

# コミックを利用した 日本語教材の 作成支援システム開発の一検討

<sup>1</sup>立命館大学大学院 情報理工学研究科

<sup>2</sup>立命館大学 情報理工学部

<sup>3</sup>サイモン・フレーザー大学

桑野 将豪<sup>1</sup> 山西 良典<sup>2</sup> 西原 陽子<sup>2</sup> 竹井 尚子<sup>3</sup>

「夕食後に家族で**1つのりんご**を分け合って食べた」のフレーズから想像されるのは...

日本だと...



家族団らんの様子が想像される



海外だと...



貧しい家族の様子が想像される



You



日常スープ©進藤 ウニ



アンバランス・トーキョー©内田 美奈子



やまとの羽根©咲 香里



手作業で  
抽出



データ化



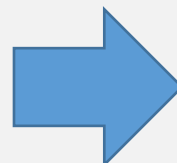
うちの猫'ず日記@gあさん

手作業で  
抽出



うちの猫'ず日記@gあさん

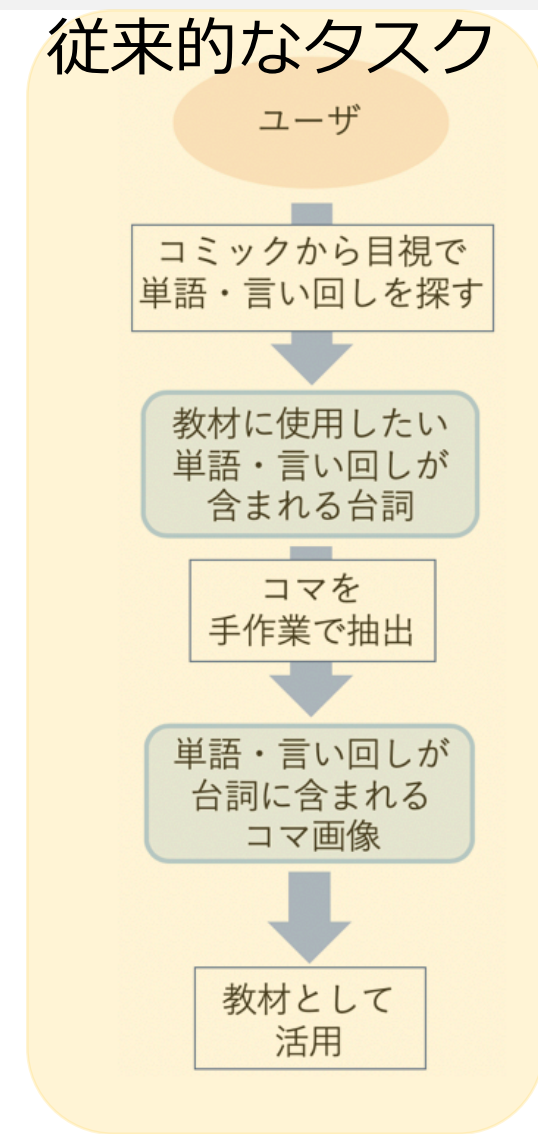
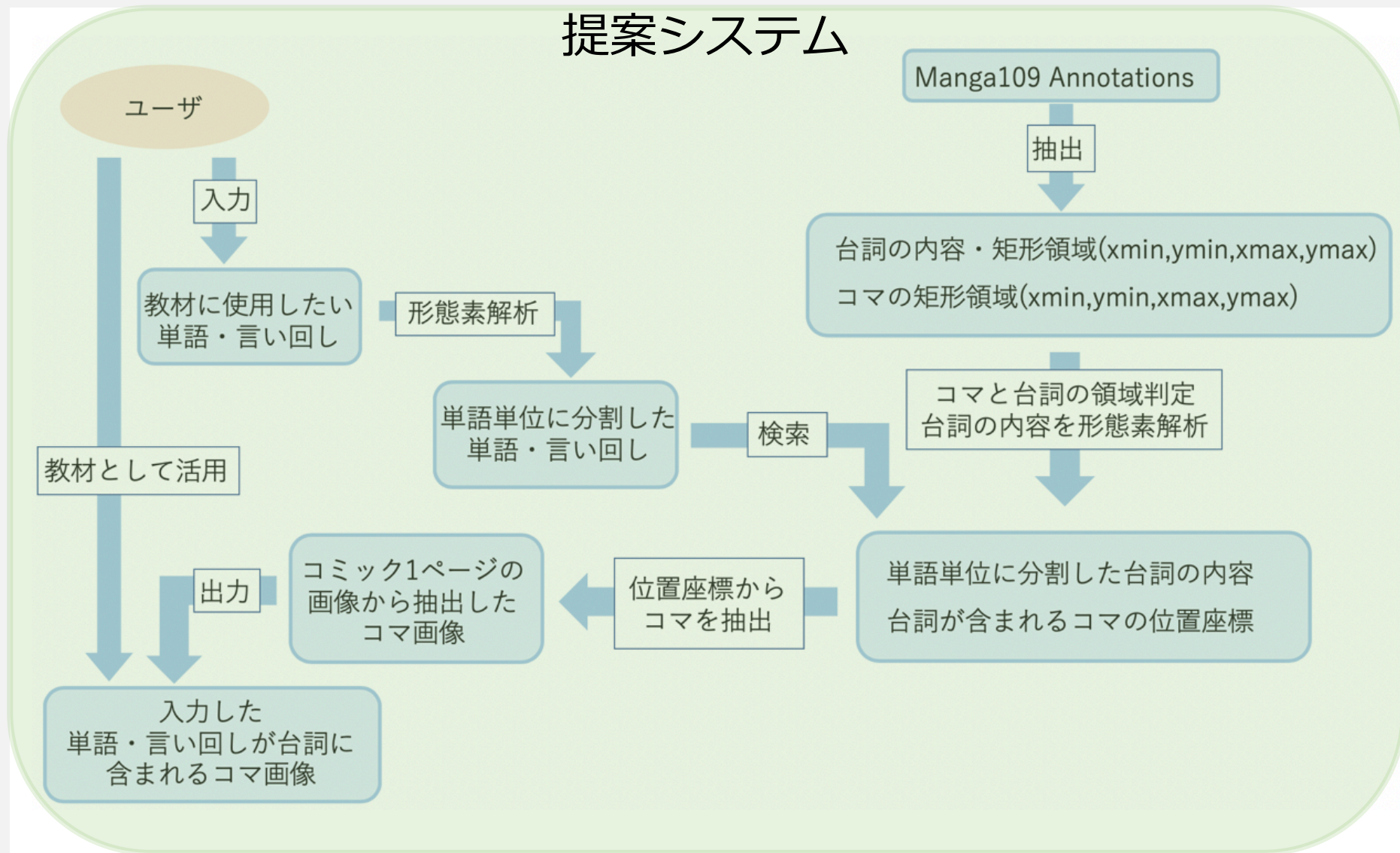
手作業だと時間がかかってしまう



コミックが使用されにくい



# 提案手法と従来のなタスクの比較



1. Manga109 Annotationsからデータを取得
2. 取得データの処理
3. 入力データの処理
4. 入力データと台詞の部分一致判定
5. コマ画像の抽出



# 1.Manga109 Annotationsからデータ取得

台詞情報 (位置・内容)  
コマ情報 (位置)

見開きページの画像情報

```
<page index="11" width="1654" height="1170">
<body id="0003613e" xmin="66" ymin="622" xmax="299" ymax="778" character="00035f85"/>
<text id="0003613f" xmin="252" ymin="626" xmax="287" ymax="752">---は・.</text>
<face id="00036140" xmin="337" ymin="901" xmax="370" ymax="938" character="00036141"/>
<body id="00036142" xmin="1101" ymin="624" xmax="1211" ymax="738" character="00035fa6"/>
<face id="00036143" xmin="529" ymin="628" xmax="614" ymax="702" character="00035f85"/>
<face id="00036144" xmin="112" ymin="1003" xmax="151" ymax="1082" character="00035f83"/>
<body id="00036145" xmin="324" ymin="928" xmax="676" ymax="1169" character="00035f85"/>
<body id="00036146" xmin="119" ymin="928" xmax="836" character="00035f85"/>
<face id="00036147" xmin="35" ymin="928" xmax="99" ymax="1082" character="00035f85"/>
<face id="00036148" xmin="113" ymin="928" xmax="698" character="00035f85"/>
<face id="00036149" xmin="68" ymin="928" xmax="76" character="00035f85"/>
<face id="0003614a" xmin="1456" ymin="961" xmax="1493" ymax="998" character="00035f85"/>
<body id="0003614b" xmin="20" ymin="794" xmax="744" ymax="1003" character="00036141"/>
<text id="0003614c" xmin="629" ymin="794" xmax="744" ymax="1003">はるかおばさん!! おひさしぶり---</text>
<text id="0003614d" xmin="913" ymin="60" xmax="1055" ymax="207">すなわち女子寮なのよ</text>
<frame id="0003614e" xmin="908" ymin="780" xmax="1449" ymax="1167"/>
<body id="0003614f" xmin="314" ymin="622" character="00036141"/>
<face id="00036150" xmin="1054" ymin="622" character="00035faa"/>
<text id="00036151" xmin="649" ymin="622" character="00035faa">おとなしく警察に行ってもらわよ</text>
<text id="00036152" xmin="931" ymin="622" character="00035faa">女子寮お〜〜!!?</text>
<body id="00036153" xmin="156" ymin="83" xmax="339" ymax="1169" character="00035f83"/>
<text id="00036154" xmin="685" ymin="374" xmax="740" ymax="472">騒々しいどうした</text>
<frame id="00036155" xmin="904" ymin="2" xmax="1651" ymax="832">
<text id="00036156" xmin="1457" ymin="64" xmax="1619" ymax="271">男子禁制の契約アパート!</text>
<text id="00036157" xmin="312" ymin="103" xmax="361" ymax="192">ん・・・?なんだ</text>
<frame id="00036158" xmin="377" ymin="81" xmax="745" ymax="347"/>
<body id="00036159" xmin="406" ymin="622" xmax="639" ymax="778" character="00035f85"/>
<body id="0003615a" xmin="872" ymin="472" xmax="1055" ymax="784" character="00035fa8"/>
<body id="0003615b" xmin="1437" ymin="932" xmax="1507" ymax="1088" character="00035f85"/>
<body id="0003615c" xmin="1068" ymin="803" xmax="1447" ymax="1169" character="00035f85"/>
<body id="0003615d" xmin="81" ymin="930" xmax="153" ymax="1169" character="00035f83"/>
```

台詞の位置情報

台詞の内容

コマの位置情報

(0,0)

X軸

Y軸



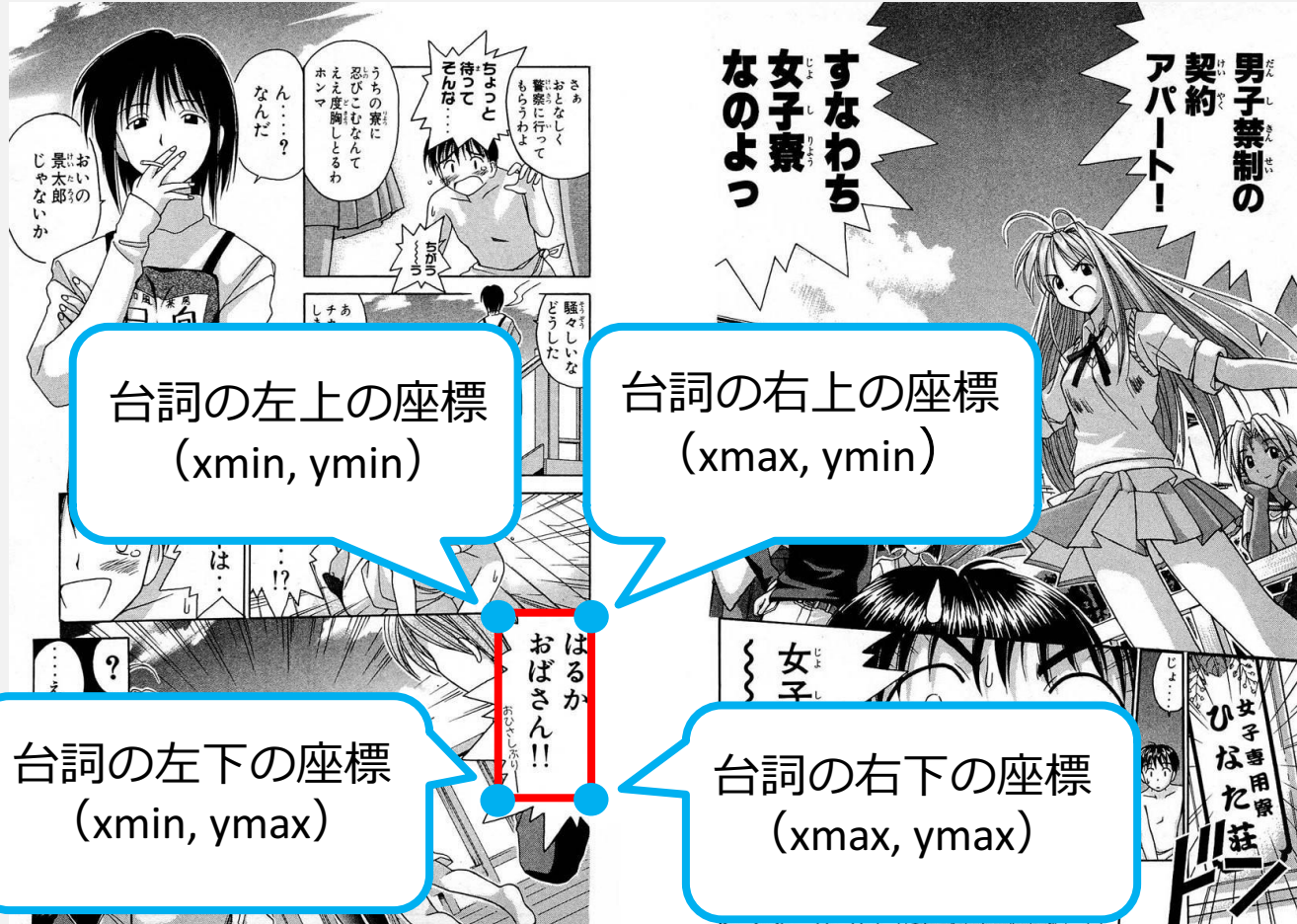


# 2.1取得データの処理（台詞の位置情報）

## 台詞の位置情報から台詞の位置座標を取得

```
<page index="11" width="1654" height="1170">
<body id="0003613e" xmin="66" ymin="622" xmax="299" ymax="778" character="00035f85"/>
^144" xmin="112" ymin="1003" xmax="151" ymax="1082" cha
^145" xmin="324" ymin="928" xmax="676" ymax="1169" cha
^146" xmin="119" ymin="836" character="00035f85"/>
^147" xmin="35" ymin="990" character="00035f85"/>
^148" xmin="113" ymin="698" character="00035f85"/>
^149" xmin="68" ymin="76" character="00035f85"/>
^14a" xmin="1456" ymin="961" xmax="1493" ymax="998" cha
^14b" xmin="20" ymin="20" xmax="244" ymax="618" charact
^14c" xmin="629" ymin="794" xmax="744" ymax="1003">はる
^14d" xmin="915" ymin="60" xmax="1095" ymax="207">すなわ
^14e" xmin="908" ymin="0" xmax="1449" ymax="1167"/>
^14f" xmin="314" ymin="1169" character="00035f85"/>
<text id="00036157" xmin="312" ymin="1003" ymax="1169" character="00035f85"/>
<frame id="00036158" xmin="377" ymin="836" ymax="990"/>
<body id="00036159" xmin="406" ymin="622" ymax="778" character="00035f85"/>
<body id="0003615a" xmin="872" ymin="478" ymax="784" character="00035fa8"/>
<body id="0003615b" xmin="1437" ymin="928" ymax="1088" character="00035f85"/>
<body id="0003615c" xmin="1068" ymin="836" ymax="1169" character="00035f85"/>
<body id="0003615d" xmin="81" ymin="930" ymax="1169" character="00035f83"/>
```

台詞の位置情報



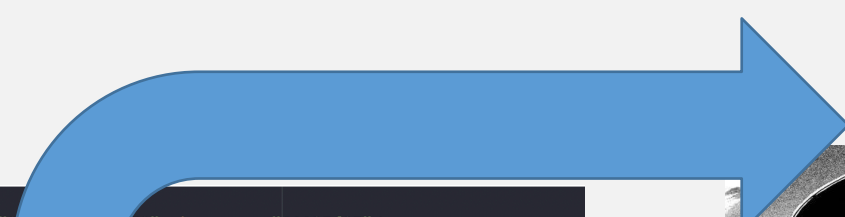


# 2.2取得データの処理（コマの位置情報）

## コマの位置情報からコマの位置座標を取得

```
<page index="11" width="1654" height="1170">  
<body id="0003613e" xmin="66" ymin="622" xmax="278" character="00035f85"/>  
<text id="0003613f" xmin="252" ymin="626" xmax="752" character="00035f85"/>  
<face id="00036140" xmin="337" ymin="901" xmax="938" character="00036141"/>  
<body id="00036142" xmin="1101" ymin="624" xmax="738" character="00035fa6"/>  
<face id="00036143" xmin="529" ymin="628" xmax="702" character="00035f85"/>  
<face id="00036144" xmin="112" ymin="1003" xmax="1082" character="00035f83"/>  
<body id="00036145" xmin="324" ymin="928" xmax="1169" character="00035f85"/>  
<body id="00036146" xmin="1199" ymin="258" xmax="836" character="00035f83"/>  
<face id="00036147" xmin="35" ymin="1011" xmax="1090" character="00035faa"/>  
<face id="00036148" xmin="1139" ymin="651" xmax="1141" ymax="698" character="00035fa6"/>  
" xmin="314" ymin="622" character="00035f85" >  
" xmin="1054" ymin="622" character="00035f85" >  
" xmin="649" ymin="622" character="00035f85" >  
" xmin="931" ymin="622" character="00035f85" >  
" xmin="156" ymin="832" xmax="339" ymax="1169" character="00035f85" >  
" xmin="685" ymin="374" xmax="740" ymax="472" >騒々しいなど  
5" xmin="904" ymin="2" xmax="1651" ymax="832" >  
" xmin="1457" ymin="64" xmax="1619" ymax="271" >男子禁制の契  
" xmin="312" ymin="103" xmax="361" ymax="192" >ん・・・?  
" xmin="377" ymin="81" xmax="745" ymax="347" >  
<body id="0003615d" xmin="81" ymin="930" xmax="153" ymax="1169" character="00035f83"/>
```

コマの位置情報





# 2.3取得データの処理 (台詞・コマの領域判定) 10

コマの枠内に台詞の枠が完全に**含まれる**



コマ内に台詞が**存在**すると判定



## 2.4取得データの処理（台詞の内容）

台詞の一例：すなわち女子寮なのよっ

単語単位に分割      形態素解析器：MeCab  
辞書：NEologd

すなわち / 女子寮 / な / の / よっ

接続詞

名詞

助動詞

名詞

動詞

ユーザが入力した教材に使用したい単語・言い回し  
入力例：せっかく

単語単位に分割      形態素解析器：MeCab  
辞書：NEologd

せっかく

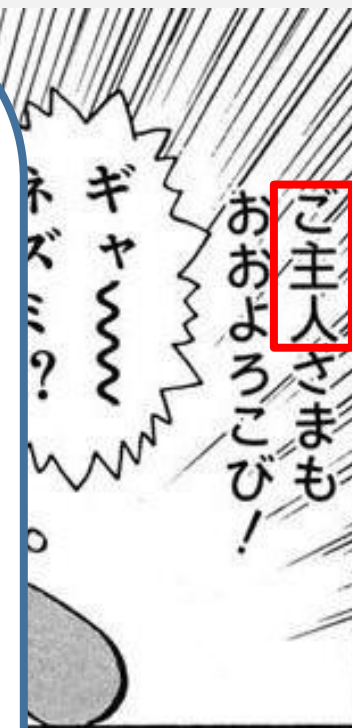
副詞



入力データ

教材に使用したい  
**単語・言い回しが**  
**台詞**に含まれていると  
判定できる

台詞



うちの猫'ず日記@がぁさん

!

# 5.コマ画像の抽出

コマの位置座標をもとに見開きページの画像からコマ画像を抽出



抽出



10作品のコミックを対象に入力データの検索を実施した

## コミック一覧

- ・ベルモンド Le Visiteur
- ・エブリデイおさかなちゃん
- ・はるかりフレイン
- ・じじばばファイト!
- ・ラブひな
- ・日常スーブ
- ・OLランチ
- ・うちの猫'ず日記
- ・アンバランス・トーキョー
- ・やまとの羽根

検索の入力データには14種類の副詞・接続詞を使用した

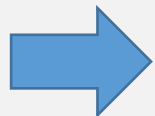
## 入力データ一覧

### 副詞

- ・いまにも
- ・おそらく
- ・すでに
- ・せっかく
- ・ますます
- ・まんがいち

### 接続詞

- ・それとも
- ・だけど
- ・つまり
- ・ところが
- ・そのうえ
- ・ところで
- ・それと
- ・なぜなら




活用が存在せず検索が容易だと考えたため

# 実験結果1 (入力データが含まれるコマの抽出) 16

## 10作品に対する検索結果

10作品のコマ数の  
合計9260コマ



抽出

73コマを  
抽出することができた

タイトル	1巻分のコマ数	検案件数
ベルモンド Le VisiteuR	921	14
エヴリデイおさかなちゃん	912	14
アンバランス・トーキョー	957	12
うちの猫'ず日記	985	10
じじばばファイト!	862	6
OLランチ	833	5
はるかりフレイン	839	4
日常スープ	845	4
やまとの羽根	1029	3
ラブひな	1077	1



本研究における会話とは、 . . . コマ中で会話が行われているコマの検索結果



- コマ中に複数のキャラクターが存在
- 複数のキャラクターが発言を行う

73コマ中20コマ抽出

タイトル	検索件数	会話が行われているコマ数
ベルモンド Le VisiteuR	14	1
エヴリデイおさかなちゃん	14	2
アンバランス・トーキョー	12	6
うちの猫'ず日記	10	0
じじばばファイト！	6	4
OLランチ	5	2
はるかりフレイン	4	3
日常スープ	4	2
やまとの羽根	3	0
ラブひな	1	0

## 「ところで」



アンバランス・トーキョー©内田 美奈子



アンバランス・トーキョー©内田 美奈子



OLランチ©さんり ようこ



日常スープ©進藤 ウニ



日常スープ©進藤 ウニ

## 「つまり」



はるかリフレイン©伊藤 伸平



はるかリフレイン©伊藤 伸平



アンバランス・トーキョー©内田 美奈子



OLランチ©さんり ようこ



ベルモンド Le Visiteur ©石岡 ショウエイ



「すでに」



エブリデイおさかなちゃん©国樹 由香



じじばばファイト! ©西川 伸司



じじばばファイト! ©西川 伸司

「せっかく」



じじばばファイト! ©西川 伸司



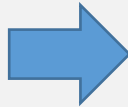
はるかりフレイン©伊藤 伸平



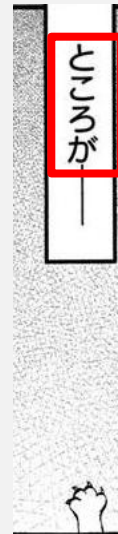
- ・ コミックからコマの自動抽出をすることができた
  - ➡ コミックを教材に使用可能なことが示唆された
- ・ 自動抽出したコマでコンテキストを把握可能なコマは僅かであった
  - ➡ 1コマでコンテキストを把握することが不可能なコマは前後のコマを同時に複数提示する必要がある



うちの猫'ず日記@gあさん



うちの猫'ず日記@gあさん



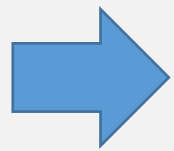
うちの猫'ず日記@gあさん



うちの猫'ず日記@gあさん

## まとめ

- ・ コミックからコマの自動抽出を可能にした



作業効率の向上を図ることができた

コミックを教材として使用可能なことが示唆された

## 今後の展望

- ・ 正規表現の利用  
「全く～ない」のような複雑な表現を検索可能にするため