Nth Entry 1 Search Corpus	s 🖲 Target 🔘 Reference
	Start Contex	t Size 20 🔻 words
Sort 1: CEN 🔻 Sort 2: 1R 🔹	Sort 3: 2R 🔻 Sort Page Si	ze All • entries
Target: Target Hits: 0	0% Font Size	
First Previous Next Last		<u>^</u>
Line	KWIC	
		E
Line	Reference	
		-

続いて, File ボタンをクリックすると, メニュー画面が現れる。

🦥 AntPConc 1.0.2	
File Help	
Load Corpus Ctrl+O	Show Nth Entry 1 Search Cornus @ Target O Reference
Build/Edit Corpus Ctrl+N	- Show with Entry - Search Corpus O Harder O Holdeneo
Save Corpus	Start Context Size 20 🛛 words
Save Results Ctrl+S	Sort 3: 2R 👻 Sort Page Size All 👻 entries
Exit Target: Target Hits: 0	0% Font Size
First Previous Next Last	
Line	KWIC

2 Build/Edit Corpus

メニュー画面の2番目のメニューBuild/Edit Corpus にカーソルを合わせてクリックする。 'AntPConc Corpus Builder'で Choose Files をクリックする。

🍓 AntPCor	nc 1.0.2
File Help	
Search Ter	m 🗹 Words 🗌 Case Show Nth Entry 1 💌 Search Corpus 💿 Target 🔿 Reference
	Start Context Size 20 💌 words
Sort 1: CE	N 💌 Sort 2: 1R 💌 Sort 3: 2R 💌 Sort Page Size All 💌 entries
Target: Ta	🧖 AntPConc – Corpus Builder
First F	Add Corpus Clear Selected Corpus Clear All Corpora Corpus 1 Corpus 2
	Display Name Target
Line	
	Choose Files Clear Selected File Clear All Files
	Cancel Update Corpus

デスクトップから「607-en.txt」(英語コーパスデータ)を指定して、「開く」をクリックする。ウィンドウに「607-en.txt」が入る。続いて、Corpus 2 をクリックし、Choose Files をクリックして、デスクトップから「607-ja.txt」(日本語コーパスデータ)をクリックする。ウィンドウに「607-ja.txt」が入る。

🐲 AntPConc - Corpus Builder	AntPConc - Corpus Builder
Add Corpus Clear Selected Corpus Clear All Corpora Corpus 1 Corpus 2 Display Name Target 607-en.txt	Add Corpus Clear Selected Corpus Clear All Corpora Corpus 1 Corpus 2 Display Name Reference 607-ja.txt
Choose Files Clear Selected File Clear All Files Cancel Update Corpus	Choose Files Clear Selected File Clear All Files Cancel Update Corpus

なお、Choose Files の作業を早めるため、「ドラッグアンドドロップ」を使うことも可能である。また、各コーパスを *AntPConc* の画面で区別するため、Display Name で表示するコーパス名を入力することができる。ここで Update Corpus のボタンをクリックすれば、検索を開始することができる

③ Save Corpus

複数のコーパスデータを読み込む手間を次回から省くために、読み込んだ複数コーパス データを1つの「データベースファイル」として保存しておく便利な方法について述べる。 ②の手順でコーパスデータを読み込んだ後、File の3番目のメニューの Save Corpus にカ ーソルを合わせてクリックする。データベースファイルの保存場所と保存ファイル名(ピ リオドで区切られた右側の文字列部分の拡張子に.db が付く)を指定して「保存」を押す。 ここでは、「607_corpus.db」という名前で保存した。

④ Load Corpus

前述の③Save Corpus の手順を済ませておくと、次回からまず、Load Corpus をクリック して、ファイル名(本稿では「607_corpus.db」)を指定し、「開く」を選択するとすぐに AntPConc が検索可能な状態になる。

授業や自習で AntPConc を使用する時には、データベースファイル(「607_corpus.db」)を 学習者に配布するとスムーズに検索作業に入ることができる。

○ 検索編

AntPConc のデフォルト設定で十分な検索が可能であるが、さらに詳細な検索設定が可能である。

Search Term : 検索語の大文字・小文字の区別

*AntPConc*のデフォルト設定では、検索の際、英語の大文字と小文字は区別しない。小文字で検索語句を"must"と入力しても、大文字の"Must"も検索結果として出力される。

一方, Search Term の Case の左の□をクリックして Z Case に設定すると,大文字と小文字 を区別した検索が実行される。例えば, Z Case 設定で"There"を検索すると,小文字の "there"を含む文は出力されず,"There is/are"の文を出力することが可能である。

AntPConc 1.0.2		x
File Help		
Search Term 🗹 Wo	ords 🔽 Case Show Nth Entry 20 🔻 Search Corpus 💿 Target 🔘 Ref	erence
There	Start Context Size 20 🔻 wo	ords
Sort 1: CEN 🔻 S	ort 2: 1R 🔻 Sort 3: 2R 🔻 Sort Page Size All 💌 er	ntries
Target: Target Hits	s: 115 100% Font Size	
Line	KWIC	
1	There are a couple of security camera	
2	There are far more flowers than there	
3	There aren't any cleaners. We have to	=
4	There is a chipmunk.	
5	There is the visitors' lobby.	-
Line	Reference	
1	防犯力メラが数台あります。	*
2	昨年の、この時期よりはるかに多くの花が咲いている。	
3	掃除機はない。わたしたちはあらゆる掃除をしなければならない。	=
4	シマリス がいる 。	
5	来場者のロビーがあります。	

Search Corpus : 検索コーパスの選択

AntPConc を用いると、複数の言語のコーパスを検索することができる。Search Corpus の ラジオボタンを指定のコーパス(ここでは Target あるいは Reference)に設定することによ って、コンコーダンス画面に表示する言語を選択できる。3 つ以上のコーパスを取り込んで いる場合、各コーパスの指定名がここに表示される。

AntPConc を使用する多くの場合, Search Corpus は Target (英語コーパス) が選択されており,検索結果が KWIC 画面(上半分)に表示される。また, Reference (日本語コーパス)の検索結果が Reference 画面(下半分)に表示される。

一方, Search Corpus を日本語コーパスに設定することも可能である。検索語句として"彼 女のもの"を検索した結果を示した。日本語の"彼女のもの"を含む検索結果が KWIC 画 面に, "hers"を含む英語の検索結果が Reference 画面に表示される。

antPConc 1.0	0.2
File Help	
Search Term 彼女のもの	✓ Words ♥ Case Show Nth Entry 1 ▼ Search Corpus ○ Target ◎ Reference Start Context Size 20 ▼ words
Sort 1: CEN	▼ Sort 2: 1R ▼ Sort 3: 2R ▼ Sort Page Size All ▼ entries
Target: Refere	ence Hits: 5 100% Font Size
Line	KWIC
1	一 面 に 掲載 した 記事 は <u>彼女</u> のもの で ある 。
2	彼の ベルトは 彼女 のものよりもよりよく見える。
3	払のものを描くつもりです、そして彼女は彼女のものを描くつもりです。
4	Liz は 彼女 のものを 持ってくる でしょう。
5	私はその自転車は彼女のものであるとは思いません。
Line	Reference
1	The story on the front page is hers.
2	His belt looks better than hers.
3	I'll paint mine, and she'll paint hers.
4	Liz will bring hers too.
5	I don't think that the bike is hers.

Show Nth Entry : 検索結果の表示件数

AntPConc のデフォルト設定では、Show Nth Entry は1に設定されており、すべての検索 結果が表示される。検索結果の件数が多い時には、スクロールしてすべてのコンコーダン スラインを観察して、そこからルールを導き出すのは学習負担が大きい。そこで、Show Nth Entry を 5、10、20、50 の中から選択して、検索結果を 5 文ごと、10 文ごと、20 文ごと、50 文ごとに表示することによって、コンコーダンスラインの概要を把握し易くすることが可 能となる。 例えば、"have to"の用法を観察する場合、Show Nth Entry を 10 に設定すると、 合計 41 行の検索結果の表示件数が 5 行の例文表示となり、少数の英語と日本語文を対比し た比較・観察が可能となる。

🦥 AntPConc 1.0.2	
File Help	
Search Term 🗹 Words 🗌 Case	Show Nth Entry 🔟 Search Corpus 💿 Target 🔘 Reference
	Context Size 20 v words
Sort 1: CEN 🕶 Sort 2: 1R 💌	Sort 3: 2R V 10 t Page Size All V entries
Target: Target Hits: 0	0% Font Size 50
First Previous Next Last	
Line	KMIC

🐲 AntPCond	: 1.0.2
File Help	
Search Terr	m 🗹 Words 🛛 Case Show Nth Entry 10 🔻 Search Corpus 💿 Target 🔘 Reference
have to	Start Context Size 20 vords
Sort 1: CE	N ▼ Sort 2: 1R ▼ Sort 3: 2R ▼ Sort Page Size All ▼ entries
Target: Tar	get Hits: 41 100% Font Size
Line	KWIC
1	When you get married you have to get used to sharing your life with
2	We've lost so much time already, so we'll have to make it up next week.
3	Tomorrow's the day when we have to make a final decision.
4	I should have to stayed on longer at school.
5	Did Nadia have to wear golf shoes last time?
Line	Reference
1	結婚 する 時 は 誰 かと 人生 を 共有 する ことに 慣れ なけれ ば なら ない 。
2	私たちは既に何回も負けてしまったので、来週は埋め合わせをしなければなりません。
3	明日は私たちが最終決定をしなければならない日です。
4	私は学校に長くとどまっていればよかった。
5	Nadiaは前回はゴルフシューズを履かなければなりませんでしたか。

Font Size:フォントサイズの調節

AntPConc では、Font Size の目盛をスライドさせることによってフォントサイズを調節す ることができる。フォントサイズを一目盛大きくすると、コンコーダンスラインが非常に 見やすくなり、学習者はコンピュータ画面から顔を離すことが可能となる。また、机間巡 視をする教師にも学習者のコンピュータ画面が見やすくなり、学習活動の進捗状況の把握 が容易になる。



Sort:並べ替え

*AntPConc*では、検索結果を観察しやすくするために、検索語および検索語の左右の語を ABC順に並べ替えるソート(Sort)機能が付いている。検索結果の表示順序を、検索語から 左右3番目の語まで、第1(Sort1)、第2(Sort2)、第3(Sort3)の3つのソートキーごと に優先順位を指定することができる。

デフォルトでは、Sort 1 が CEN (検索語), Sort 2 が 1R (検索語の右 1 語目), Sort 3 が 2R (検索語の右 2 語目) (いわゆる右ソート) と設定されており, Search ボタンをクリックするとその優先順位でソートした結果が表示される。また、一度出力されたコンコーダンスラインに対して、新たにソートの基準を変えて Search ボタンをクリックすると新たな表示 順序に並べ替えることができる。

🦥 AntPConc 1.0.2			
File Help			
Search Term 🗹 Words	Case Show Nth Entr	y 1 💌 Search Corpus 📀	Target 🔿 Reference
		Start Context Size	20 🖌 words
Sort 1: CEN 💌 Sort 2:	1R 🕶 Sort 3: 2R 👻	Sort Page Size	All 🖌 entries
Target: Target Hits: 0	2L 0% Font Size	<u> </u>	
First Previous Nex	1L CEN		
Line	1R	KMC	
	2R		
l	3R		

Page Size :ページサイズの調節

大量の出力結果を表示する際,1ページ(1画面)に表示する行数を選択することができる。*AntPConc*では,検索結果のPage Size を 10 行,50 行,100 行,500 行,1000 行, All の 6 通りから選択できる。デフォルトは All に設定されており,1ページに検索結果のすべての行が表示されるので,スクロールして見る必要がある。

例えば、"as * as"の検索では、Page Size を 10 に設定した検索結果を示した。1ページ に 10 行ずつ結果が表示される。2ページ目を見るには KWIC 画面の上に表示されたページ ボタンのタブの 2 をクリックする。3 をクリックすると3ページ目が表示される。スクロー ルしなくてよいので、出力結果の件数が多い時には、この機能が便利である。

AntPConc 1.0.2	ור
le Help	
earch Term 🗹 Words 🗋 Case Show Nth Entry 1 🍸 Search Corpus 💿 Target 🔘 Refere	ence
Start Context Size 20 💌 words	;
ort 1: CEN 🗸 Sort 2: 1R 🗸 Sort 3: 2R 🖌 Sort Page Size All 💙 entri	es
arget: Target Hits: 0 0% Font Size 50	
First Previous Next Last	
Line KMIC 1000	
All	
AntPConc 1.0.2	۲.
File Help	
Search Term 🖉 Words 🖉 Case Show Nth Entry 💶 🛪 Search Cornus @ Target @ Reference	
as * as Start Context Size 20 vords	
Sort 1: CEN 🔻 Sort 2: 1R 🔻 Sort 3: 2R 🔻 Sort Page Size 10 💌 entrie	25
	_
Total Tota	
First Previous 1 2 3 4 Next Last	
Line KWIC	
1 Is a tiger as big as a lion?	
2 Its bill is almost as big as its body, exclaimed Kim.	
3 Good grief. That one is as big as a house!	
4 Motor-racing is as dangerous as boxing.	
5 Cuba isn't as dry as Malta.	
6 Run as fast as you can.	
7 I am picking the apples as fast as I can.	
8 Hurray! The sail is as good as new!	
9 Jazz isn't as good as opera.	
10 Clap as hard as you can.	-
Line Reference	
1 とらは ライオンのように 大きいです か。	
2 「そのくちばしはその体とほぼ同じ大きさだ。」とキムは叫びました。	
3 信じられない。それは、家と同じくらい大きい!	
8 万歳!帆は新品同様だ!	
9 ジャズ はオペラ ほどょくありません 。	

Hits :検索ヒット数

Target Hits の右側に,検索結果の件数が表示される。例えば,教師のデモンストレーションの検索結果の件数を見ることで,学習者は自分が正しく検索できているかを自己チェックできる。また,検索結果の件数に応じて,Show Nth Entry 等の設定を変えて,コンコーダンスラインを観察することが可能となる。

AntPConc 1.0.2	
File Help	
Search Term 📝 W	ords 🗷 Case Show Nth Entry 🔟 🔻 Search Corpus 💿 Target 💿 Reference
as * as	Start Context Size 20 🔻 words
Sort 1: CEN 🔻 S	ort 2: 1R V Sort 3: 2R V Sort Page Size 10 V entries
Target: Target Hits	s: 39 100% Font Size
First Previous	1 2 3 4 Next Last
Line	KWIC
1	Is a tiger as big as a lion?
2	Its bill is almost as big as its body, exclaimed Kim,

Context Size :1行の長さの調節

Context Size は KWIC 画面での検索語句を含む左右の語数を示す。デフォルトでは 20 に設定されているが、 5 words, 10 words, 15 words, 20 words, 50 words の中から選択できる。

😻 AntPConc 1.0.2		
File Help		
Search Term 🗹 Words 🗌 Case Show Nth Entr	y 🔟 💙 Search Corpus 💿	Target 🔘 Reference
	Start Context Size	20 🕶 words
Sort 1: CEN 🕶 Sort 2: 1R 💌 Sort 3: 2R 💌	Sort Page Size	5 10 entries
Target: Target Hits: 0 0% Font Size	<u> </u>	15 20
First Previous Next Last		50
Line	KMIC	

その他

ワイルドカード:検索語句の入力にはワイルドカード(wildcard)と呼ばれる特殊記号 を利用することができる。「*」は任意の文字列を表し、例えば、as * as と入力すると, as bigas, as dangerous as, as dry as などが検索できる。

ハイライト: KWIC 画面あるいは Reference 画面のコンコーダンスラインをクリックすると,対応する Reference 画面あるいは KWIC 画面の両方のコンコーダンスラインがハイライト表示される。

○ 自作コーパスデータの作成編

AntPConc を検索する英語と日本語の2言語コーパス(あるいは多言語コーパス)データの作成方法を述べる。自分で収集した英文と日本語文からコーパスデータが作成可能である。

Step 1:「エクセル」で1つの列に英語の文をタイプする。隣の列に英語に対応する日本 語の対訳文をタイプする。

Step 2: 英語列をコピーして、「メモ帳」に貼りつけ、 UTF-8 形式で保存して、英語フ アイルを作成する。次に、日本語列をコピーして、「メモ帳」に貼りつけ、UTF-8 形式で保 存して、日本語ファイルを作成する。*AntPConc*で検索可能なコーパスデータが作成できる。

Step 3: このステップはとばすことも可能である。日本語テキストは、B列のように「分かち書き」をすると、検索の際、単語ごとにきれいに検索やソートが可能となる。分かち書きとは、文において、語の区切りに空白をはさんで区切って書き表すことである。英語の場合は単語と単語の間にスペースが入っているので分かち書きは不要である。分かち書きプログラムには、TECDICP(http://www.te-com.biz/download/page1.html)やJapanese Tutor(http://www.vector.co.jp/soft/dl/winnt/edu/se462303.html)が利用可能である。

X 🖢			
ファイノ	ファイル ホーム 挿入 ベージレイアウト 数式 データ 校閲 表示 活用しよう!エクセル		
標準		マロン ご <th>】並べて比!</th>	】並べて比!
B26 ▼ (<i>k</i>			
	A	В	(
1	The girl looks.	その少女は見ます。	
2	Two fish swim.	二匹の魚が泳ぐ。	
3	A bird sings.	鳥が歌う。	
4	The rabbits eat.	ウサギ は 食べ ます 。	
5	The baby sleeps.	赤ちゃん は 眠る 。	
6	Rusty stays home.	Rustyは、家にいる。	
7	The boys walk.	少年たちは歩く。	
8	Mom waves.	母は手を振る。	
9	A cat looks out.	猫 が 外 を 見る 。	
10	Birds eat food.	鳥は餌を食べます。	
11	The cow moos.	牛はモーと鳴く。	
12	Some ducks run.	何羽かのカモが走る。	
13	The dog barks.	犬 が ほえる 。	
14	The kitten jumps.	子 ネコ が ジャンプし ます 。	
15	The fish jumps.	魚 が ジャンプし ます 。	
16	That woman runs.	あの女は走る。	
17	They fly a kite.	彼ら は たこ を 上げる 。	
18	He sells ice cream.	彼 は アイスクリーム を 販売し て いる 。	
19	The little girl digs.	少女が 穴を 掘る。	
20	A man eats an orange.	男 は オレンジ を 食べる 。	

参考文献:

- Anthony, L. 3AntPConc (Version 1.0.2) [Computer Software]. Tokyo, Japan: Waseda University. Available from http://www.antlab.sci.waseda.ac.jp/, 2013.
- Anthony, L. Chujo, K., Kenji, Y. and Atsushi, M., Broadening the Scope of Parallel Corpus Tools: Using AntPConc in the DDL Classroom, Second Asia Pacific Corpus Linguistics Conference (APCLC 2014), 2014, March.
- Chujo, K., Anthony, L. and Oghigian, K., Teaching Remedial Grammar through DDL using AntPConc, AACL 2013 - American Association for Corpus Linguistics, San Diego State University, California, USA, 2013, January.
- Chujo, K., Chikako, N. and Oghigian, K., A DDL Approach to Remedial Grammar Using the AntPConc Parallel Concordancer, The 11th Asia TEFL International Conference, Ateneo de Manila University, Philippines, 2013, October.
- 5) 中條清美,西垣知佳子,「多言語検索ツール AntPConc を利用したリメディアル文法授業の実践」外国語教育メディア学会(LET)第53回全国研究大会,2013年8月9日,文 京学院大学.
- Chujo, K., Anthony, L., Oghigian, K. and Yokota, K., Teaching Remedial Grammar through Data-Driven Learning Using AntPConc, *Taiwan International ESP Journal*, 5(2), 2013, 65-90.

謝辞:本研究は平成 25-28 年度科学研究費補助金基盤研究(B)(25284108)を受けています。