

## 表紙 , 目次

著者	国立国語研究所コーパス開発センター
雑誌名	言語資源活用ワークショップ発表論文集
巻	1
ページ	1-1
発行年	2017
URL	<a href="http://id.nii.ac.jp/1328/00001451/">http://id.nii.ac.jp/1328/00001451/</a>

# 言語資源活用ワークショップ 2016

## 発表論文集

2017年3月6日(月) 『語彙資源活用シンポジウム』  
2017年3月7・8日(火・水) 『言語資源活用ワークショップ2016』

大学共同利用機関法人 人間文化研究機構  
国立国語研究所 コーパス開発センター 編

# Programme

# Programme: 語彙資源活用シンポジウム

2017年3月6日(月)

## 【セッション1】(2F 講堂)

- 10:10-10:15 趣旨説明  
..... 浅原正幸(国立国語研究所)
- 10:15-10:45 『UniDic』の拡張計画  
..... 岡照晃(国立国語研究所)
- 10:45-11:15 単語分かち書き用辞書『mecab-ipadic-NEologd』を公開して得た  
知見について  
..... 佐藤敏紀(LINE)
- 11:15-11:45 拡張型NLP『JMAT』における実利用に向けた形態素解析のリソー  
スチューニング  
..... 北浦雅子・紀伊馬章(ジャストシステム)
- 11:45-13:00 休憩

## 【セッション2】(2F 講堂)

- 13:00-13:30 『JUMAN++』の大規模語彙獲得へ向けた取り組み  
..... 森田一(京都大学)
- 13:30-14:00 『分類語彙表』の特徴と問題点  
..... 山崎誠(国立国語研究所)
- 14:00-14:15 休憩

## 【セッション3】(2F 講堂)

- 14:15-14:45 『日本語歴史コーパス』に出現した新規語の『UniDic』への登録に  
ついて  
..... 鴻野知暁(国立国語研究所)
- 14:45-15:15 『日本国語大辞典』の編集方法—これまでとこれから  
..... 佐藤宏(小学館)
- 15:15-15:45 中型国語辞典『大辞林』編集と見出し語の収集・選定について—未  
知語・新語を中心に  
..... 山本康一(三省堂辞書出版部)
- 15:45-16:00 休憩

## 【パネルセッション】(2F 講堂)

- 16:00-17:00 パネルセッション・総合討論

## Programme: 言語資源活用ワークショップ 2016

**2017年3月7日(火)**

- 10:00-10:15 ■挨拶 (2F 講堂) 前川喜久雄
- 10:15-11:05 ■口頭発表 A グループ (2F 講堂)
- [O-A-1]  
国語教科書と高校生作文の複文構造比較—従属節の構造と節形式の量的比較—  
..... 松本理美 (立命館大:学生)
- [O-A-2]  
友人への「断り」に対する評価に関する質的考察 —日本語母語話者と中国人日本語話者の評価を通して—  
..... 藤越 (東京大:学生)
- 11:05-11:55 ■招待講演 (2F 講堂)
- [I-1]  
講演・講義の音声認識と字幕作成へのコーパスの活用  
..... 秋田祐哉 (京都大学)
- 12:00-13:00 休憩
- 13:30-15:00 ■『国語研日本語ウェブコーパス』検索系『梵天』デモ (2F セミナー室 238 室)

- 13:00-14:15 ■ポスター発表 A グループ (2F フロア・多目的室)
- [P-A-1]  
もし小学生が『現代日本語書き言葉均衡コーパス』並みに漢字を使ったら  
..... 今田水穂 (文部科学省)
- [P-A-2]  
コーパス構築における発話アライメントの現状  
..... 石本祐一 (国語研)
- [P-A-3]  
発話文への発話者情報付与の基本設計 — 『現代日本語書き言葉均衡コーパス』収録の小説を対象に—  
..... 宮寄由美・柏野和佳子・山崎誠 (国語研)
- [P-A-4]  
夢梅本『倭玉篇』全文テキストデータベースの構築  
..... 高橋大希・劉冠偉 (北海道大:学生)・池田証壽 (北海道大)
- [P-A-5]  
『日本語諸方言コーパス』の構築について  
..... 木部暢子・佐藤久美子・中西太郎 (国語研)  
中澤光平 (与那国町与那国語辞典編集業務嘱託員)
- [P-A-6]  
相談における談話構造 — 修辞機能と脱文脈化の観点からの分析—  
..... 田中弥生 (国語研・東京大:学生)
- [P-A-7]  
『UniDic』と『分類語彙表』の見出し対応表データの構築  
..... 近藤明日子 (国語研)・田中牧郎 (明治大)
- [P-A-8]  
『名大会話コーパス』の比較に基づく教室談話における「中途終了型発話」の特徴  
..... 矢田真菜 (東京学芸大:学生)
- 14:15-14:20 休憩 (ポスター切替)

14:20-15:35	<p>■ポスター発表 B グループ (2F フロア・多目的室)</p> <p>[P-B-1] 『多言語母語の日本語学習者横断コーパス』の母語話者データにおけるタスクと産出語彙の関連 ..... 小西円 (国語研)</p> <p>[P-B-2] 『現代日本語書き言葉均衡コーパス』に対する分類語彙表番号アノテーションの試行 ..... 加藤祥・浅原正幸・山崎誠 (国語研)</p> <p>[P-B-3] 「大規模日常会話コーパス」プロジェクト—コーパスに基づく話し言葉の多角的研究— ..... 小磯花絵 (国語研)</p> <p>[P-B-4] 日本語語構成情報データベースの構築 ..... 浅尾仁彦 (情報通信研究機構)</p> <p>[P-B-5] 発話文自動生成のための日本語表現文型辞書の作成 夏目和子 (名古屋大)・刀山将大 (名古屋大:学生)・佐藤理史 (名古屋大)</p> <p>[P-B-6] スマホで古辞書 — 『篆隸万象名義』のIDS 検索を例に— ..... 劉冠偉・李媛 (北海道大:学生)・池田証壽 (北海道大)</p> <p>[P-B-7] 機械翻訳用超大規模辞書データ資源 ..... 春遍雀來 (日中韓辞典研究所)</p> <p>[P-B-8] モンゴル語アクセント研究のためのデータベース ..... 玉栄 (内モンゴル大・国語研)・西川賢哉・前川喜久雄 (国語研)</p> <p>[P-B-9] 多重の読みを持つテキストのコーパス化 ..... 小木曾智信 (国語研)</p>
15:35-15:45	休憩

15:45-17:25

■口頭発表 B グループ (2F 講堂)

[O-B-1]

次元形容詞にみる母語話者らしい日本語形容詞の使用

..... 西内沙恵 (国語研・立教大)

[O-B-2]

日本語コーパスの包括的検索環境の実現に向けて

前川喜久雄・浅原正幸・小木曾智信・小磯花絵・木部暢子・迫田久美子 (国語研)

[O-B-3]

機能語用例文データベース『はごろも』の今後の展開

..... 堀恵子 (東洋大・筑波大)・内丸裕佳子 (岡山大)・加藤恵梨 (朝日大)

小西円・山崎誠 (国語研)・江田すみれ (日本女子大)

建石始 (神戸女学院大)・中俣尚己 (京都教育大)・李在鎬 (早稲田大)

[O-B-4]

日本語学習者コーパスの教育応用における留意点—『多言語母語の

日本語学習者横断コーパス』に見る母語話者 L1 産出データの安定性

検証を中心に—

..... 石川慎一郎 (神戸大)

18:00-19:30

■懇親会



2017年3月8日(水)

- 10:10-11:00 ■口頭発表 Cグループ(2F 講堂)  
[O-C-1]  
漢語の仮名表記—実態と背景—  
..... 間淵洋子(明治大:学生)  
[O-C-2]  
『日本語歴史コーパス』短単位アノテーション作業効率化に向けた形態素解析用辞書『UniDic』の段階的特殊化の検討—近松コーパスを例として—  
..... 岡照晃(国語研)
- 11:00-11:50 ■招待講演(2F 講堂)  
[I-2]  
言語資源の設計・再設計と言語資源を活用した実習授業の設計  
..... 松吉俊(電通大)
- 11:50-13:00 休憩
- 13:00-15:30 ■『国語研日本語ウェブコーパス』検索系『梵天』デモ(2F セミナー室 238 室)

13:00-14:15

■ポスター発表 C グループ (2F フロア・多目的室)

[P-C-1]

全文検索システム『ひまわり』における言語分析支援機能の拡張

..... 山口昌也 (国語研)

[P-C-2]

児童生徒の「手」作文に於ける経年変化の計量的分析

阿部藤子 (東京家政大)・今田水穂 (文部科学省)・宗我部義則 (お茶の水女子大付属中)

富士原紀絵 (お茶の水女子大)・松崎史周 (日本女子体育大)・宮城信 (富山大)

[P-C-3]

『日本語日常会話コーパス』収録の進捗状況

..... 田中弥生・柏野和佳子・角田ゆかり (国語研)

伝康晴 (千葉大)・小磯花絵 (国語研)

[P-C-4]

『分類語彙表』の類義語と分散表現を利用した all-words 語義曖昧

性解消

..... 鈴木類 (茨城大:学生)・古宮嘉那子 (茨城大)・浅原正幸 (国語研)

佐々木稔・新納浩幸 (茨城大)

[P-C-5]

形態素解析ソフトウェア『Web 茶まめ』の改良と Web API の試

作

..... 川口寛治・薦田龍輝 (東京電機大:学生)・堤智昭 (東京電機大)

[P-C-6]

『現代日本語書き言葉均衡コーパス』を用いた「～ていく」「～てくる」構文の意味分析

..... 加藤麟太郎 (東京大:学生)・藤井聖子 (東京大)

[P-C-7]

明治初期教科書『物理階梯』のコーパス作成による語彙の考察

..... 田中牧郎 (明治大)・島田むつみ・高橋雄太 (明治大:学生)

[P-C-8]

話し言葉コーパスの転記タグ:『多言語母語の日本語学習者横断コーパス』と『日本語話し言葉コーパス』の比較

..... 西川賢哉 (国語研)

[P-C-9]

『日本語日常会話コーパス』の転記基準と作業工程

..... 川端良子 (国語研・千葉大:学生)・臼田泰如・西川賢哉 (国語研)

徳永弘子 (国語研・東京電機大)・小磯花絵 (国語研)

14:15-14:20

休憩 (ポスター切替)

14:20-15:35

■ポスター発表 D グループ (2F フロア・多目的室)

[P-D-1]

『現代日本語書き言葉均衡コーパス』と『分類語彙表』を利用した漢字 3 文字略熟語の抽出

..... 山崎誠 (国語研)

[P-D-2]

名詞項構造付与データの構築

..... 竹内孔一 (岡山大)

[P-D-3]

『名大会話コーパス』中納言版・ひまわり版公開データの作成

..... 柏野和佳子・西川賢哉・小磯花絵 (国語研)

[P-D-4]

『現代日本語書き言葉均衡コーパス』に対する節の意味分類情報アノテーション—基準策定, 仕様書作成の必要性について—

..... 松本理美 (立命館大:学生)・浅原正幸 (国語研)・有田節子 (立命館大)

[P-D-5]

『日本語話し言葉コーパス』における発声様式の自動分類

森大毅 (宇都宮大)・藤本雅子 (国語研)・浅井拓也 (北陸先端大:学生)・前川喜久雄 (国語研)

[P-D-6]

近代文語文の通時的変化の分析 —語種率・品詞率に着目して—

..... 近藤明日子 (国語研)

[P-D-7]

結合の強度を測る指標としての Log-r の有用性: 日・英語のバイグラムデータに基づく MI、LLR などとの比較

..... 藤村逸子 (名古屋大)・青木繁伸 (群馬大)

[P-D-8]

語彙・文型調査を目的とした『幼稚園の配布文書コーパス』の作成

..... 長谷川守寿 (首都大)・西尾広美 (国語研)

[P-D-9]

固有表現抽出におけるアノテーション手法の比較

..... 鈴木雅也 (茨城大:学生)・古宮嘉那子 (茨城大)

岩倉友哉 (富士通研)・佐々木稔・新納浩幸 (茨城大)

- 15:35-15:45 休憩
- 15:45-16:35 ■口頭発表 D グループ (2F 講堂)
- [O-D-1]
- 『現代日本語書き言葉均衡コーパス』への情報構造アノテーションの  
分析  
.....宮内拓也 (国語研・東京外大:学生)・浅原正幸 (国語研)  
中川奈津子 (千葉大・学振)・加藤祥 (国語研)
- [O-D-2]
- 読み時間と情報構造について (ちょっとながめ)  
..... 浅原正幸 (国語研)
- 16:35-17:00 ■クロージング (2F 講堂)

## 目次

国語教科書と高校生作文の複文構造比較—従属節の構造と節形式の量的比較—	[O-A-1]	
松本理美 (立命館大:学生)		2
友人への「断り」に対する評価に関する質的考察 —日本語母語話者と中国人日本語話者の評価を通して—	[O-A-2]	
藤越 (東京大:学生)		10
もし小学生が『現代日本語書き言葉均衡コーパス』並みに漢字を使ったら	[P-A-1]	
今田水穂 (文部科学省)		20
コーパス構築における発話アライメントの現状	[P-A-2]	
石本祐一 (国語研)		30
発話文への発話者情報付与の基本設計 —『現代日本語書き言葉均衡コーパス』収録の小説を対象に—	[P-A-3]	
宮寄由美・柏野和佳子・山崎誠 (国語研)		38
夢梅本『倭玉篇』全文テキストデータベースの構築	[P-A-4]	
高橋大希 (北海道大:学生)・劉冠偉・池田証壽		49
『日本語諸方言コーパス』の構築について	[P-A-5]	
木部暢子・佐藤久美子・中西太郎 (国語研)・中澤光平 (与那国町与那国語辞典編集業務嘱託員)		57
相談における談話構造 —修辞機能と脱文脈化の観点からの分析—	[P-A-6]	
田中弥生 (国語研・東京大:学生)		69
『UniDic』と『分類語彙表』の見出し対応表データの構築	[P-A-7]	
近藤明日子 (国語研)・田中牧郎 (明治大)		79
『名大会話コーパス』の比較に基づく教室談話における「中途終了型発話」の特徴	[P-A-8]	
矢田真菜 (東京学芸大:学生)		87
『多言語母語の日本語学習者横断コーパス』の母語話者データにおけるタスクと産出語彙の関連	[P-B-1]	
小西円 (国語研)		95
『現代日本語書き言葉均衡コーパス』に対する分類語彙表番号アノテーションの試行	[P-B-2]	
加藤祥・浅原正幸・山崎誠 (国語研)		104
『日常会話コーパス』プロジェクト—コーパスに基づく話し言葉の多角的研究—	[P-B-3]	
小磯花絵 (国語研)		114
日本語語構成情報データベースの構築	[P-B-4]	
浅尾仁彦 (情報通信研究機構)		120
発話文自動生成のための日本語表現文型辞書の作成	[P-B-5]	
夏目和子 (名古屋大)・刀山将大 (名古屋大:学生)・佐藤理史 (名古屋大)		126
スマホで古辞書 —『篆隸万象名義』のIDS 検索を例に—	[P-B-6]	
劉冠偉・李媛 (北海道大:学生)・池田証壽 (北海道大)		140

機械翻訳用超大規模辞書データ資源	[P-B-7]	
春遍雀來 (日中韓辞典研究所)		148
モンゴル語アクセント研究のためのデータベース	[P-B-8]	
玉栄 (内モンゴル大・国語研)・西川賢哉・前川喜久雄 (国語研)		154
多重の読みを持つテキストのコーパス化	[P-B-9]	
小木曾智信 (国語研)		159
次元形容詞にみる母語話者らしい日本語形容詞の使用	[O-B-1]	
西内沙恵 (国語研・立教大)		163
日本語コーパスの包括的検索環境の実現に向けて	[O-B-2]	
前川喜久雄・浅原正幸・小木曾智信・小磯花絵・木部暢子・迫田久美子 (国語研)		170
機能語用例文データベース『はごろも』の今後の展開	[O-B-3]	
堀恵子 (東洋大・筑波大)・内丸裕佳子 (岡山大)・加藤恵梨 (朝日大)・小西円・山崎誠 (国語研)・江田すみれ (日本女子大)・建石始 (神戸女学院大)・中俣尚己 (京都教育大)・ 李在鎬 (早稲田大)		180
日本語学習者コーパスの教育応用における留意点—『多言語母語の日本語学習者横断コーパ ス』に見る母語話者 L1 産出データの安定性検証を中心に—	[O-B-4]	
石川慎一郎 (神戸大)		190
漢語の仮名表記—実態と背景—	[O-C-1]	
間淵洋子 (明治大:学生・学振)		201
『日本語歴史コーパス』短単位アノテーション作業効率化に向けた形態素解析用辞書『UniDic』 の段階的特殊化の検討—近松コーパスを例として—	[O-C-2]	
岡照晃 (国語研)		214
全文検索システム『ひまわり』における言語分析支援機能の拡張	[P-C-1]	
山口昌也 (国語研)		226
児童生徒の「手」作文に於ける経年変化の計量的分析	[P-C-2]	
阿部藤子 (東京家政大)・今田水穂 (文部科学省)・宗我部義則 (お茶の水女子大付属 中)・富士原紀絵 (お茶の水女子大)・松崎史周 (日本女子体育大)・宮城信 (富山大)		234
『日本語日常会話コーパス』構築における会話収録方法と進捗状況	[P-C-3]	
田中弥生 (国語研・東京大:学生)・柏野和佳子・角田ゆかり (国語研)・伝康晴 (千葉 大)・小磯花絵 (国語研)		248
『分類語彙表』の類義語と分散表現を利用した all-words 語義曖昧性解消	[P-C-4]	
鈴木類 (茨城大:学生)・古宮嘉那子 (茨城大)・浅原正幸 (国語研)・佐々木稔・新納浩幸 (茨城大)		258
形態素解析ソフトウェア『Web 茶まめ』の改良と Web API の試作	[P-C-5]	
川口寛治・薦田龍輝 (東京電機大:学生)・堤智昭 (東京電機大)		265
『現代日本語書き言葉均衡コーパス』を用いた「～ていく」「～てくる」構文の意味分析	[P-C-6]	
加藤麟太郎 (東京大:学生)・藤井聖子 (東京大)		273

明治初期教科書『物理階梯』のコーパス作成による語彙の考察	[P-C-7]	
田中牧郎(明治大)・島田むつみ・高橋雄太(明治大:学生)		282
話し言葉コーパスの転記タグ:『多言語母語の日本語学習者横断コーパス』と『日本語話し言葉コーパス』の比較	[P-C-8]	
西川賢哉(国語研)		288
『日本語日常会話コーパス』の転記基準と作業工程	[P-C-9]	
川端良子(国語研・千葉大:学生)・白田泰如・西川賢哉(国語研)・徳永弘子(国語研・東京電機大)・小磯花絵(国語研)		296
『現代日本語書き言葉均衡コーパス』と『分類語彙表』を利用した漢字3文字略熟語の抽出	[P-D-1]	
山崎誠(国語研)		307
名詞項構造付与データの構築	[P-D-2]	
竹内孔一(岡山大)		317
『名大会話コーパス』中納言版・ひまわり版公開データの作成	[P-D-3]	
柏野和佳子・西川賢哉・小磯花絵(国語研)		324
『現代日本語書き言葉均衡コーパス』に対する節の意味分類情報アノテーション—基準策定, 仕様書作成の必要性について—	[P-D-4]	
松本理美(立命館大:学生)・浅原正幸(国語研)・有田節子(立命館大)		336
『日本語話し言葉コーパス』における発声様式の自動分類	[P-D-5]	
森大毅・藤本雅子(国語研)・浅井拓也・前川喜久雄(国語研)		347
近代文語文の通時的变化の分析 —語種率・品詞率に着目して—	[P-D-6]	
近藤明日子(国語研)		355
結合の強度を測る指標としての Log-r の有用性:日・英語のバイグラムデータに基づく MI, LLR などとの比較	[P-D-7]	
藤村逸子(名古屋大)・青木繁伸(群馬大名誉教授)		364
語彙・文型調査を目的とした『幼稚園の配布文書コーパス』の作成	[P-D-8]	
長谷川守寿(首都大)・西尾広美(国語研)		377
固有表現抽出におけるアノテーション手法の比較	[P-D-9]	
鈴木雅也(茨城大:学生)・古宮嘉那子(茨城大)・岩倉友哉(富士通研)・佐々木稔・新納浩幸(茨城大)		385
『現代日本語書き言葉均衡コーパス』への情報構造アノテーションの分析	[O-D-1]	
宮内拓也(国語研・東京外大:学生)・浅原正幸(国語研)・中川奈津子(千葉大・学振)・加藤祥(国語研)		404
読み時間と情報構造について(ちょっとながめ)	[O-D-2]	
浅原正幸(国語研)		416

発表論文