

真正性のある接触場面  
コーパスが明らかにする  
学習者と母語話者の発話の違い

中俣尚己 (京都教育大学)

# 1. はじめに

- ◎ 「はごろも」は  
**「機能語」** 用例文  
データベースである。

# 機能語 V.S. 実質語

- ◎ 語彙が大切(山内2012)
- ◎ 「は」と「が」を教えるか、「焼肉」と「ビール」を教えるか。
- ◎ 実質語の方が教育では重要なのか？

# 第1の問い

- ◎ 外国語学習・教育において、機能語とはどういう位置づけか？
- ◎ コーパスのデータから上記の問いを考えてみる。

## 2. 『日中Skype会話コーパス』 の紹介

# 「日中Skype会話コーパス」とは

- ◎ 2012年5月～7月に、東京・実践女子大学と長沙・湖南大学の学生間で行ったSkypeを利用した遠隔日本語会話活動(中俣ほか2013)を録音、文字化したもの。接触場面会話コーパス。



# 「日中Skype会話コーパス」とは

- ◆ 中国人学習者は全員 2 年生。  
日本人は 3 年生～M1。
- ◆ 9ペア。38会話。
- ◆ 総会話時間46:48:35。  
1 会話あたり平均1:13:55。
- ◆ 語数は約20万語。  
(「雪だるま」で解析)
- ◆ 助成：JSPS(26770180)



# 利用するには

中俣のホームページから無料でダウンロードできます。

中俣 尚己 Naoki Nakamata

TOP

研究テーマ

業績一覧

著書

データベース

## データベース

### プロフィール



中俣 尚己  
Naoki NAKAMATA

#### 生年月日

1981年8月24日生まれ

#### 出身地

大阪府

#### 所属

京都教育大学教育学部准教授

#### メールアドレス

[nakamata@kyokyo-u.ac.jp](mailto:nakamata@kyokyo-u.ac.jp)

### 『日中 Skype 会話コーパス』

#### 『日中 Skype 会話コーパス』とは何ですか？

このコーパスは2012年5月から7月にかけて、日本・東京の実践女子大学と中国・長沙市の湖南大学との間で行われた日本語でのSkype会話交流活動の内容を、日本学術振興会の科研費若手研究(B)「**縦断型接触場面コーパスの構築とそれを用いた日本語教育のための談話研究 (課題番号26770180、研究代表者中俣尚己)**」の助成を受けて録音、文字化したものです。会話を楽しむことを主目的とした活動の録音であり、真正性のある言語行動のコーパスといえます。この活動自体の実践報告につきましては、中俣尚己ほか(2013)「[Skypeを活用した日中会話交流プログラム](#)」(『**実践國文學**』83所収)を御覧ください。

#### 『日中 Skype 会話コーパス』の概要

# 特徴 1 : 真正性がある

- コーパスのための活動ではない。
- Skypeを用いた会話活動を通し、中国の学習者には学んだ日本語を使う機会を提供するとともに学習意欲を継続させること、日本の母語話者には外国人と文化交流をしたり日本語を教えたりしながら、日本語について考えてもらうことが第一の目的。
- → 真正性のある  
接触場面の雑談。

# 表1 KYコーパスと日中Skype会話コーパスの出現数の比較

語	KYコーパス	Skypeコーパス
明後日	0	7
木曜	6	41
すごい	77	211
すごく	190	86
すげえ	0	4

「明後日」や「木曜」は基本語であるのに、コーパスに出現しにくい(北村・富岡・川村 2009)

## 特徴 2 : 縦断的なデータ

- 会話活動は 1 週間に 1 回、継続的に行った。
- 最も多いペアで 7 回分の会話がある。

## 特徴 3 : 一種の電話場面

- ◆ 終結部などは電話場面そのものの展開が観察される(橋内1999)。
- ◆ コミュニケーション・ブレイクダウンや沈黙の研究にもどうぞ。

# 特徴4：話題が指定されている

1	ポップカルチャー	6	伝統・行事
2	料理	7	夏休み・夏の予定
3	家庭・家族・子供	8	大学生活
4	故郷・ 今住んでいる場所	0	指定なし・トピック認定できず
5	敬語		

必ずしも厳密に守られているわけではなく、  
話がそれたり日本語についての質問も。

### 3. 学習者と母語話者の発話の違い

# 学習者の発話 VS 母語話者の発話

	学習者	母語話者
Token頻度	104,156	100,325
Type頻度	5,434	5,217
TTR	0.052	0.052

学習者と母語話者の（使用）語彙量に、差異は存在しない。



NS/NNSの違い  
話題の違い

# 表1 品詞ごとの学習者と母語話者の使用頻度

	学習者	母語話者
副詞	3,477	5,646
助動詞	9,693	14,421
動詞	7,807	10,531
接続詞	578	748
助詞	26,510	29,133
その他	1,389	1,310
数量詞	1,416	1,309
※形容詞	7,002	6,163
名詞	23,853	19,098
感動詞	22,431	11,966

# 図 品詞ごとの母語話者を1とした時の 学習者のToken頻度の割合



# まとめ

- ◎ 名詞・形容詞の差は少ない。
- ◎ 副詞・助動詞・助詞の差は大きい。
- ◎ 動詞の差も大きいですが、これは「～という人」のような「という」の使用に学習者と母語話者に大きな差が見られたため。
  
- ◎ 実質語の差は小さく、機能語の差は大きい。
- ◎ なぜか？

- ◎ J：（略）パスタ、日本料理じゃなくて、パスタ。＜うーん＞パスタ分かります？ パスタ。
- ◎ C：さった？ [サラダ？
- ◎ J：[パスタ。＜うん＞パスタ。
- ◎ C：んー？
- ◎ J：待ってね。《ポーズ 8秒》日本料理じゃないですよ、もう。パスタ。
- ◎ C：**ああ、パスタ**。＜うん＞ピザですね。ん？ ピザです。
- ◎ J：あの、麺。ピザとはちょっと違うかな。＜うーん＞麺、＜麺＞のパスタとか。[うん。
- ◎ C：あー。
- ◎ J：ピザとか、お昼とかはそういうのみんな食べますね。今の若い子たちは。
- ◎ C：**このパスタは＜うん＞食材はなんですか。** うーん。
- ◎ J：麺です。パスタ。パスタ。＜麺＞パスタ。知らない？ うん。[あの。
- ◎ C：[麺ですか。あー。

- ◎ C : もし、じゅ、授業中、あー、先生の授業を、き、あー、聞かないで、ほかの人、ほかの人と、あー、しゃべるの、えー、のとき、えー、先生は、私、私たちの耳を、あー、しぼる？
- ◎ J : 耳を？ ぴゅ？
- ◎ C : はい、●●です。
- ◎ J : ああ。それはしないですかね。ふーん。
- ◎ C : ●。
- ◎ J : しぼる。えーと、つー、つねる。
- ◎ C : ●あの、●●●●なんですか。あ、つねる。
- ◎ J : うん。こうやって、こう、ぐっていうのね。

# Q1. 言語学習における機能語の位置づけとは？

- ◎ 実質語は意味交渉の対象になり、現実のコミュニケーションにおいても学習される。
- ◎ 機能語は意味交渉の対象とならない。コミュニケーションを通じて習得することが難しい。

	学習者	母語話者
とても	159	19
すごく	13	76
すごい	12	204

- ◎ 機能語の教育・明示的な解説の重要性

## 4. 学習者と母語話者の 機能語の使用の違い



# 回数の差をどう表現するか？

◎ 学習者 2回

◎ 母語話者 10回

◎ 8回の差.....？

◎ 5倍の差.....？

# 対数尤度比(LLR)を計算

## ◎ 田中・近藤(2011)の補正值

$$2(a \ln a + b \ln b + c \ln c + d \ln d - (a+b) \ln(a+b) - (a+c) \ln(a+c) - (b+d) \ln(b+d) - (c+d) \ln(c+d) + (a+b+c+d) \ln(a+b+c+d))$$

- ◎ a : 当該資料での当該語の度数    b : 参照資料での当該語の度数  
c : 当該資料の延べ語数 - a    d : 参照資料の延べ語数 - b
- ◎ lnは自然対数を表す。aまたはbが0の場合、 $a \ln a$ または $b \ln b$ を0として計算する。
- ◎  $ad - bc < 0$  の場合の場合、-1 を乗じる補正を行う。
- ◎ この値を使って、学習者がよく使う語、母語話者がよく使う語をピックアップした。

# 母語話者の方が多い語

順位	語	NNS	NS	LLR	品詞
1	ん	292	1,858	1506.03	助詞
2	ね	837	2,732	1350.799	助詞
3	って	81	976	1023.868	助詞
4	いう	91	909	895.6327	動詞
5	だ	542	1,769	871.0089	助動詞
6	そう	650	1,904	833.2901	副詞
7	です	3,262	5,256	806.1719	助動詞
8	てる	97	845	789.9937	助動詞
9	んだ	21	339	384.4708	助動詞
10	なんか	30	350	362.7626	助詞

# 母語話者の方が多い語

順位	語	NNS	NS	LLR	品詞
11	けど	159	591	328.1095	助詞
12	とか	371	889	297.2483	助詞
13	か	1882	2723	295.3339	助詞
14	かな	28	290	288.8451	助詞
15	こう	34	303	285.3543	副詞
16	へー	74	393	285.3003	感動詞
17	ので	67	348	249.0346	助動詞
18	凄い	17	206	215.9501	形容詞
19	ふーん	62	297	201.2889	感動詞
20	て	776	1253	189.0765	助詞

# 母語話者の方が多い語

順位	語	NNS	NS	LLR	品詞
21	ない	365	720	172.4779	助動詞
22	何	227	529	169.2611	名詞
23	なの	16	162	160.0019	助詞
24	ごめん	11	144	154.4288	名詞
25	結構	9	134	148.836	副詞
26	たり	44	209	140.7661	助詞
27	よ	389	687	129.7973	助詞
28	言う	91	278	126.7176	動詞
29	【NNSの名前】	51	209	125.9171	英文字
30	じゃあ	146	354	120.1478	接続詞

# 母語話者の方が多い語

順位	語	NNS	NS	LLR	品詞
31	御	528	829	115.2778	接頭辞
32	ちょっと	346	595	105.4008	副詞
33	ねえ	23	136	105.2087	助詞
34	もう	142	327	102.6044	副詞
35	で	826	1122	95.27596	助詞
36	やつ		58	88.6091	名詞
37	ふん		57	87.08114	感動詞
38	やる	59	185	86.73167	非自立 動詞
39	だった	13	94	80.76294	助動詞
40	な	9	84	80.48723	助詞

# 母語話者の方が多い語

順位	語	NNS	NS	LLR	品詞
41	が	1622	1909	79.94952	助詞
42	なさる	11	82	71.46721	非自立 動詞
43	みたいに	36	130	70.16532	助動詞
44	も	681	898	67.26558	助詞
45	後	114	240	64.77871	名詞 副 詞
46	そうそう	3	54	62.7535	感動詞
47	出る	31	112	60.4734	動詞
48	とかで	1	45	60.36489	助詞
49	た	744	942	58.53392	助動詞
50	凄く	13	76	58.39751	形容詞

# 学習者の方が多い語

順位	語	NNS	NS	LLR	品詞
1	ええ	1286	118	-987.064	感動詞
2	ああ	3285	1153	-808.738	感動詞
3	は	4470	2004	-670.986	助詞
4	えー	1923	584	-587.231	感動詞
5	うーん	2015	707	-493.62	感動詞
6	【日本人の名前】	421	12	-437.189	英文字
7	え	968	212	-428.352	感動詞
8	えっ	1288	375	-415.298	感動詞
9	あの	2360	1025	-378.698	形容詞
10	と	3702	1937	-353.621	助詞



# 学習者の方が多い語

順位	語	NNS	NS	LLR	品詞
11	あー	200	9	-190.512	感動詞
12	そして	165	3	-181.535	接続詞
13	あ	2266	1267	-168.362	感動詞
14	この	567	180	-161.603	形容詞
15	の	4483	2941	-150.339	助詞
16	私たち	130	4	-133.255	複合名詞
17	とても	159	19	-107.611	副詞
18	中国	603	256	-101.154	名詞
19	私	937	471	-100.561	名詞
20	あの	438	160	-99.3597	感動詞

# 学習者の方が多い語

順位	語	NNS	NS	LLR	品詞
21	えと	93	3	-94.5644	名詞
22	おお	137	17	-90.8819	感動詞
23	だから	226	57	-86.3025	助詞
24	長沙	70	4	-62.8078	名詞
25	さん	427	194	-60.7942	非自立名詞
26	彼	79	10	-51.8323	名詞
27	あのお	396	187	-50.8887	感動詞
28	あああ	139	38	-48.3155	感動詞
29	ただ	100	20	-47.8718	副詞
30	お	157	50	-44.4618	感動詞

# 学習者の方が多い語

順位	語	NNS	NS	LLR	品詞
31	ている	410	214	-38.8384	助動詞
32	つまり	35	1	-36.2954	副詞
33	日本語	122	38	-35.681	複合名詞
34	日本	750	465	-34.3565	名詞
35	ません	374	198	-33.7351	助動詞
36	だいたい	134	47	-32.5791	名詞 副詞
37	ルームメート	25		-31.3615	名詞
38	いろいろ	143	54	-30.4842	副詞
39	彼女	35	3	-27.4985	名詞
40	習慣	49	8	-27.4514	名詞

# 学習者の方が多い語

順位	語	NNS	NS	LLR	品詞
31	う	85	25	-26.909	感動詞
32	についての	30	2	-25.7262	助詞
33	いいえ	33	3	-25.3277	感動詞
34	それから	20		-25.0889	助詞
35	田舎	68	18	-24.5589	名詞
36	はい	2027	1498	-24.152	感動詞
37	以前	19		-23.8344	名詞 副詞
38	彼ら	17		-21.3254	複合名詞
39	学生	71	22	-20.9159	名詞
40	単語	22	1	-20.8984	名詞

# まとめ

- ◎ 述語に接続する機能語  
母語話者 >>> 学習者  
特に終助詞（ね、かな、な）
- ◎ 接続詞  
学習者 >> 母語話者
- ◎ 感動詞  
学習者 >>> 母語話者  
「へー」「ふーん」「ふんふん」は母語話者のみ
- ◎ 副詞  
学習者に多いものと母語話者に多いもの。

## 5. おわりに

# 学習者と母語話者の差の考察

- ◎ 意味が学習者にとって把握しにくい、学習者が使っていない形式。＝終助詞  
JFL環境の学習者が習熟するのは難しい。
- ◎ 文化的な要因 （母語話者の「ごめん」の多用）
- ◎ 母語の影響 （学習者の一人称の多用）
- ◎ 学習者と母語話者で同じ機能を表現しているが、品詞が異なる。  
（学習者＝接続詞、母語話者＝接続助詞）

# 大きな問題点

- ◎ 学習者と母語話者で同じ機能を表現  
していて、品詞まで同じだが、使う形式が  
異なる。
- ◎ 学習者「いいえ」VS  
母語話者「いや」「いえ」
- ◎ 学習者「とても」VS母語話者「すごく」
- ◎ 学習者「なければならない」VS母語話者  
「なければいけない」
- ◎ 学習者にみられる  
「について」「にとって」



# 初級にしては硬すぎる語

- ◎ 初級で原因・理由表現としてきちんとした表現として、「ため」を導入しよう！とはまず考えないはず。
- ◎ 文体的には「ため」と同等である「なければならない」が初級で出現するのは全く不合理

# まとめ

	実質語	機能語
役割	タスクを達成するために必要	流暢に、正確に、適切に伝えるために必要
自然習得	易しい	難しい
何によって決まるか	話題 選択される	スタイル(文体)、話者？ 選択する
	選択される	選択する

# 参考文献

- 北村達也・富岡洋介・川村よし子(2009)「IDF を用いた単語レベル判定システムの構築と検証」『日本語教育方法研究会誌』16(1), pp.52-53
- 田中牧郎・近藤明日子(2011)「教科書コーパス語彙表」『言語政策に役立つ、コーパスを用いた語彙表・漢字表等の作成と活用』pp.55-63, 2011 文部科学省科学研究費特定領域研究「代表性を有する大規模日本語書き言葉コーパスの構築：21世紀の日本語研究の基盤整備」言語政策班
- 中俣尚己(2015)「「日中Skype会話コーパス」を用いた話題別語彙の抽出―「食」の場合―」『第8回コーパス日本語学ワークショップ予稿集』pp.11-18
- 橋内武(1999)『ディスコース 談話の織りなす世界』くろしお出版
- 山内博之(2012)「非母語話者の日本語コミュニケーション能力」野田尚史(編)『日本語教育のためのコミュニケーション研究』pp.125-144, くろしお出版
- Long, M.H. (1983) Native speaker/non-native speaker conversation and the negotiation of comprehensible input, "Applied Linguistics" 4(2):126-141
- Tono, Y., Yamazaki, M., Maekawa, K. (2013) *A Frequency Dictionary of Japanese*, Routledge

# 参考サイト

- ◎ 日中Skype会話コーパス

<http://nakamata.info/database.html>

- ◎ 雪だるまプロジェクト

<http://snowman.jnlp.org/>

発表は以上です。  
ご意見・ご質問よろしく申し上げます。

利用希望者は  
「Skypeコーパス」で検索！

# おまけ (山内2012)

カテゴリー	実質語A	実質語B	実質語C	機能語
例	アパート	首都	昔	を
具体・抽象	具体物	抽象概念		
話題との関係	話題に従属する		話題に従属しない	
辞書的意味	あり			なし