



Language Resources Workshop

言語資源活用 ワークショップ

2016

クロージング

主催 国立国語研究所 コーパス開発センター
2017年3月7日-3月8日

通常

運営

国立国語研究所 コーパス開発センター

センター長 前川喜久雄

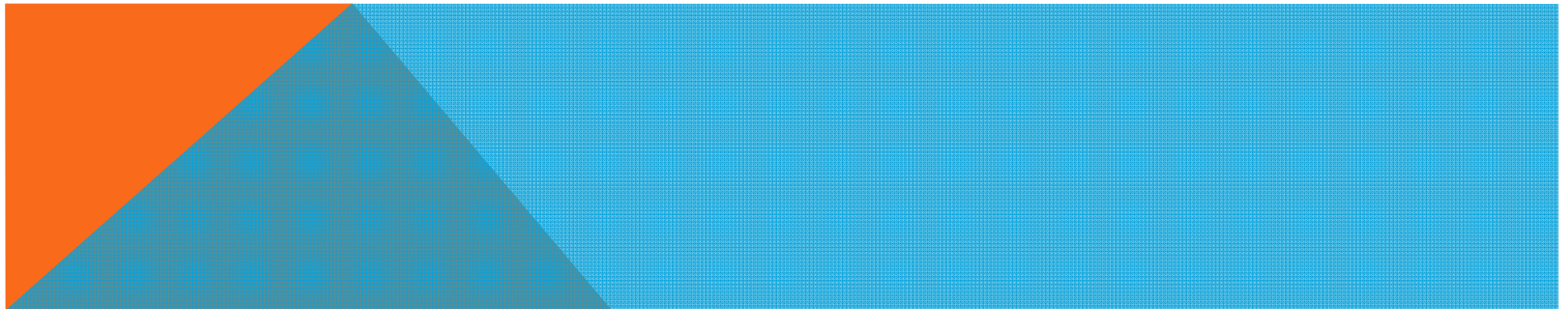
専任准教授 浅原正幸

特任助教 石本佑一 岡照晃

研究員 加藤祥 近藤明日子 西川賢哉

DBエンジニア 中村壮範

併任職員 小木曾智信 柏野和佳子 小磯花絵 山口昌也 山崎誠



運営→指定討論の廃止

- 指定討論

ワークショップの最後に数人の研究者が各発表の概要を示す

- 代わりにビジネスミーティング形式のクロージング

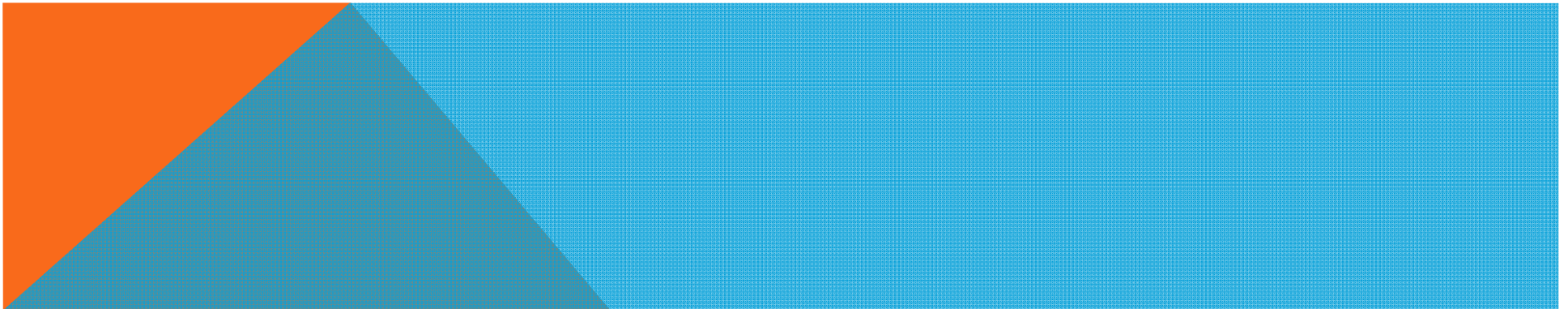
ワークショップの運営に関する議論

これから各種情報を示しますので

フロアからコメントをおねがいします

もしくはメールでご連絡ください：

lrw@ninjal.ac.jp



運営→日程

2016年11月1日(火) 発表応募開始

2016年12月26日(月) 12:00 発表応募締切

→ 2017年1月5日(木) 12:00 発表応募締切

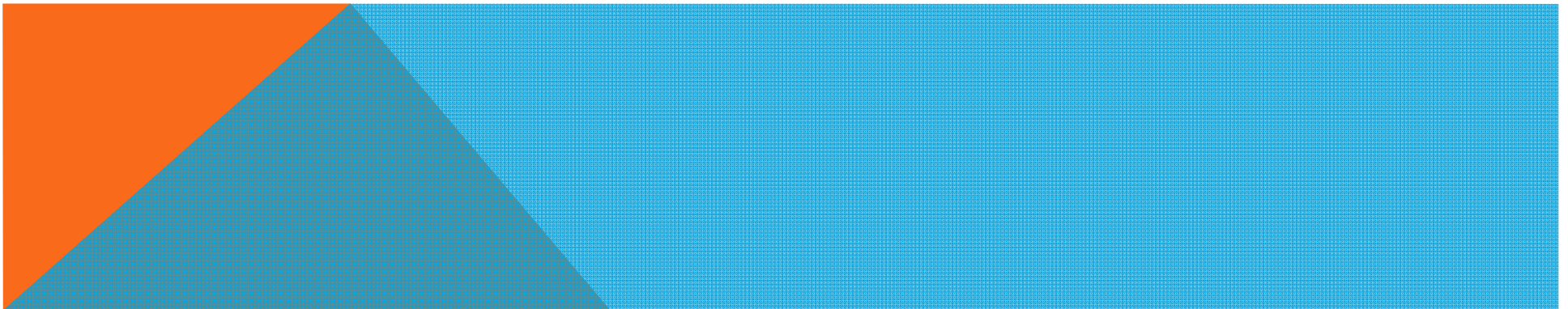
2017年 1月 6日(金) 採否の通知

→ 2017年1月12日(木) 採否の通知

2017年 2月 3日(金) 17:00 (JST) 予稿集掲載原稿締切

→2017年2月6日(月) 9:00 (JST)

2017年 3月 7 日(火), 8 日(水) ワークショップ開催



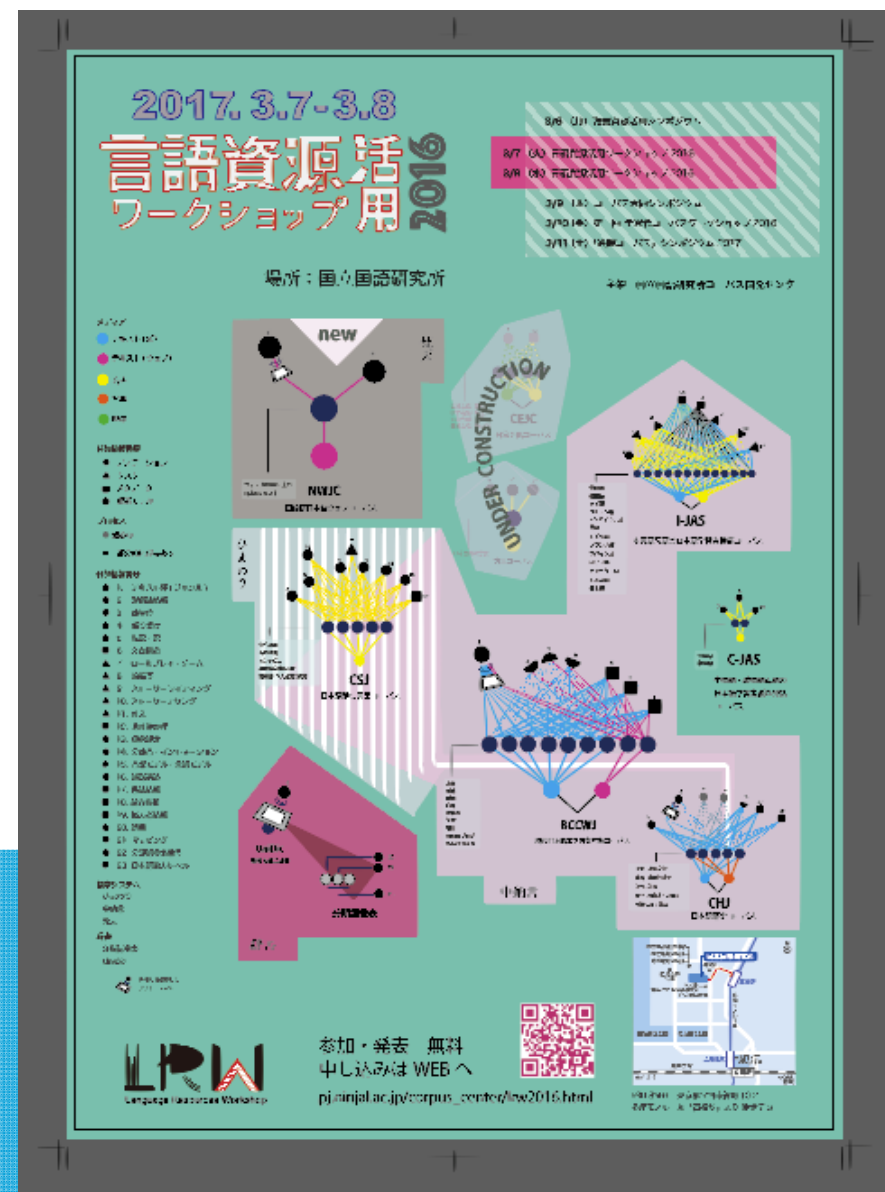
運営→ポスターデザイン

インフォグラフィック

情報の可視化

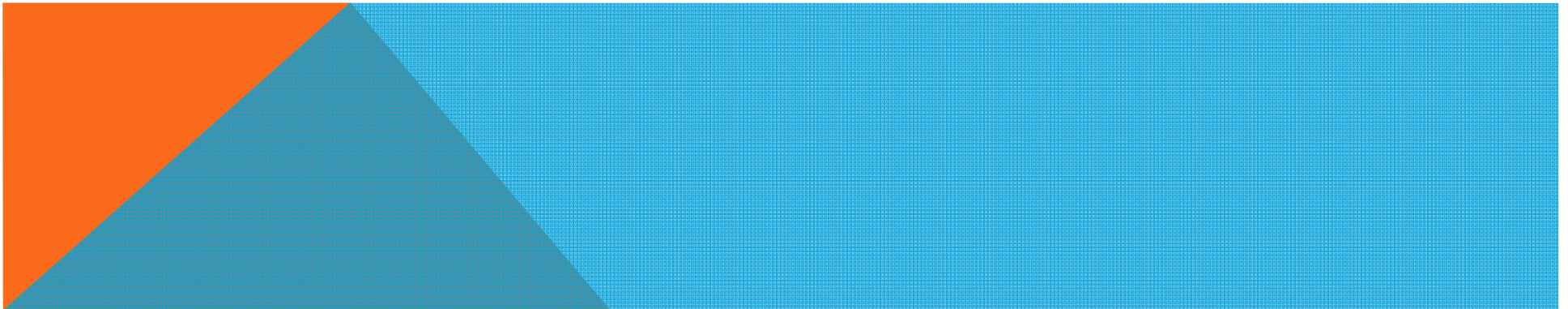
国語研で開発される
コーパスの成長過程を
図示したもの

5年間でこの図がどのように
成長していくか



運営→書店出店

- 朝倉書店 3/6(月)-3/8(水)
- くろしお出版 3/7(火), 3/9 (木), 3/10 (金)
- 勉誠出版 3/6(月)-3/8(水)
- ひつじ書房 3/6(月)-3/8(水)

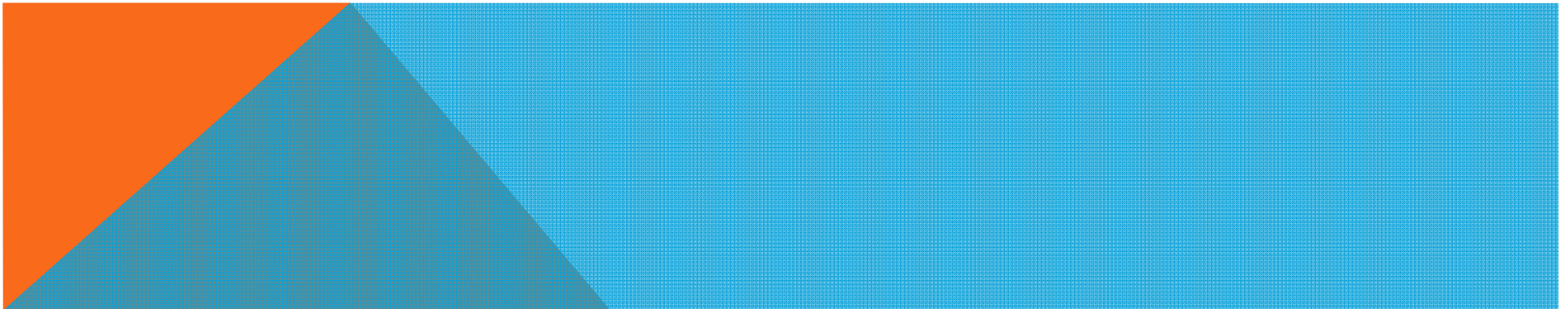


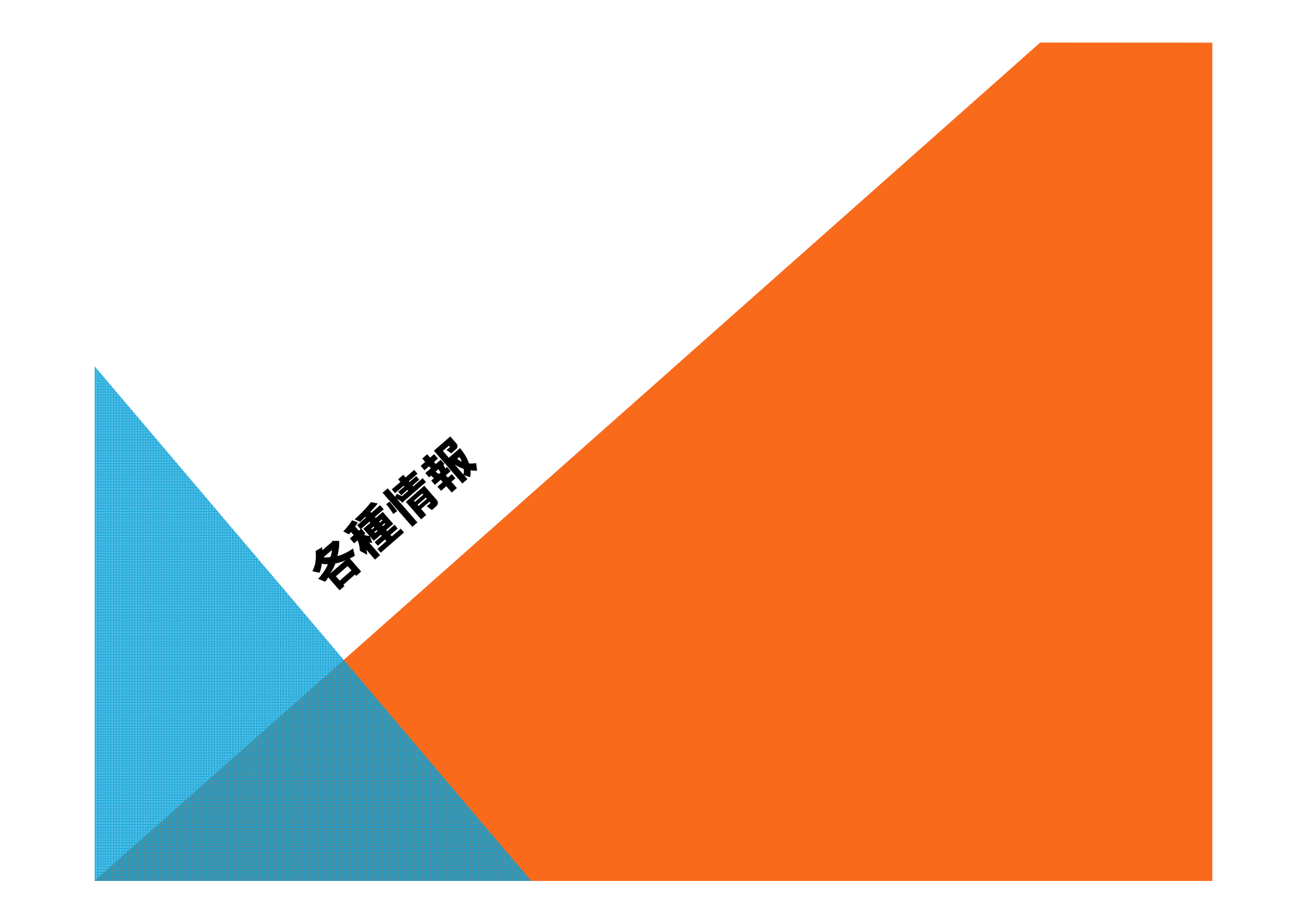
運営上のおねがい

連絡先は

lrw@ninja.ac.jp をお願いします。

添付ファイル自動保存 script などさまざまなしかけがほどこしてあります
特任助教も登録されています



The background features a large orange polygon on the right side, which is a rectangle with its top-left corner cut off by a diagonal line. On the left side, there are two overlapping triangles: a light blue one on top and a darker blue one below it, both with a fine grid pattern. The text '各種情報' is positioned in the white space between these blue shapes and the orange shape.

各種情報

参加者数

語彙資源活用シンポジウム

事前申込者数：89名（2/27 現在）

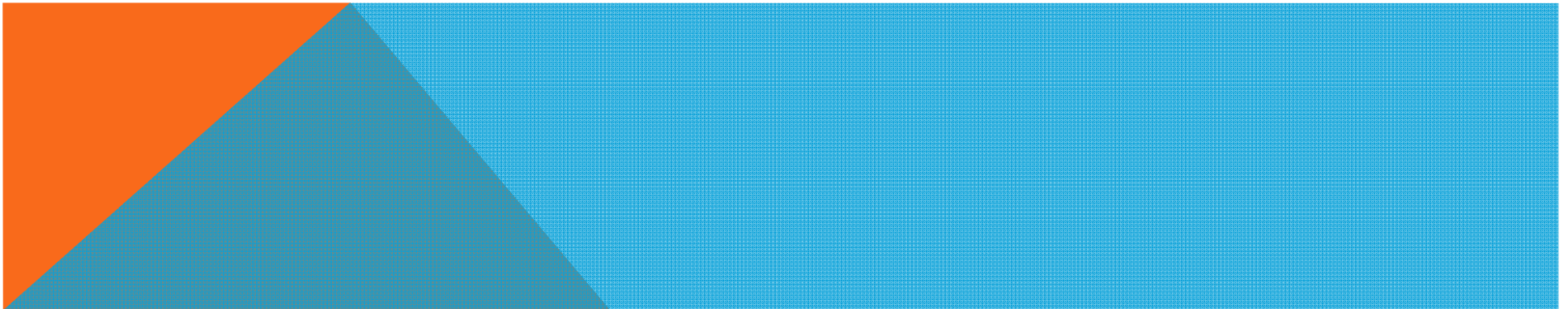
実参加者数：（3/8 15:00 現在）
100名

言語資源活用ワークショップ

事前申込者数：108名（2/27 現在）

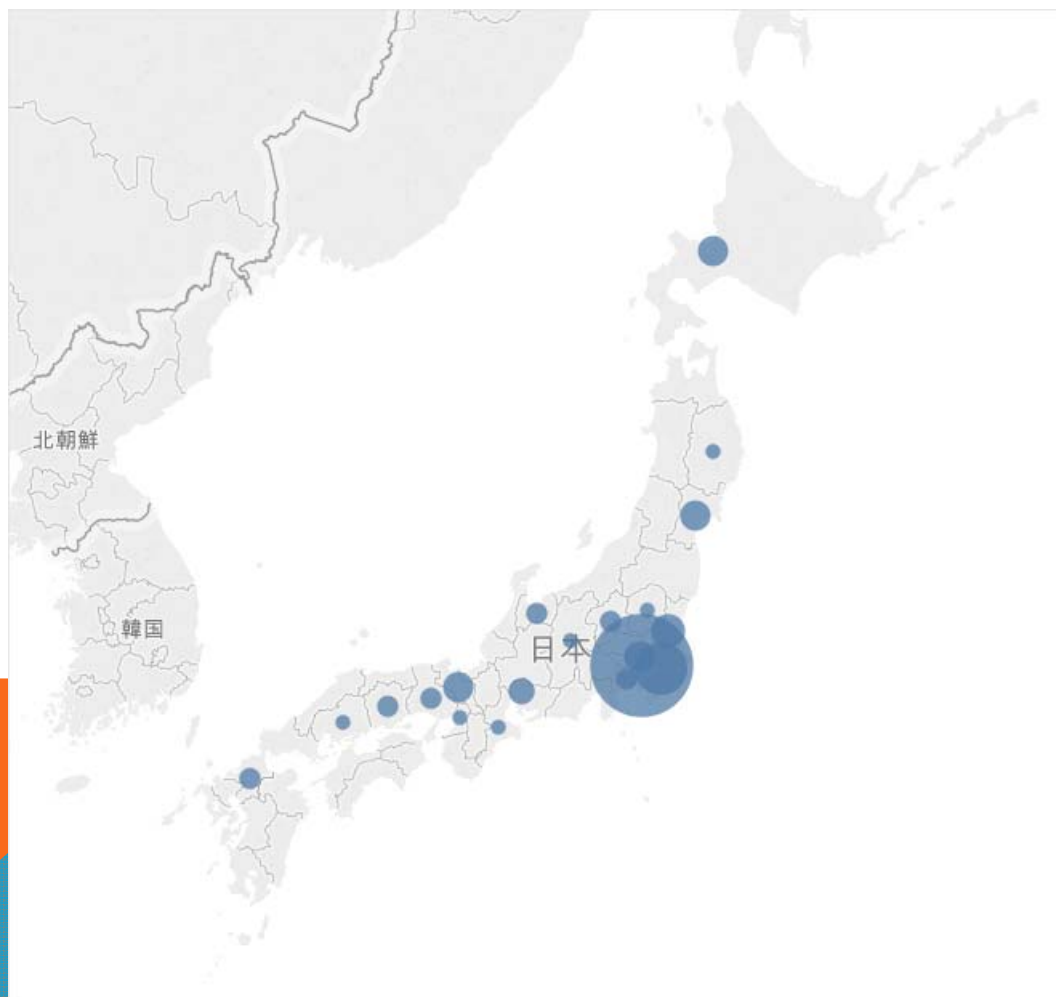
懇親会参加者 35名

実参加者数：（3/8 15:00 現在）
160名



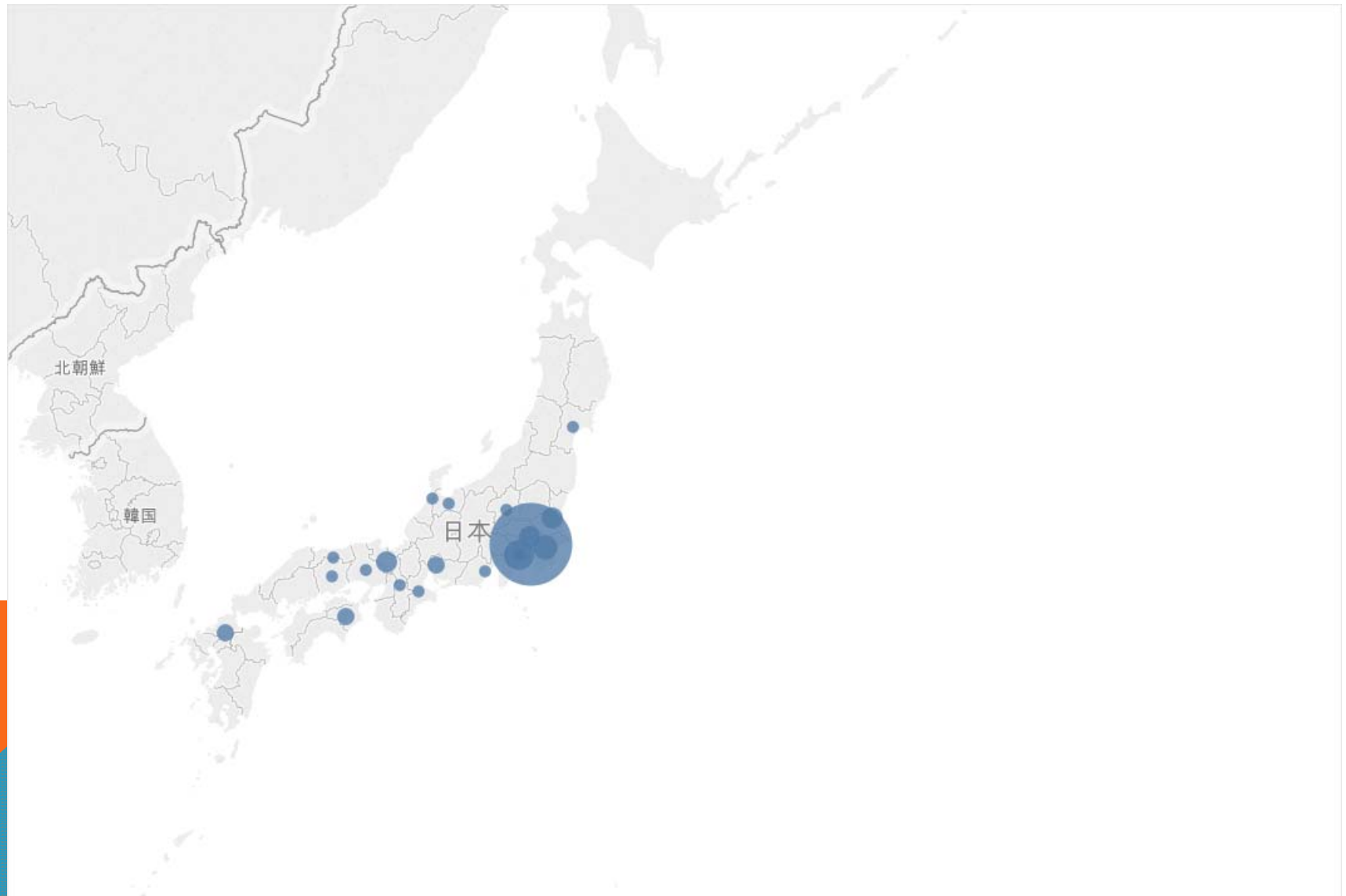
参加者数→在住地

言語資源活用ワークショップ 事前申込者数：108名（2/27 現在）



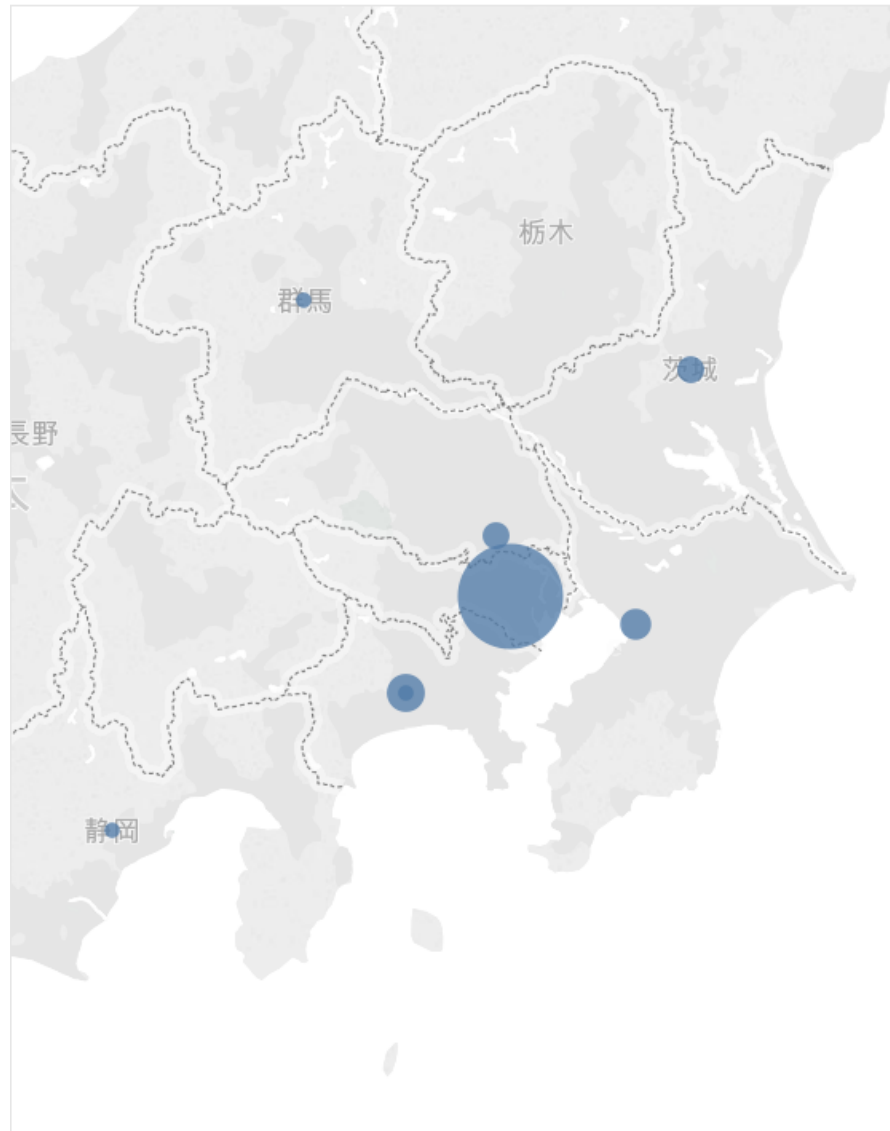
参加者数→在住地

語彙資源活用シンポジウム 事前申込者数：89名（2/27 現在）

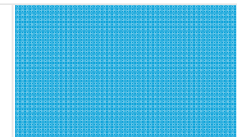
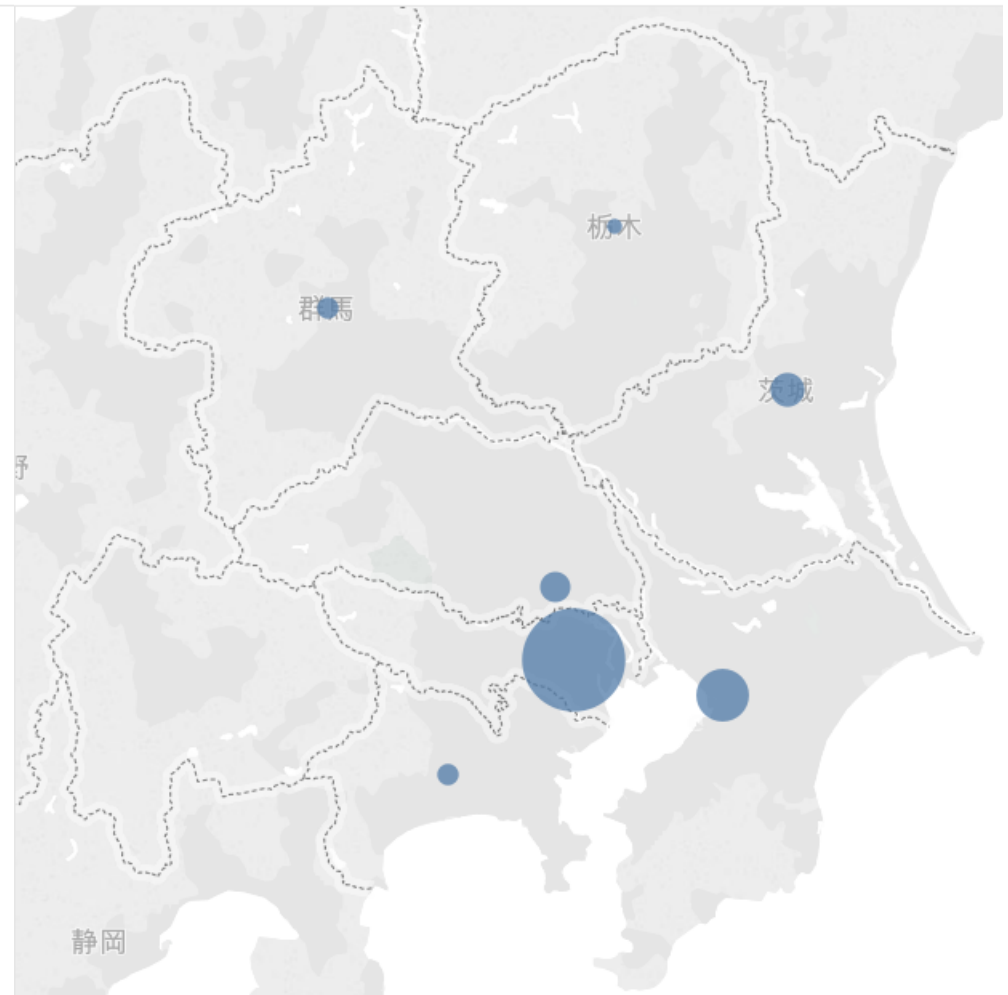


参加者数→在住地（首都圏）

語彙資源活用シンポジウム(3/6) 参加者分布



言語資源活用ワークショップ(3/7-8) 参加者分布

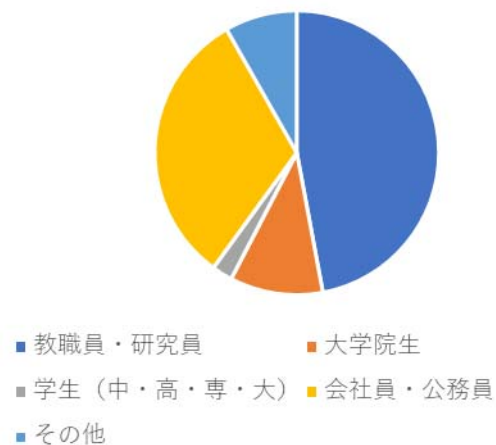


参加者数→職種

語彙資源活用シンポジウム

事前申込者数：89名（2/27 現在）

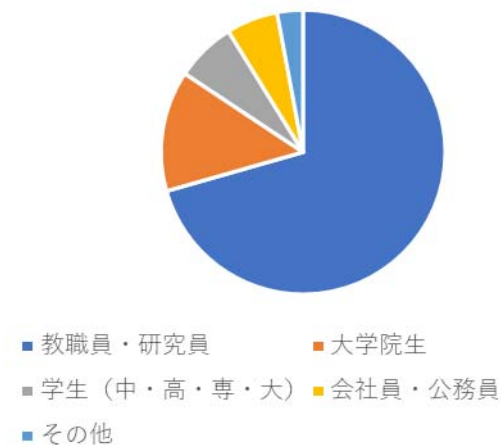
語彙資源シンポジウム(3/6)



言語資源活用ワークショップ

事前申込者数：108名（2/27 現在）

言語資源ワークショップ(3/7,8)



発表件数

3/6 語彙資源シンポジウム：

招待講演 8 件

+ パネルセッション

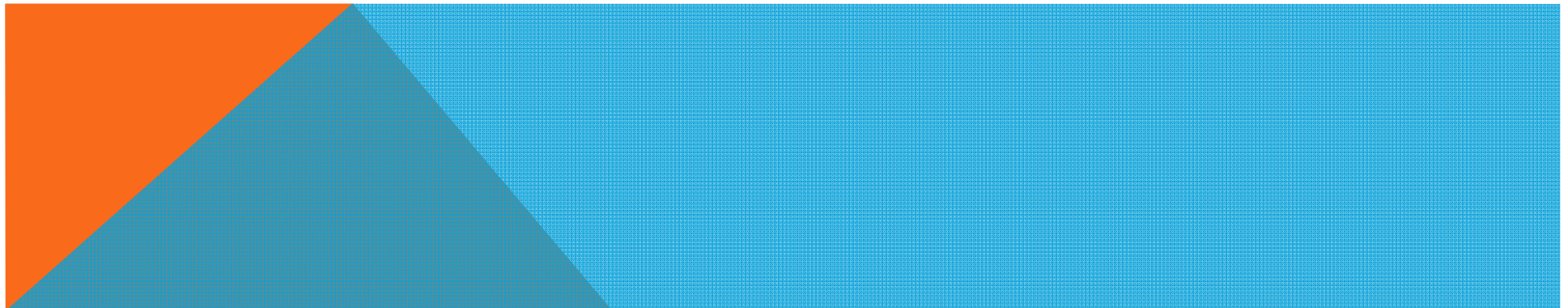
3/7-8 言語資源ワークショップ：

総発表件数 48件（学生発表 14件）

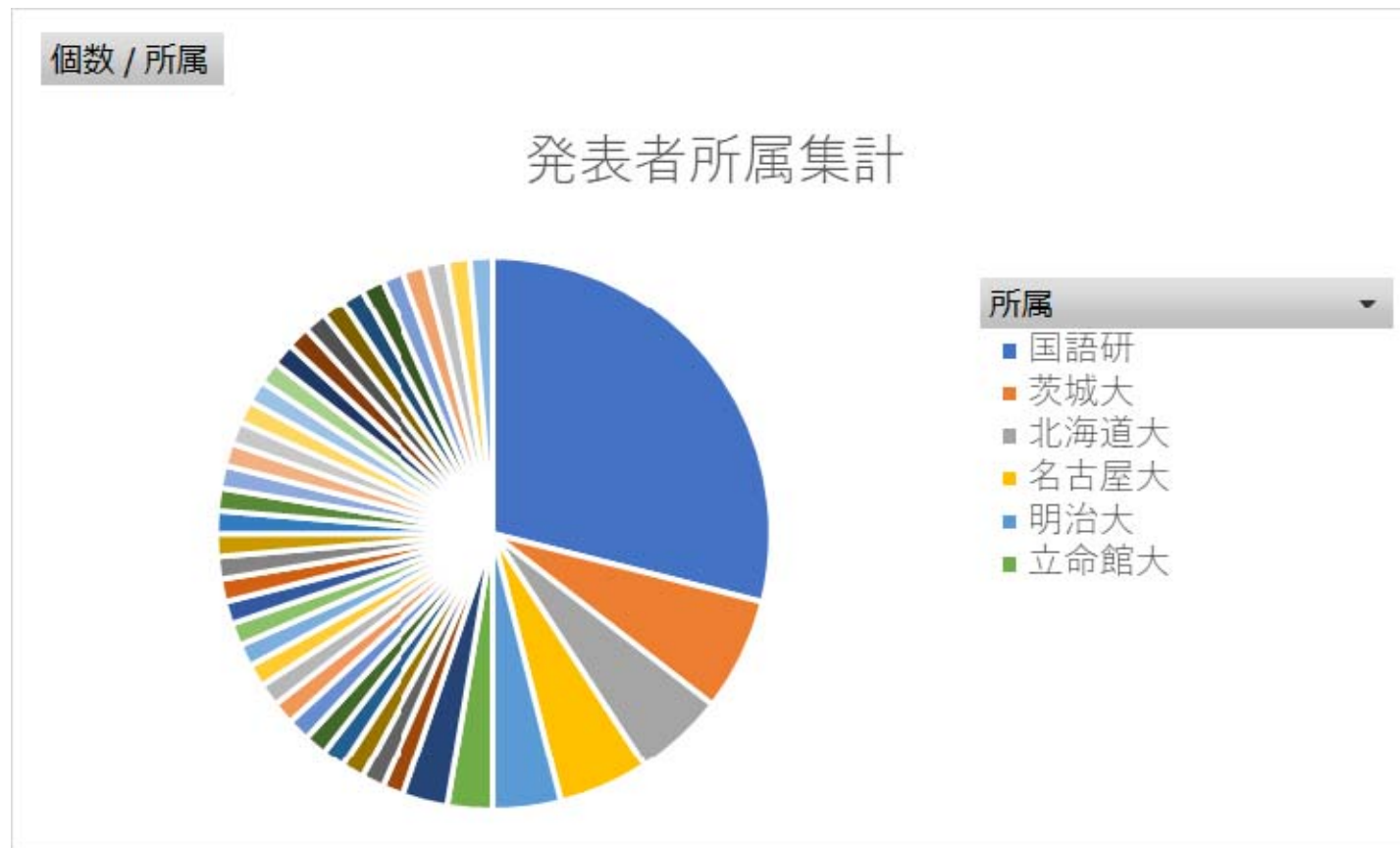
口頭発表件数 12件（招待講演 2件, 学生発表4件）

ポスター発表件数 35件（学生発表10件）

デモンストレーション 1件

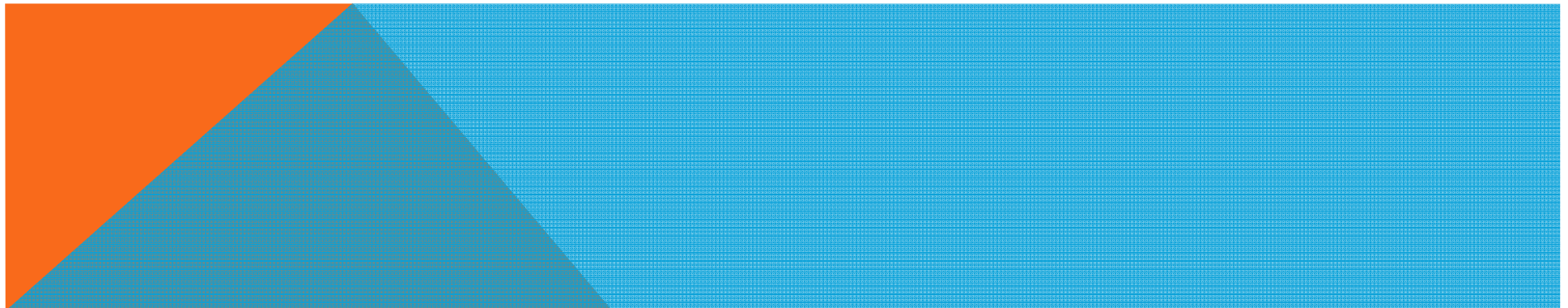


発表件数→発表者の所属



発表件数→利用言語資源（頻度3以上）

利用言語資源	件数
『現代日本語書き言葉均衡コーパス(BCCWJ)』	13
『UniDic』	9
『日本語歴史コーパス(CHJ)』	5
『日本語日常会話コーパス(CEJC)』	4
『多言語母語の日本語学習者横断コーパス(I-JAS)』	4
『名大会話コーパス』	3
『分類語彙表』	3
『日本語話し言葉コーパス(CSJ)』	3



The background features a large orange triangle on the right side, pointing towards the top right. On the left side, there are two overlapping triangles: a light blue one on top and a darker blue one below it, both pointing towards the bottom left. The text '発表論文集' is positioned in the white space between these two blue triangles.

発表論文集

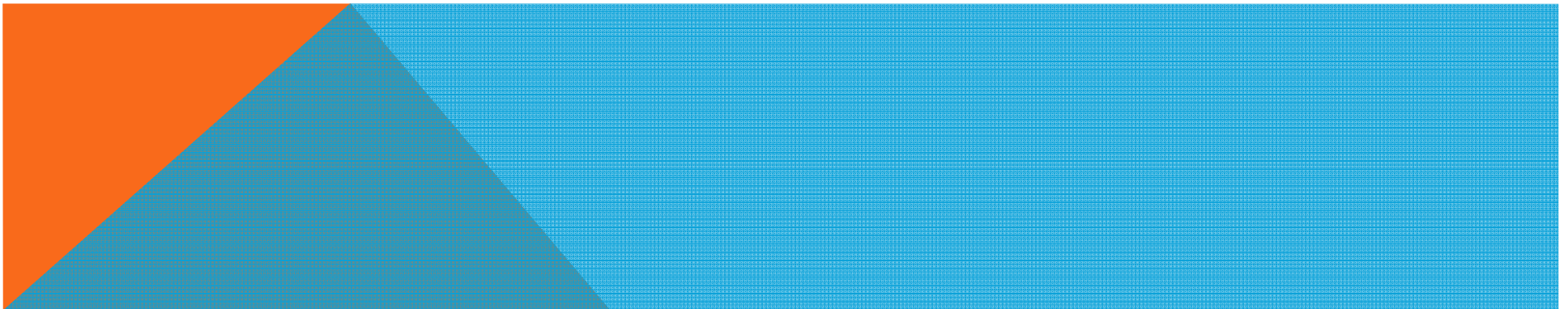
発表論文集

http://pj.ninjal.ac.jp/corpus_center/lrw/LRW2016_proceedings.pdf

国立国語研究所学術情報リポジトリにて公開すべく調整中

<https://repository.ninjal.ac.jp/>

(DOI を付与)



今後の予定

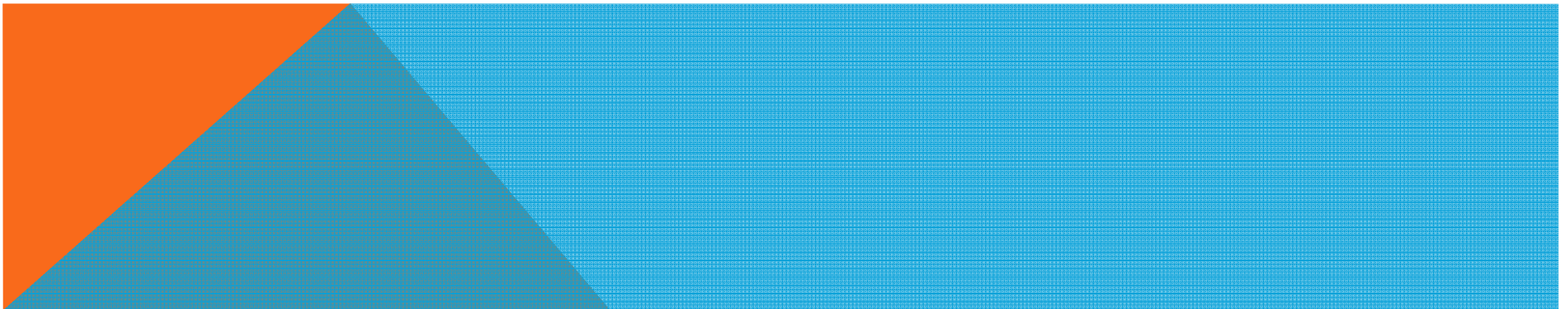
第2回 2017年9月5日(火)・6日(水)

第3回 2018年9月

第4回 2019年9月

第5回 2020年9月

第6回 2021年9月



今後の予定→ 第2回

2017年9月3日(日) 国語研研究系音声言語研究領域 経年変化班 国際シンポ (仮)

2017年9月4日(月) 国語研研究系音声言語研究領域 経年変化班 国際シンポ

2017年9月5日(火) 言語資源活用ワークショップ 2017

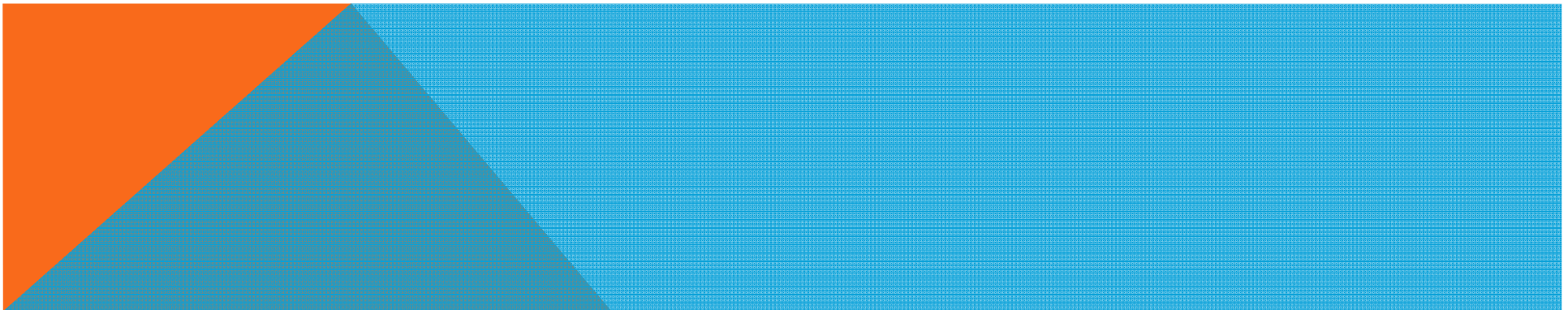
2017年9月6日(水) 言語資源活用ワークショップ 2017

2017年9月7日(木) コーパス開発センター主催シンポジウム (仮)

2017年9月8日(金) コーパス合同シンポジウム

2017年9月9日(土) 第3回学習者コーパス・ワークショップ (I-JAS)

2017年9月25日(月)

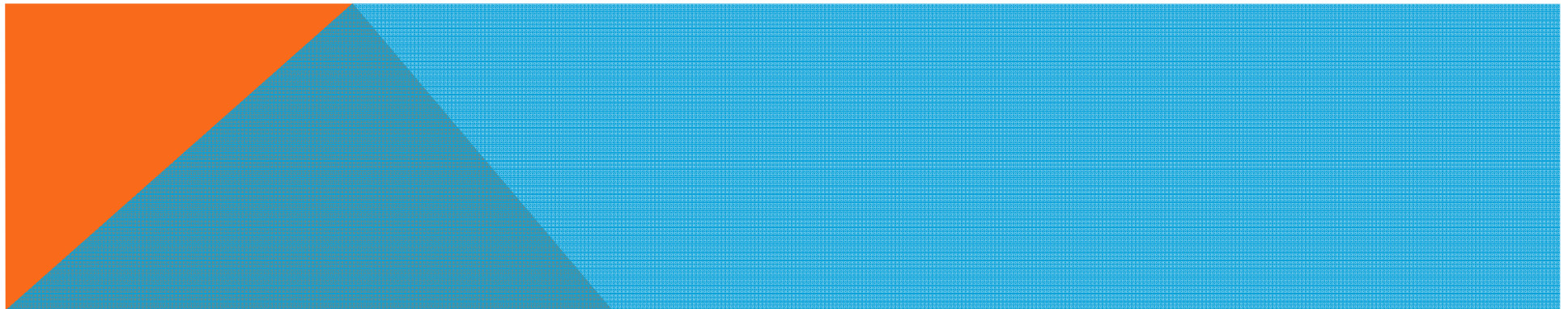


今後の予定→ 第2回 ワークショップ

2017年9月5日(火) 言語資源活用ワークショップ 2017

2017年9月6日(水) 言語資源活用ワークショップ 2017

招待講演者： 東京大学 美馬秀樹先生
あとひとり さきほどひとり依頼しました



今後の予定→ 第2回→併設シンポ

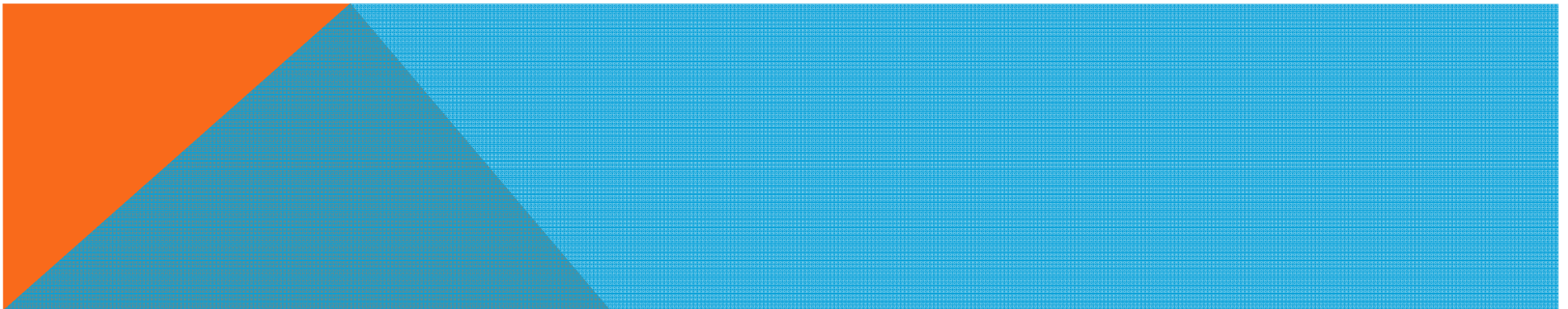
2017年9月7日(木) コーパス開発センター主催シンポジウム (仮)

- ネタ切れ?
 - 語彙資源 関連(3/6 実施)
 - Universal Dependency 関連 (3/13 言語処理学会チュートリアル)
 - 分類語彙表 & all-word WSD 関連 (3/14 言語処理学会テーマセッション)
- 案
 - N2I (NII-NINJAL-ISM)
 - NII CODH (人文学オープンデータ共同利用センター)
 - L1コーパス関連
 - ウェブコーパス関連
 - 読み時間 心理言語関連
 - 助教2人にまかせる

一般企業の方が
参加可能な内容
がよい？

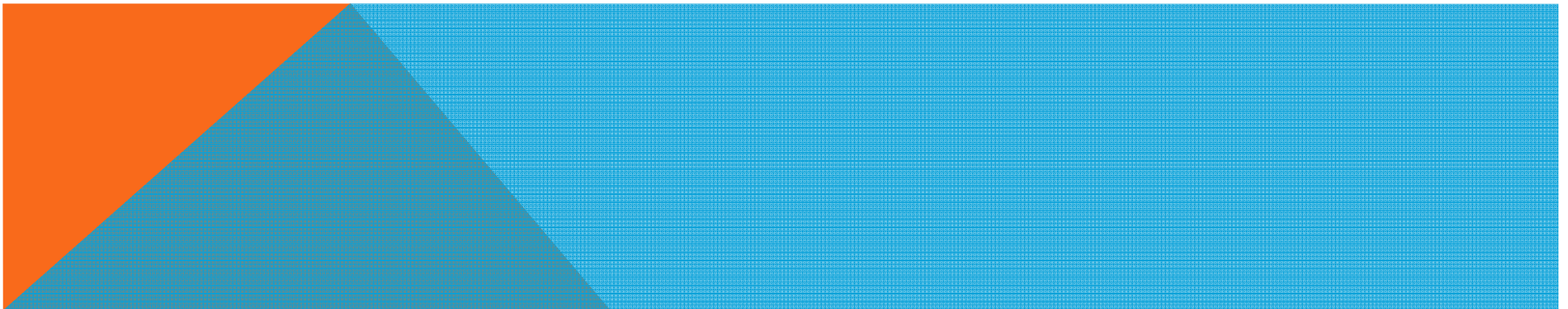
今後の予定→ 第2回→講習会案

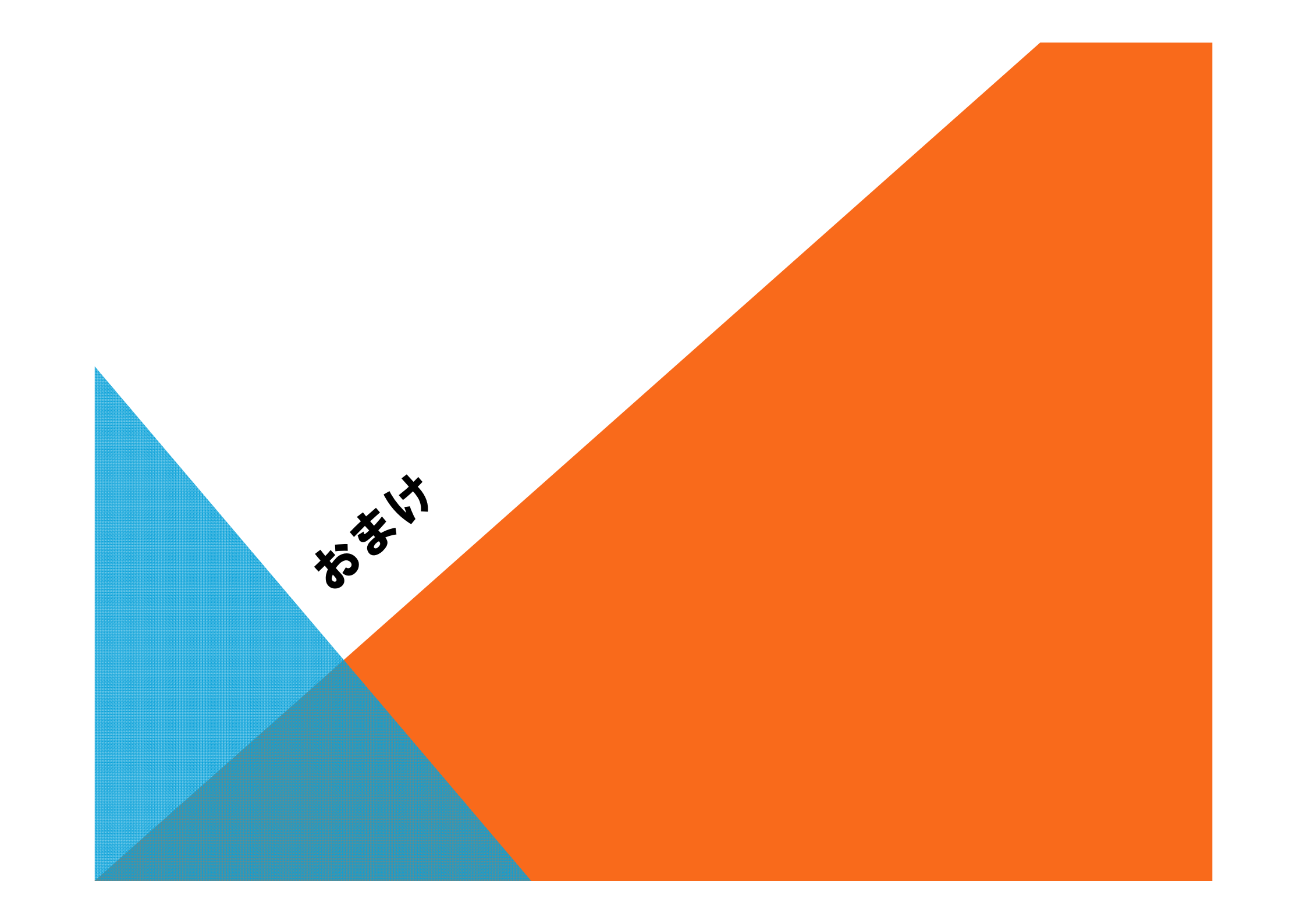
- praat 講習会
- 「梵天」→「ChaKi.NET」講習会
- プログラミング講習会
 - python
 - ruby
 - R
 - XML
- 視線走査装置 EyeLink 講習会



今後の予定→ 第2回→輪読会

- Corpus Linguistics Conference 2017
 - 論文を読む
 - Summer School を参考にする
- 統計処理関連
 - Baayen 本：Analyzing Linguistic Data: A Practical Introduction to Statistics using R





おまけ

「単独」

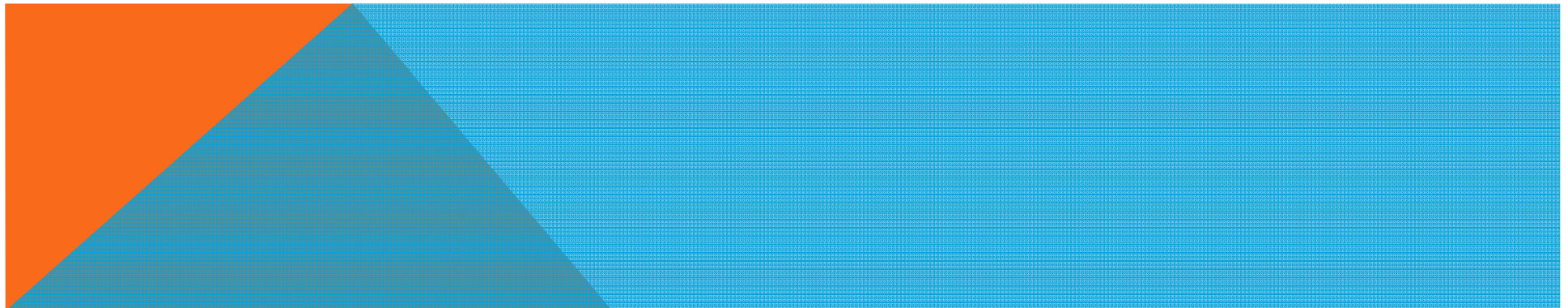
「数独」

= 「数字は独立に限る」(ニコリ社の商標)

一般名称：「ナンバープレース」
の単語(語彙素)バージョン

縦・横の各列及び、太線で囲まれた3×3のブロック内に同じ品詞の単語が複数入ってはいけない

彼の		逆も	弄る	れる			然う 然う	狸
まで	私				狐			或る
だ	否	猫	色んな		憂鬱			頑張る
		より	其れ		らしい	今晚は	黒ずむ	為
為る				仮令			だけ	
新しい	只管		ほど					せる
		同じ		遡る	やら		固より	其処
		無い	川獺		うえい	所謂		
	を	ない			漸く			御早う



「単独」

配置する品詞9種：UniDic 品詞体系（用法主義）

名詞（名詞-普通名詞-副詞可能, 名詞-普通名詞-形状詞可能を含む）

動詞

形容詞（名詞-普通名詞-形状詞可能を含む）

副詞（名詞-普通名詞-副詞可能を含む）

連体詞

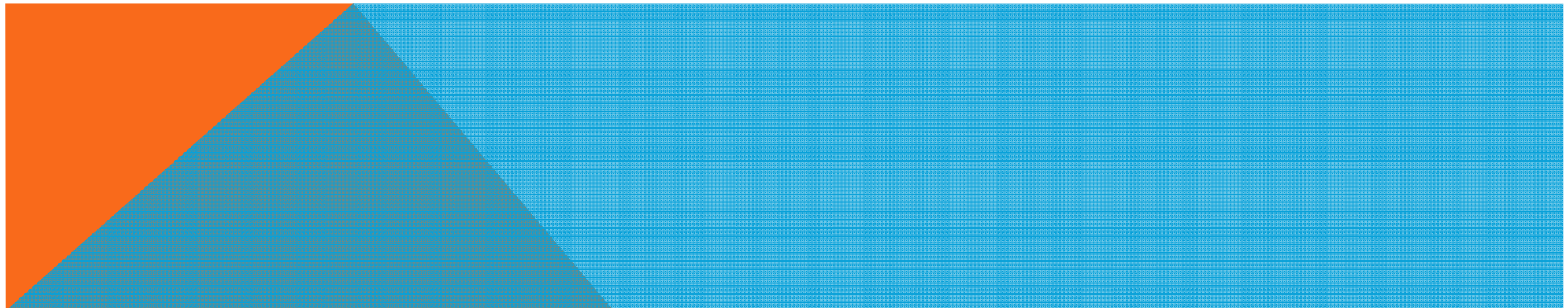
感動詞

代名詞

助動詞

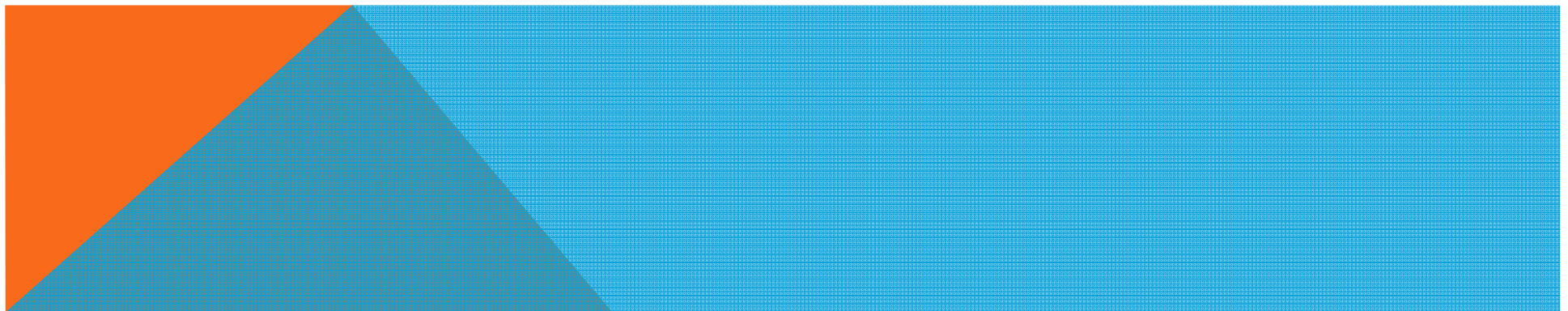
助詞

彼の		逆も	弄る	れる			然う 然う	狸
まで	私				狐			或る
だ	否	猫	色んな		憂鬱			頑張る
		より	其れ		らしい	今晚は	黒ずむ	為
為る				仮令			だけ	
新しい	只管		ほど					せる
		同じ		遡る	やら		固より	其処
		無い	川獺		うえい	所謂		
	を	ない			漸く			御早う



回答（１）

彼の	形容詞	迎も	弄る	れる	代名詞	助詞	然う 然う	狸
まで	私	動詞	副詞	感動詞	狐	助動詞	形容詞	或る
だ	否	猫	色んな	助詞	憂鬱	副詞	代名詞	頑張る
名詞	連体詞	より	其れ	形容詞	らしい	今晚は	黒ずむ	為
為る	助動詞	代名詞	感動詞	仮令	連体詞	名詞	だけ	形容詞
新しい	只管	感動詞	ほど	名詞	動詞	代名詞	連体詞	せる
感動詞	名詞	同じ	助動詞	遡る	やら	形容詞	固より	其処
副詞	動詞	無い	川獺	代名詞	うえい	所謂	助動詞	助詞
代名詞	を	ない	形容詞	連体詞	漸く	動詞	名詞	御早う



回答（2）

連体詞	形容詞	副詞	動詞	助動詞	代名詞	助詞	感動詞	名詞
助詞	代名詞	動詞	副詞	感動詞	名詞	助動詞	形容詞	連体詞
助動詞	感動詞	名詞	連体詞	助詞	形容詞	副詞	代名詞	動詞
名詞	連体詞	助詞	代名詞	形容詞	助動詞	感動詞	動詞	副詞
動詞	助動詞	代名詞	感動詞	副詞	連体詞	名詞	助詞	形容詞
形容詞	副詞	感動詞	助詞	名詞	動詞	代名詞	連体詞	助動詞
感動詞	名詞	連体詞	助動詞	動詞	助詞	形容詞	副詞	代名詞
副詞	動詞	形容詞	名詞	代名詞	感動詞	連体詞	助動詞	助詞
代名詞	助詞	助動詞	形容詞	連体詞	副詞	動詞	名詞	感動詞

