

つかみ・本ネタ・オチから構成される 漫才ロボット台本自動生成手法の提案

甲南大学自然科学研究科知能情報学専攻

真下 遼
梅谷 智弘
北村 達也
灘本 明代

背景



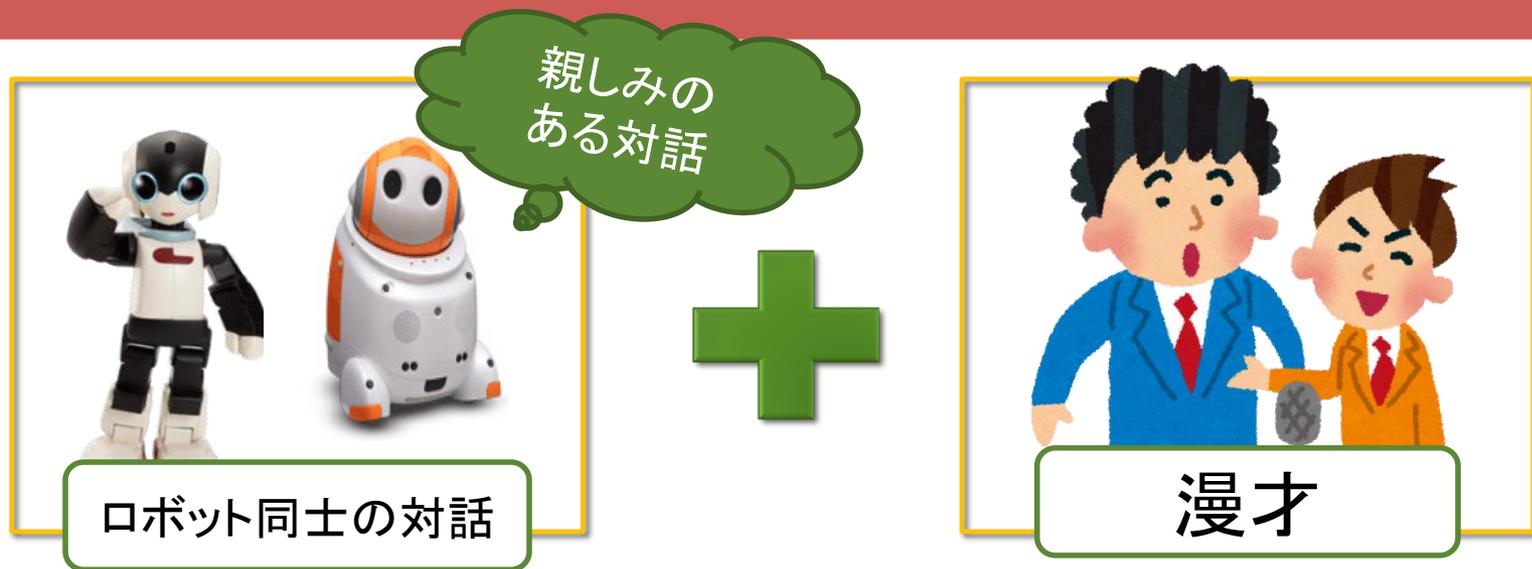
しかし

人がロボットをコミュニケーションの対象とするには抵抗がある

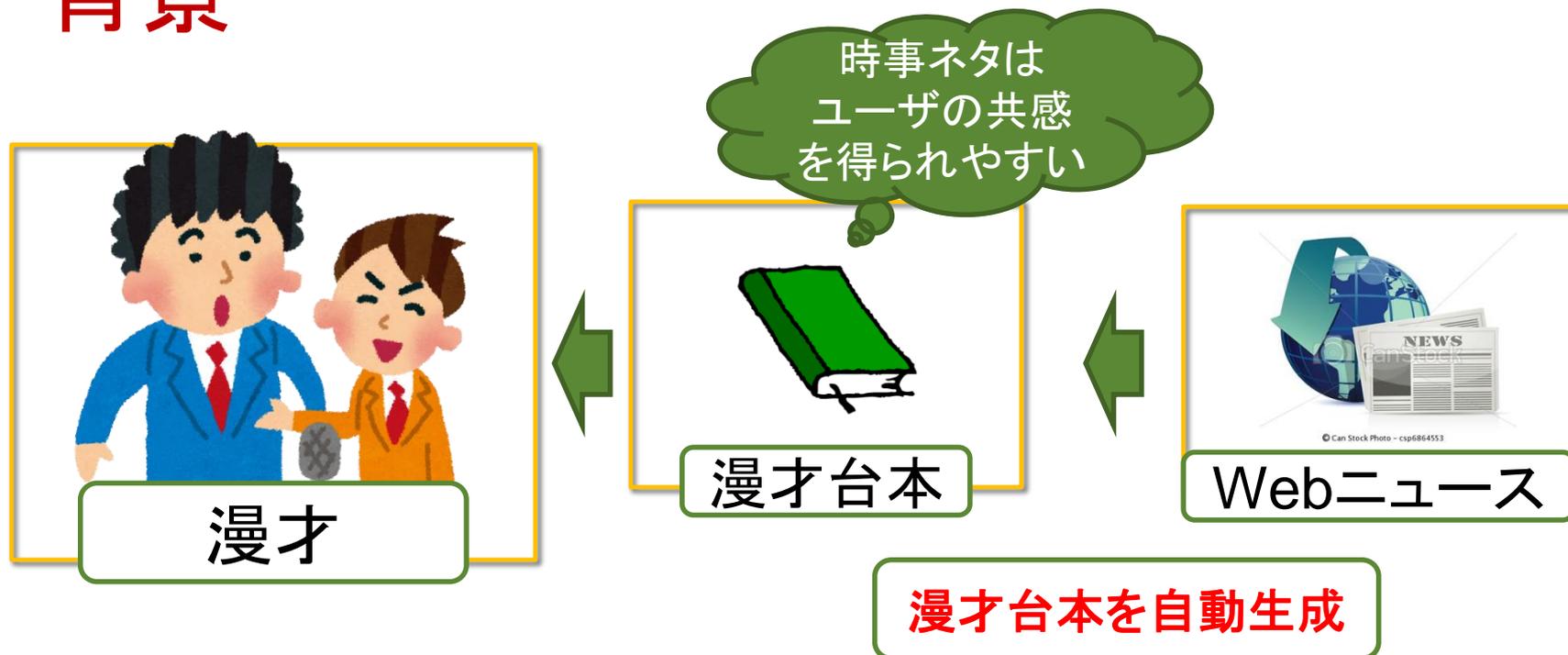
人とロボットとの円滑なコミュニケーションの手助けとなるものが必要

背景

ロボット同士の対話観察を人が行うことで、人とロボットとの自然で円滑なコミュニケーションが可能になる



背景



漫才ロボットによる漫才では、漫才台本を
Webニュースから自動生成して漫才を行う

漫才ロボットの利点

娯楽的なロボット同士の対話観察をユーザに提供し、ロボットに対する親しみや発話理解が向上

漫才の題材に時事ネタを取り入れているのでニュース離れの克服を期待できる

テレビよりもさらに、速報性の高い情報提供が可能

情報視野の拡大



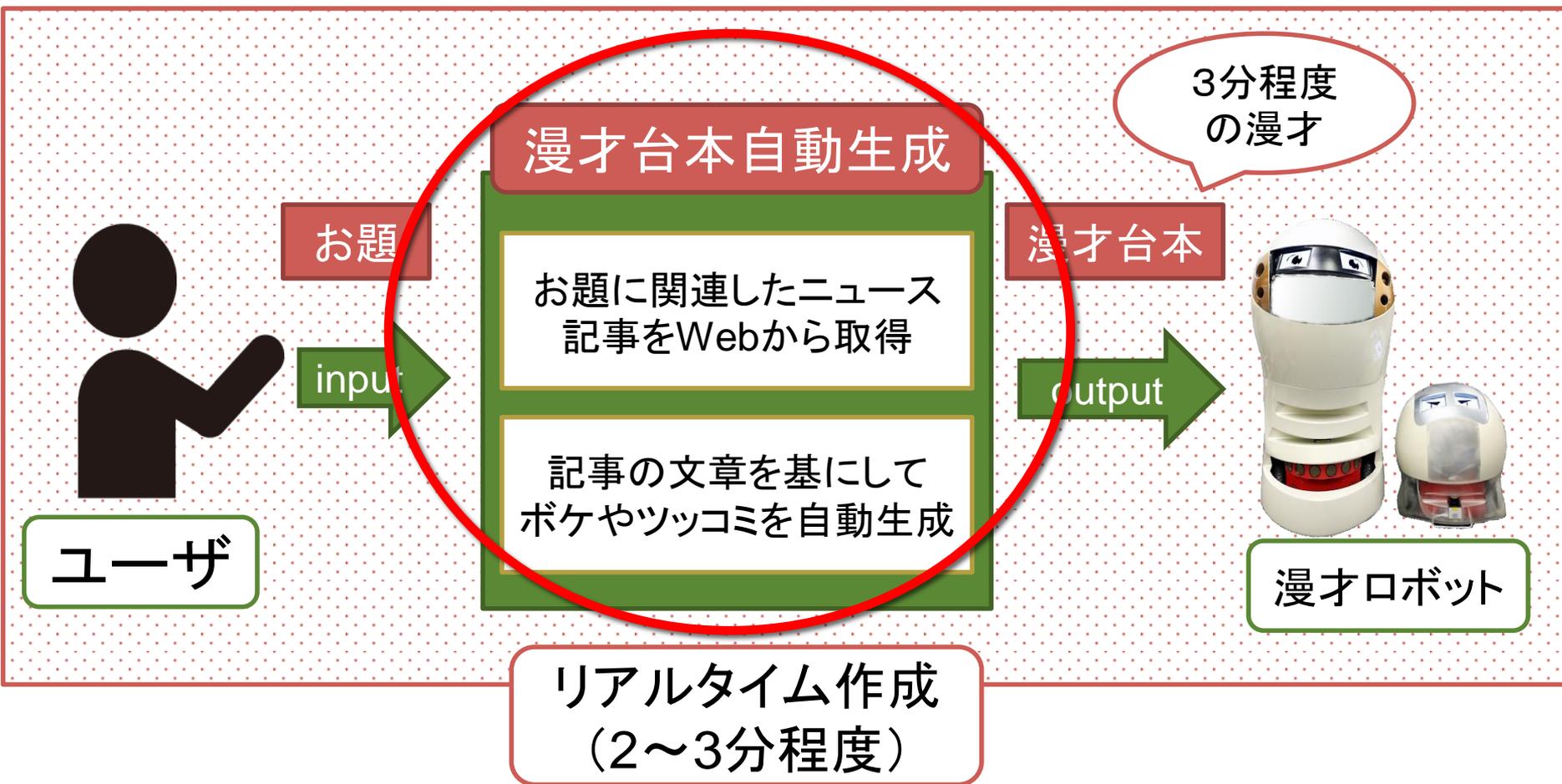
アイちゃん
(ツッコミ)

なんでやねん！

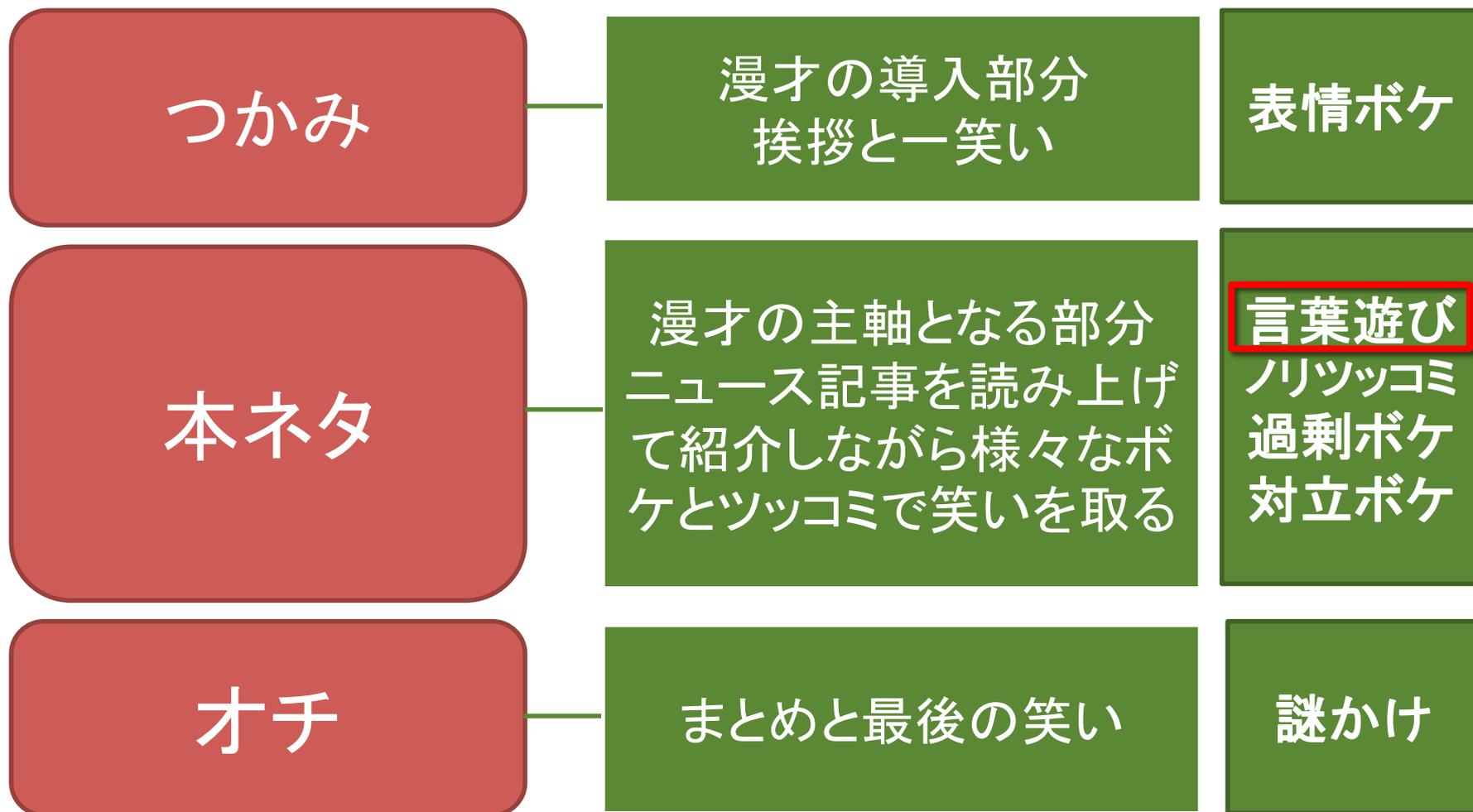


ゴン太
(ボケ)

漫才ロボットの流れ

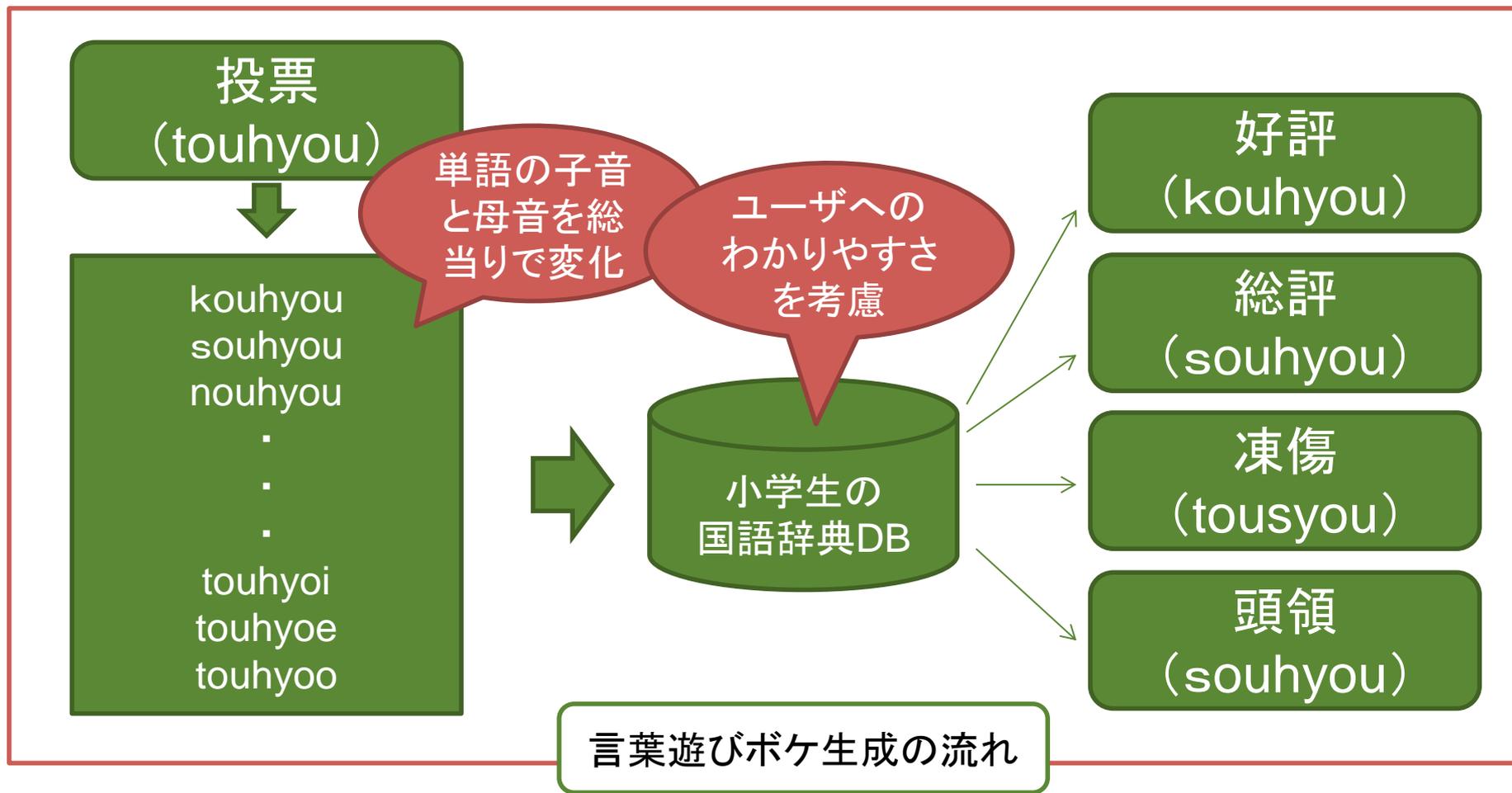


漫才の構成(三段構成)

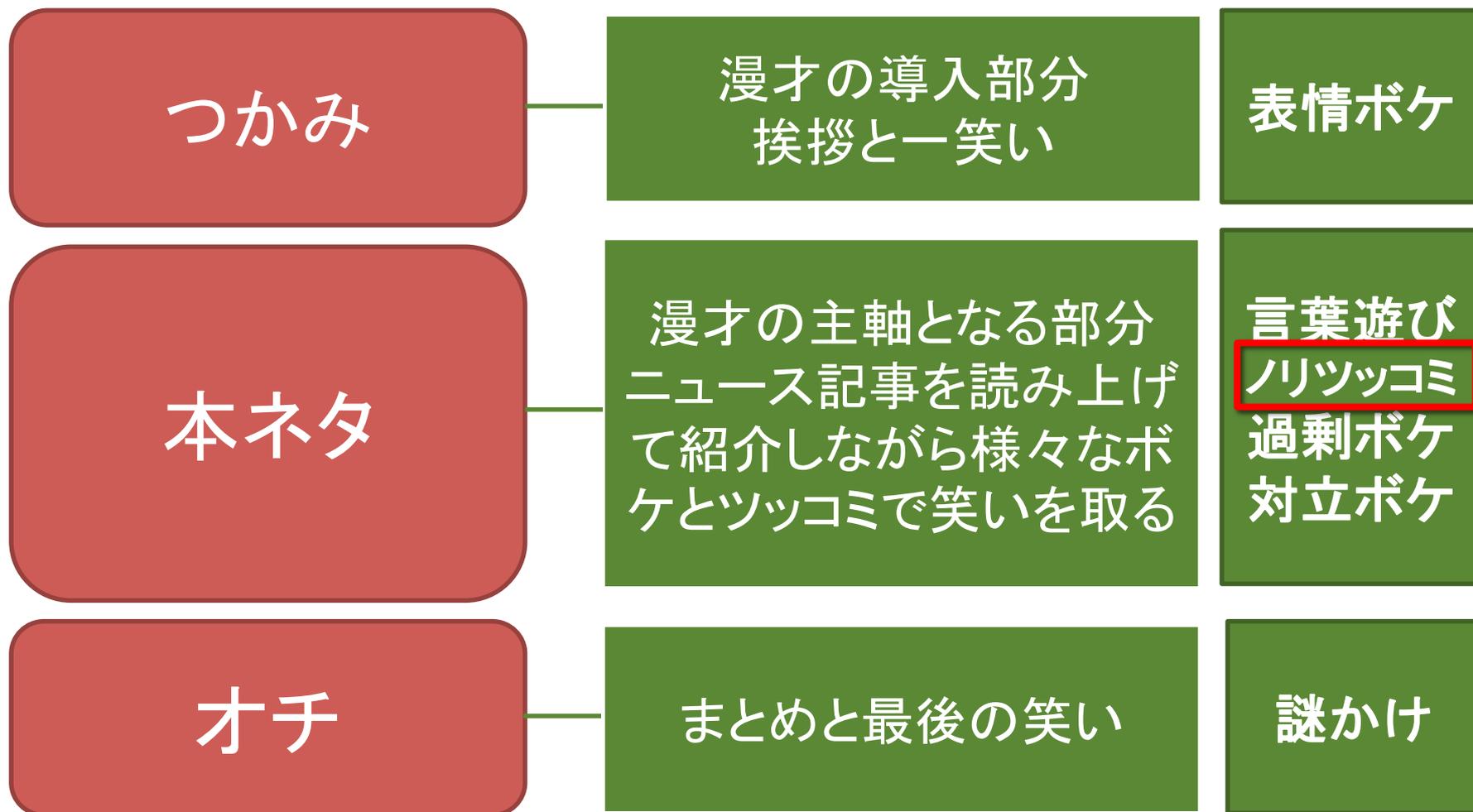


言葉遊びボケ

- 単語を別の単語と読み間違えるボケ



漫才の構成(三段構成)



ハリツッコミ

ツッコミ役が一度ボケの内容に同調し話題を展開した後に、改めて正しいツッコミを行うツッコミ

ボケ

言葉遊びボケ

最下位を決める”凍傷”で
落選した

おっと、うっかりしてもうた

ツッコミ

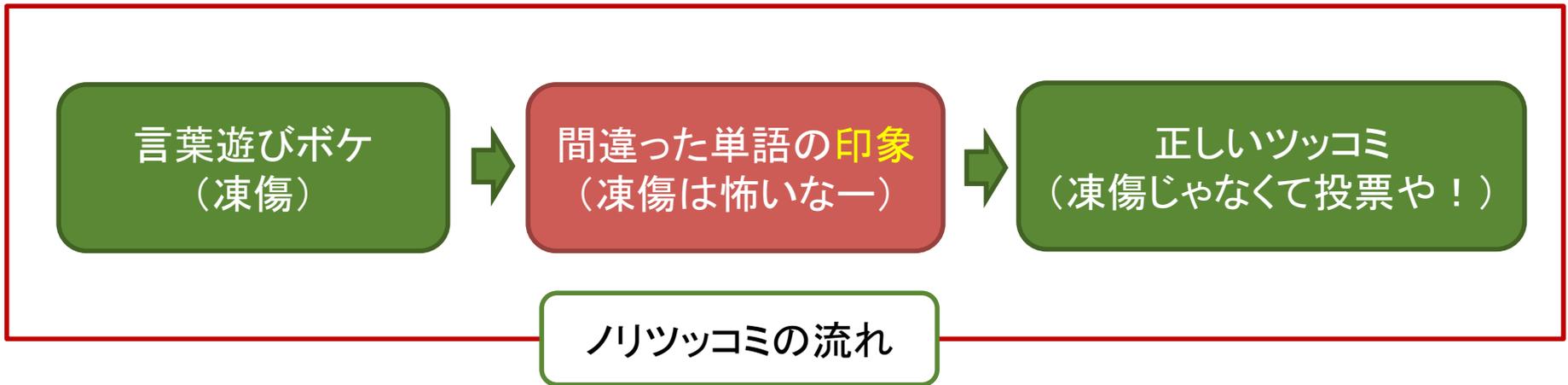
同調

ハリツッコミ

そうそう、”凍傷”は
ホント”怖い”なー

ってなんでやねん!”凍傷”って!
それは低温が原因で生じる皮
膚や皮下組織の傷害やろ!
”凍傷”ちゃうくて”投票”や!

ノリツッコミ



ノリツッコミでは間違った単語の印象を用いて同調する

印象の抽出

- **印象 ≒ 形容詞**

印象を抽出する単語をクエリとしてWeb検索を行い、上位100件のスニペットを取得

クエリ = 凍傷

取得したスニペットからクエリと共起する形容詞とその共起頻度を取得

痛い[8], 怖い[15], 寒い[12].....

形容詞を共起頻度によってランキングを行い最上位の形容詞を単語の印象として抽出

怖い[15]

印象の抽出結果例

Adjective	1st	2nd	3rd
凍傷	怖い	寒い	白い
豆腐	おいしい	美味しい	やわらかい
洋風	かわいい	明るい	おいしい
賛辞	すごい	素晴らしい	うれしい
テレビ	楽しい	めざまし	幅広い

概ね単語の印象となるものを取得できている

ハリツッコミ

ボケ

言葉遊びボケ

最下位を決める”凍傷”で
落選した

ツッコミ

ハリツッコミ

説明

そうそう, ”凍傷”は
ホント”怖い”なー

ってなんでやねん! ”凍傷”って!
それは低温が原因で生じる皮
膚や皮下組織の傷害やろ!
”凍傷”ちゃうくて”投票”や!

おっと, うっかりしてもうた

ツッコミによる説明

- ツッコミにより間違った単語の説明を補足

説明としては堅い

凍傷: 低温が原因で生じる皮膚や皮下組織の傷害



間違った単語のWikipedia記事
記事最初の1文を用いる

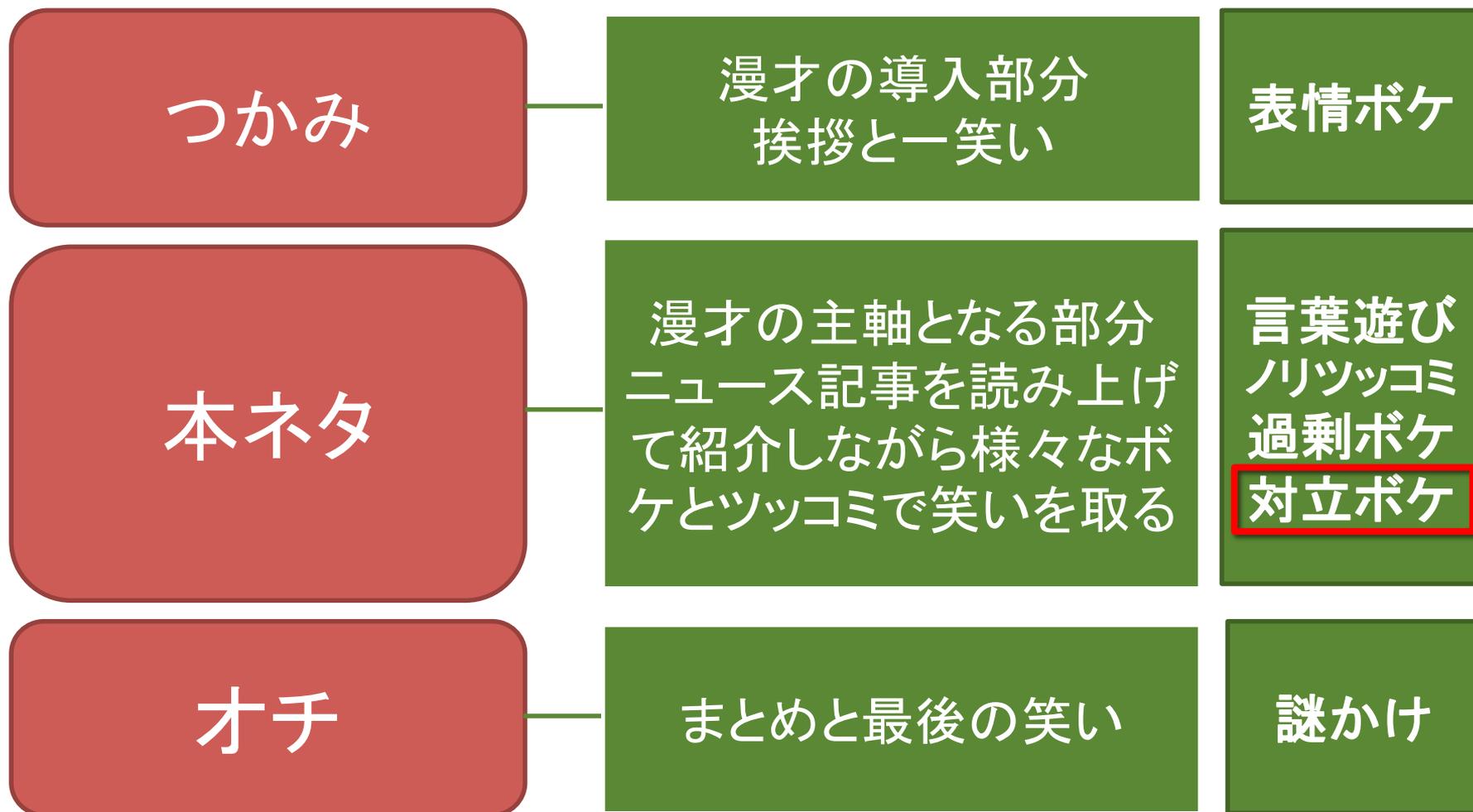


WIKIPEDIA

グライスの協調の原理より

量の公理に違反するため、漫才台本における堅い説明文には
笑いを誘う効果が期待される

漫才の構成(三段構成)



対立ボケ

ニュース記事だけでは得られない知識の拡大
対立関係は直感的に把握しやすくボケとして明確

ボケ

ツッコミ

「2020年夏季五輪の開催都市は**東京**が選ばれた」

なるほど
ところで**東京**ってどんなかわか
ってるか？

対立関係

あれやろ？
近畿地方に属する日本の都道府県の一つやろ？

お前それ**大阪**と勘違いしてるやろ。
東京は、日本の関東地方南部に位置し、1869年2月11日以来、日本の首都が置かれている都市、地域や！

そうなんか、まあでも似たようなもんやろ

どこがやねん、怒られるで！

対立関係の定義

対立関係を持つ語同士は共通の**上位概念**を持つ

- 東京 ⇔ 大阪 : 日本の都市
- 野球 ⇔ サッカー : 球技

対立関係を持つ語同士の**認知度**が近似している

- **野球 ⇔ サッカー : 野球と同じくらいメジャー**
- 野球 ⇔ フットサル : 野球程普及していない

「東京」の上位概念語

共通の上位概念を持つ語
同位語

対立関係を持つ語同士は共通の上位概念を持つ

「東京」の上位概念

日本の都市

ミリオンモンキーズに登場する実在都市

就航都市

就航地

グループ

過去に存在した店舗

オリジナルアルバム

曲

シングル

収録曲

登場人物

作品

Wikipedia
の階層構造
より

上位概念語の重み

- 上位概念語にそれぞれ重みを持たせる
- 少数の概念だけに含まれる上位概念語の方がより重要性が高い
 - 日本の都市 下位概念語数: 24語
 - 作品 下位概念語数: 120257語

重要

$$\text{重み} = 1 - \log \frac{\text{その概念語が持つ下位概念語の数}}{\text{全ての概念語の数}}$$

「東京」の上位概念語と重み

「東京」の上位概念	下位概念語数	重み
ミリオンモンキーズに登場する実在都市	9	12.6
日本の都市	24	11.7
就航都市	240	9.4
就航地	258	9.3
グループ	333	9.1
過去に存在した店舗	347	9.0
オリジナルアルバム	1100	7.8
曲	3717	6.7
シングル	19519	5.0
収録曲	35981	4.4
登場人物	95662	3.4
作品	120257	3.2

同位語の重み計算

重みを加えた上位概念語の下位概念語を取得

日本の都市→大阪, 神戸.....



取得した下位概念語をキーワードの同位語とする

同位語→大阪, 神戸.....



同位語の重みを計算する

\sum キーワードと共通する上位概念語の重み

「東京」の同位語と重み

「東京」の同位語	重み
ニューヨーク	35.1
シドニー	34.2
ロンドン	33.9
パリ	33.9
居酒屋	31.8
バンクーバー	30.7
北京	30.7
ひまわり	30.6
上海	26.3
大阪	24.5

認知度比

対立関係を持つ語同士の認知度は互いに近い

- 認知度 \div 検索結果数
- キーワードと同位語それぞれの検索結果数の比率から対立語を取得する

$$\text{認知度比} = 1 - \frac{\text{abs} (|X| - |Y|)}{\text{max} (|X|, |Y|)}$$

|X|: キーワードの検索結果数

|Y|: 対立候補語の検索結果数

「東京」の同位語と重み

「東京」の同位語	検索結果数	認知度比
東京	324000000	1.00
ニューヨーク	16100000	0.05
シドニー	2750000	0.01
ロンドン	16200000	0.05
パリ	17600000	0.05
居酒屋	40700000	0.17
バンクーバー	2430000	0.01
北京	625000000	0.52
ひまわり	9050000	0.03
上海	130000000	0.41
大阪	470000000	0.69

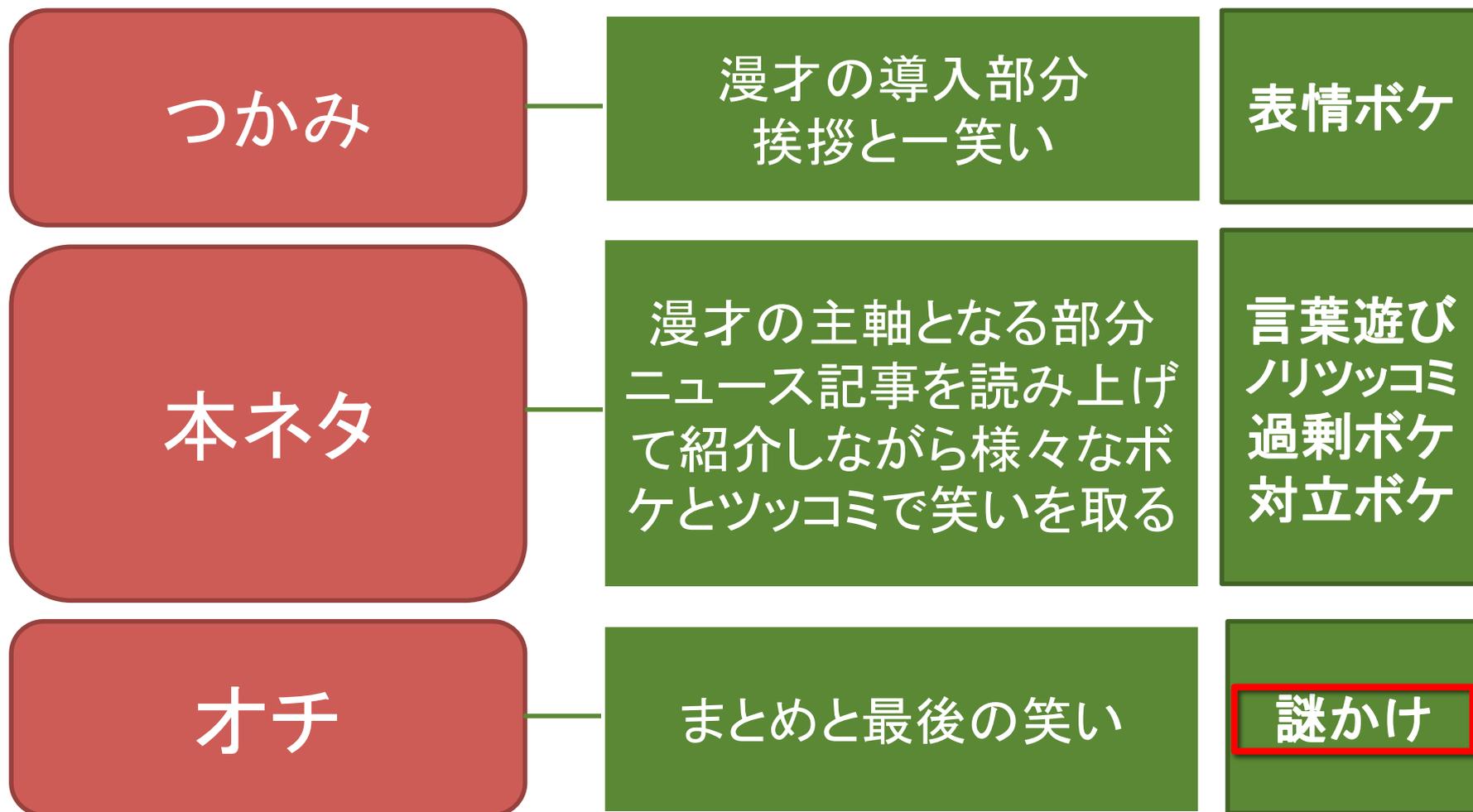
「東京」の対立語抽出結果

「東京」の同位語	重み	重み正規化	認知度比	相乗平均
ニューヨーク	35.1	100%	5%	22.4%
シドニー	34.2	97%	1%	9.8%
ロンドン	33.9	97%	5%	22.0%
パリ	33.9	97%	6%	24.1%
居酒屋	31.8	90%	17%	39.1%
バンクーバー	30.7	87%	1%	9.3%
北京	30.7	87%	45%	62.5%
ひまわり	30.6	87%	3%	16.6%
上海	26.3	75%	51%	61.8%
大阪	24.5	70%	60%	64.8%

対立語抽出結果例

キーワード	対立語
日本	中国
野球	サッカー
チェス	オセロ
犬	兎
イチロー	松井秀喜
紫式部	清少納言
バラク・オバマ	ジョージ・ワシントン
インターネット	新聞
マイクロソフト	東芝
エベレスト	阿里山

漫才の構成(三段構成)



謎かけ

東京五輪とかけて果物と解く
どちらも夏季(柿)がつきものです

A とかけて B と解く

どちらも C (C') がつきものです

A B C C' の4つの語を取得

同音異義語

A

の取得(Aとかけて)

A

ニュース記事のタイトル中からキーワードを取得

2020年「東京五輪」に決定

2013.9.8 05:20 [2020年夏季五輪]

2020年夏季五輪の開催都市を決める国際オリンピック委員会（IOC）総会は7日（日本時間8日）、ブエノスアイレスで行われ、開催都市に東京を選んだ。東京は決選投票でイスタンブールを破り、1964年以来2度目となる開催を決めた。マドリードは1回目の投票でイスタンブールと同票となり、最下位を決める投票で落選した。日本での五輪は72年札幌、98年長野の冬季五輪と合わせ4度目の開催となる。



東京の開催都市決定を示すIOCのロゲ会長＝7日、ブエノスアイレス（ロイター）



の取得 (Cがつきものです)

A をクエリとして
Web検索

スニペットから共
起頻度の高い
単語をランキング

ランキング上位の
単語を C にする

Query	1st	2nd	3rd
東京五輪	オリンピック	開催	夏季

C'

の取得 (C'がつきものです)

同音異義語を探す

Query	1st	2nd	3rd
東京五輪	オリンピック	開催	夏季



C

夏季
(kaki)



柿
(kaki)

C'

B

の取得 (Bと解く)

C' をクエリとして
Web検索

スニペットから共
起頻度の高い
単語をランキング

ランキング上位の
単語を (B) にする

Query	1st	2nd	3rd
柿	果物	富有	産地

謎かけ

ボケ

A

最後に**東京五輪**で一句
よませてもらうわ

B

東京五輪とかけて
果物と解きます

C

C'

どちらも**夏季(柿)**がつきものです

ツッコミ

ほお～, やってみ.

その心は・・・?

・・・もうええわ!

動画

エンタメ

エンタメ総合

音楽

映画

ゲーム

アジア・韓流

滝川クリステル、流行語大賞受賞に感謝！7年後の東京五輪開催にも願い

シネマトゥデイ 2013年12月3日(火)10時34分配信



オリンピック招致活動の最終プレゼンテーションに参加した滝川クリステル - Ian Walton / [Getty Images Sport](#)

フリーアナウンサーの滝川クリステルが、「お・も・て・な・し」が「2013ユーキャン新語・流行語大賞」年間大賞を受賞したことへの感謝をオフィシャルブログにつづった。その一方で、「お・も・て・な・し」が生まれるきっかけとなった東京オリンピックの開催は7年後。「決して一過性の言葉として忘れてしまうのではなく、7年後には、大輪の花を咲かせ、その後もこの『おもてなし』という文化がずっと続いていくことを願っています」と願いを込めている。

滝川クリステル、こける…フォトギャラリー

滝川は年間大賞が発表された2日夜、1月以上ぶりにブログを更新すると、「本日、『おもてなし』が2013ユーキャン新語・流行語『年間大賞』に選ばれました。ありがとうございます！」と感謝。続けて、受賞は自分だけのものではなく、招致に関わった全員のものであると強調すると、「招致メンバーとして声を掛けていただけたこと、そして、この言葉を世界の舞台上で発することができたこと、まさに奇跡だったような気がします」と改めて大舞台を振り返った。

「日本人が元々持っている、大事な大事な心の機微がなしえる『おもてなし』を改めて再認識し、自信を持たれたことが何よりもうれしいことです」という滝川。7年後、そしてその後も「お・も・て・な・し」の文化が続くことを願っている。

滝川は9月7日にアルゼンチン・ブエノスアイレスで行われたオリンピック招致活動の最終プレゼンテーションで「お・も・て・な・し」とスピーチ。開催都市が東京に決定したこともあり、「お・も・て・な・し」は一躍ブームとなっていた。(編集部・福田麗)

まとめと今後の課題

- まとめ

- 人とロボットの円滑なコミュニケーションの実現を目的に漫才ロボットを提案し、そのための漫才台本自動生成手法を提案した。
- 漫才台本をつかみ・本ネタ・オチに分割し、主に言葉遊びボケ、ノリツッコミ、対立ボケ、謎かけの自動生成手法を提案した。

- 課題

- 漫才ロボットの有用性の確認
- 自動生成の精度向上とそのための技法の確立