

『日本語日常会話コーパス』 における職場会話

小磯花絵

2018年9月3日



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

National Institute for Japanese Language and Linguistics

NINJAL

『日本語日常会話コーパス』(構築中)

■ 対象: 日常場面の中で自然に生じた会話200時間

■ 設計: 多様な会話をバランスよく納めたコーパス

■ 規模

時間	200時間	異なり話者数(推定)	600人
会話数(推定)	450	延べ話者数(推定)	1200人

■ 収録法:

個人密着法(中心) 約180時間	<ul style="list-style-type: none">➤ 40名の協力者(5世代*男女*各4名)➤ 協力者自身の日常場面の会話を約15時間収録➤ バランス等を考慮して4~5時間/1人を選定
特定場面法 約20時間	<ul style="list-style-type: none">➤ 個人密着法で収録の難しい場面の会話を補填 例) 接客場面, 職場の会話

■ 公開データ: 映像・音声・転記テキスト・各種アノテーション

個人密着法による会話の収録例

両親と



家族と夕食



オフィス日常業務



ボーイスカウト
屋外で



ボーイスカウト
打合せ



50代男性

- 配偶者・子供2人と同居
- 両親が近くに居住
- 地元で印刷会社を経営
- ボーイスカウト活動

取引先と打合せ



接待



地元の友達と



床屋で



同僚
送別会



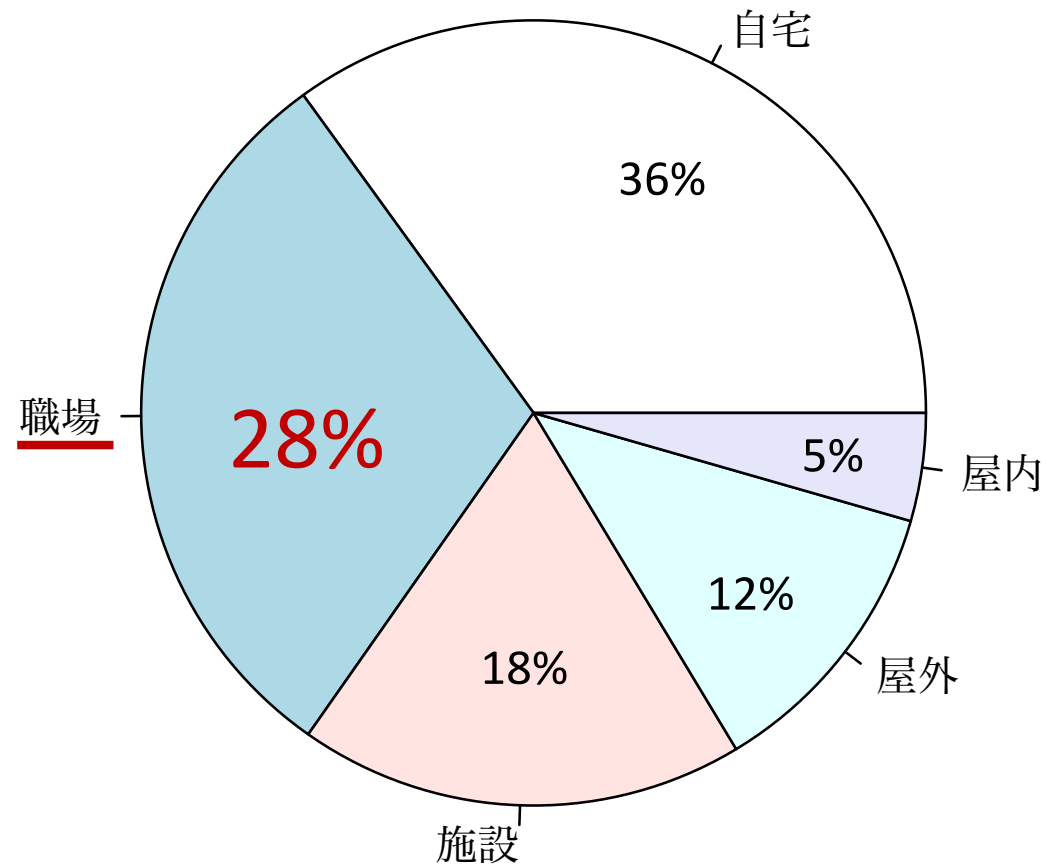
会話行動調査

私たちは普段、どのような場面の会話を、どの程度、行っている？

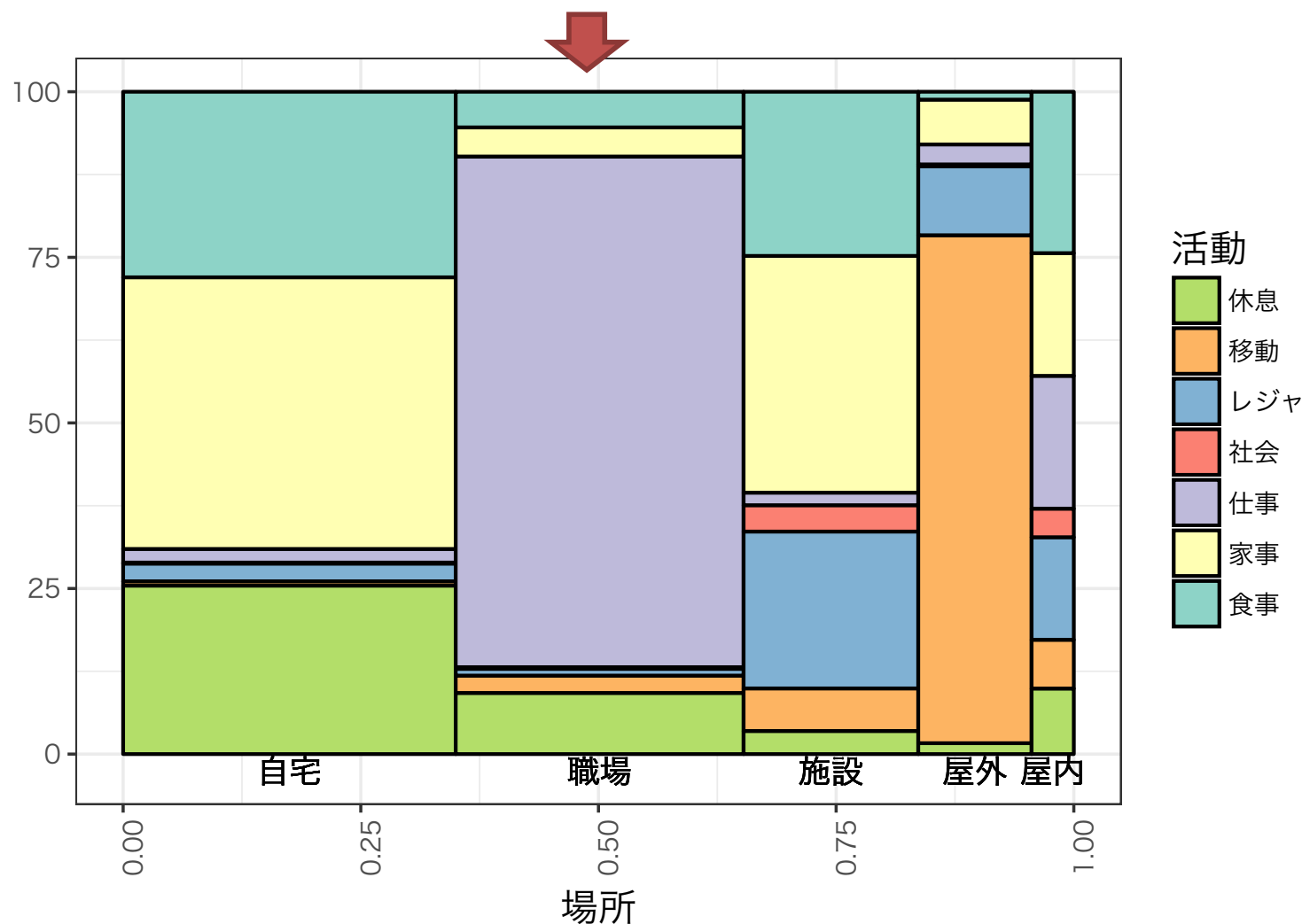
- 目的：日常会話の多様性を明らかにし、それに立脚して多様な日常会話をバランスよく納めたコーパスを設計
- 実施時期：H26.11～H27.2
- 対象：243人
(年齢・性別バランス)
- 調査日：平日2日・休日1日
(計3日／1人)
- 総会話数：9272会話

① どんな会話か			
<自由にメモしてください> レストランで1人でランチ中、旅行について友人と電話で相談			
② いつ (1つ選択)			
<input type="checkbox"/> 午前	<input checked="" type="checkbox"/> 午後	<input type="checkbox"/> 夜(午後6時頃～)	
③ どのくらい (1つ選択)			
<input type="checkbox"/> 5分未満	<input type="checkbox"/> 5～15分	<input checked="" type="checkbox"/> 15～30分	<input type="checkbox"/> 30分～1時間
<input type="checkbox"/> 1～2時間	<input type="checkbox"/> 2～5時間	<input type="checkbox"/> 5～10時間	<input type="checkbox"/> 10時間以上
④ どこで (1つ選択)			
<input type="checkbox"/> 自宅	<input type="checkbox"/> 職場・学校	<input checked="" type="checkbox"/> 公共商業施設	<input type="checkbox"/> 交通機関
<input type="checkbox"/> それ以外の屋内		<input type="checkbox"/> それ以外の屋外	
⑤ だれと (あてはまるものそれぞれに人数を記入)			
家族：_____人	親戚：_____人	先生・生徒：_____人	
仕事・学業関係：_____人	公共商業関係：_____人		
友人・知人： <u>1</u> 人	顔見知り・見知らぬ人：_____人		
⑥ 何をしながら (1つ選択)			
<input checked="" type="checkbox"/> 食事	<input type="checkbox"/> 家事・雑事	<input type="checkbox"/> 身周りの用事	<input type="checkbox"/> 療養
<input type="checkbox"/> 仕事・学業	<input type="checkbox"/> 業務外・課外活動	<input type="checkbox"/> 社会参加	
<input type="checkbox"/> レジャー活動	<input type="checkbox"/> 付き合い	<input type="checkbox"/> 移動	<input type="checkbox"/> 休息
⑦ どんな種類 (1つ選択)			
<input type="checkbox"/> 雑談	<input checked="" type="checkbox"/> 用談・相談	<input type="checkbox"/> 会議・会合	<input type="checkbox"/> 授業・レッスン・講演
⑧ その他 (あてはまるものすべて選択)			
<input checked="" type="checkbox"/> 電話・スカイプなどの遠隔での音声・映像会話			
<input type="checkbox"/> 外国人を含む会話		<input type="checkbox"/> 外国語を含む会話	

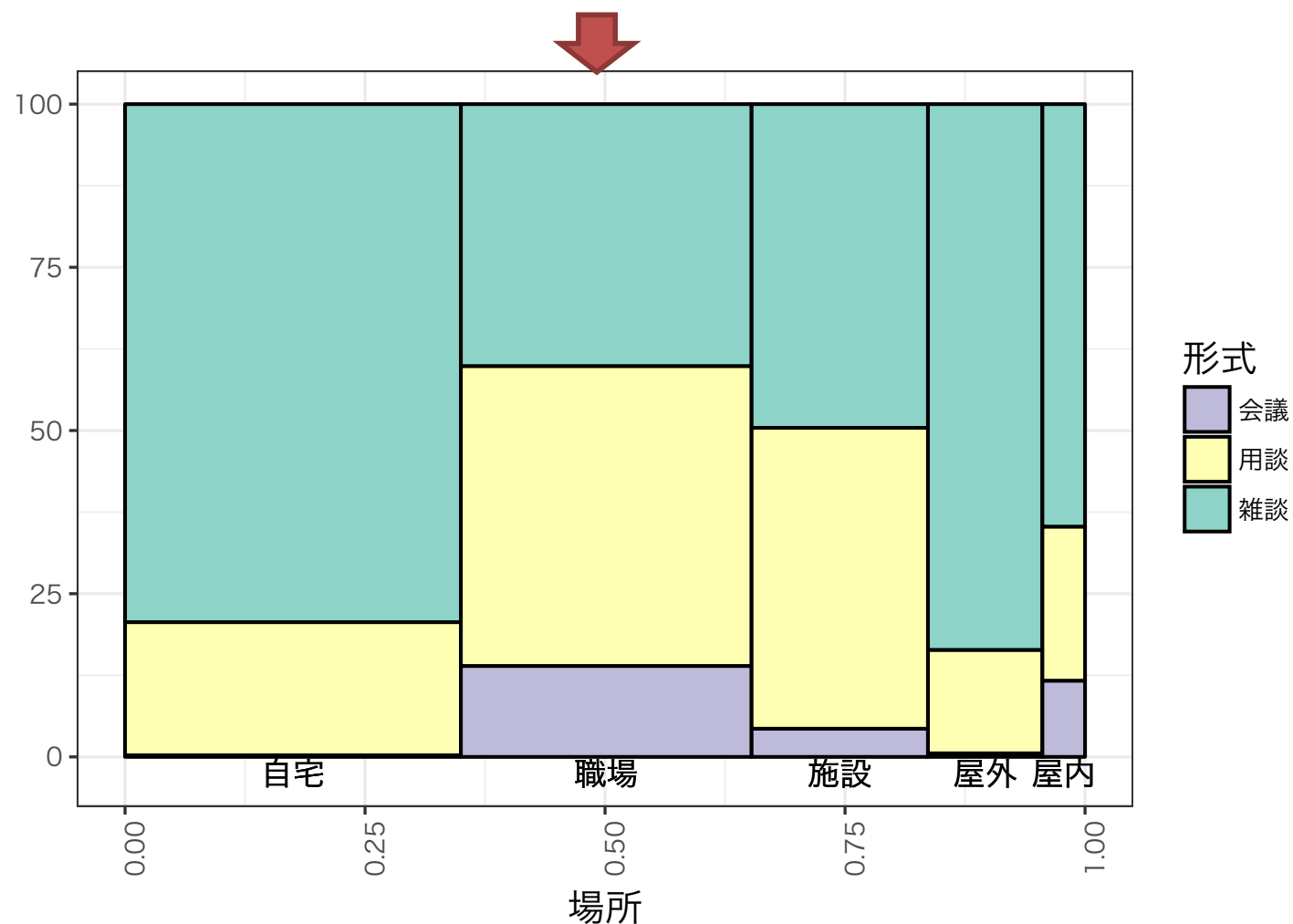
調査結果 -場所-



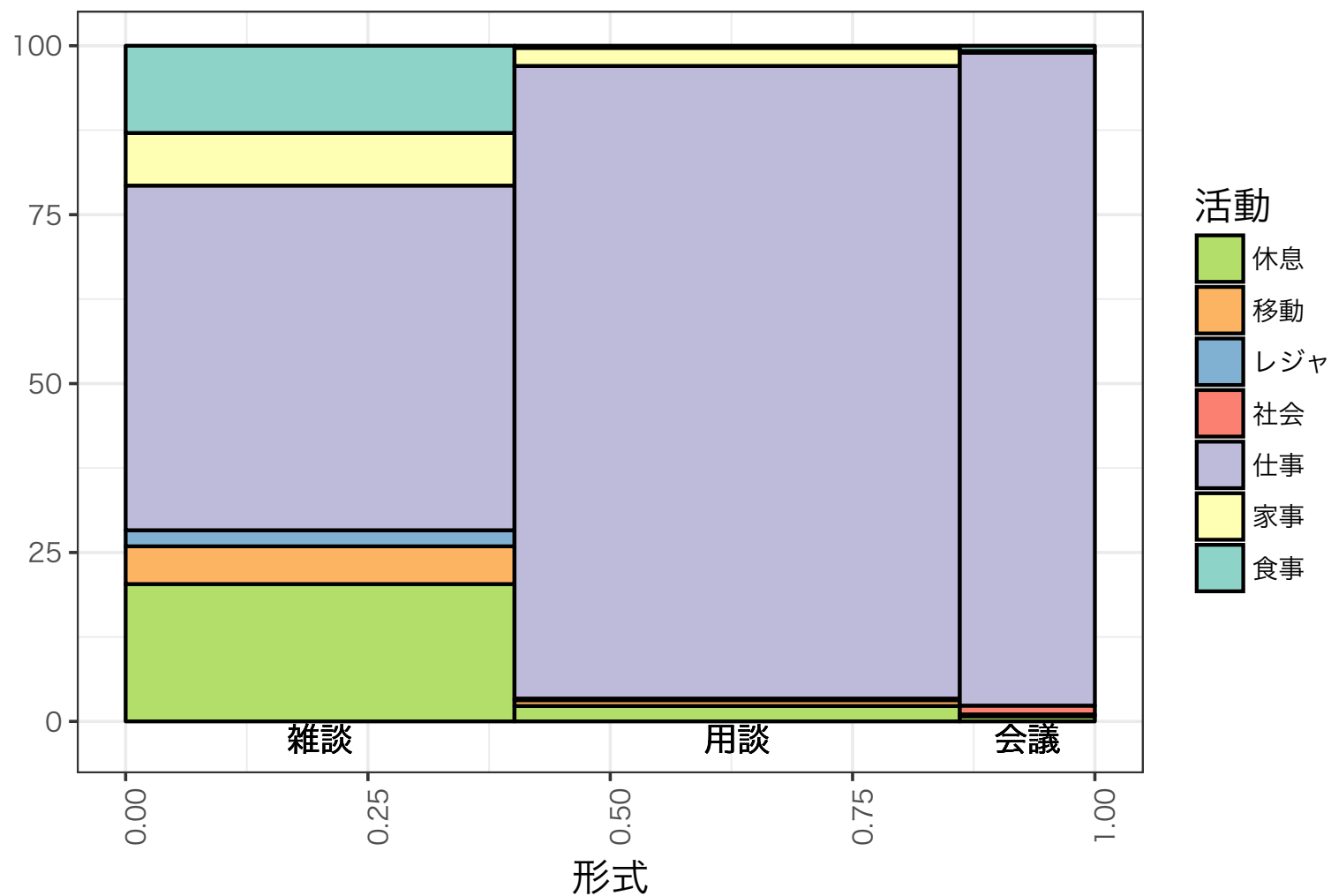
調査結果 -場所 × 活動-



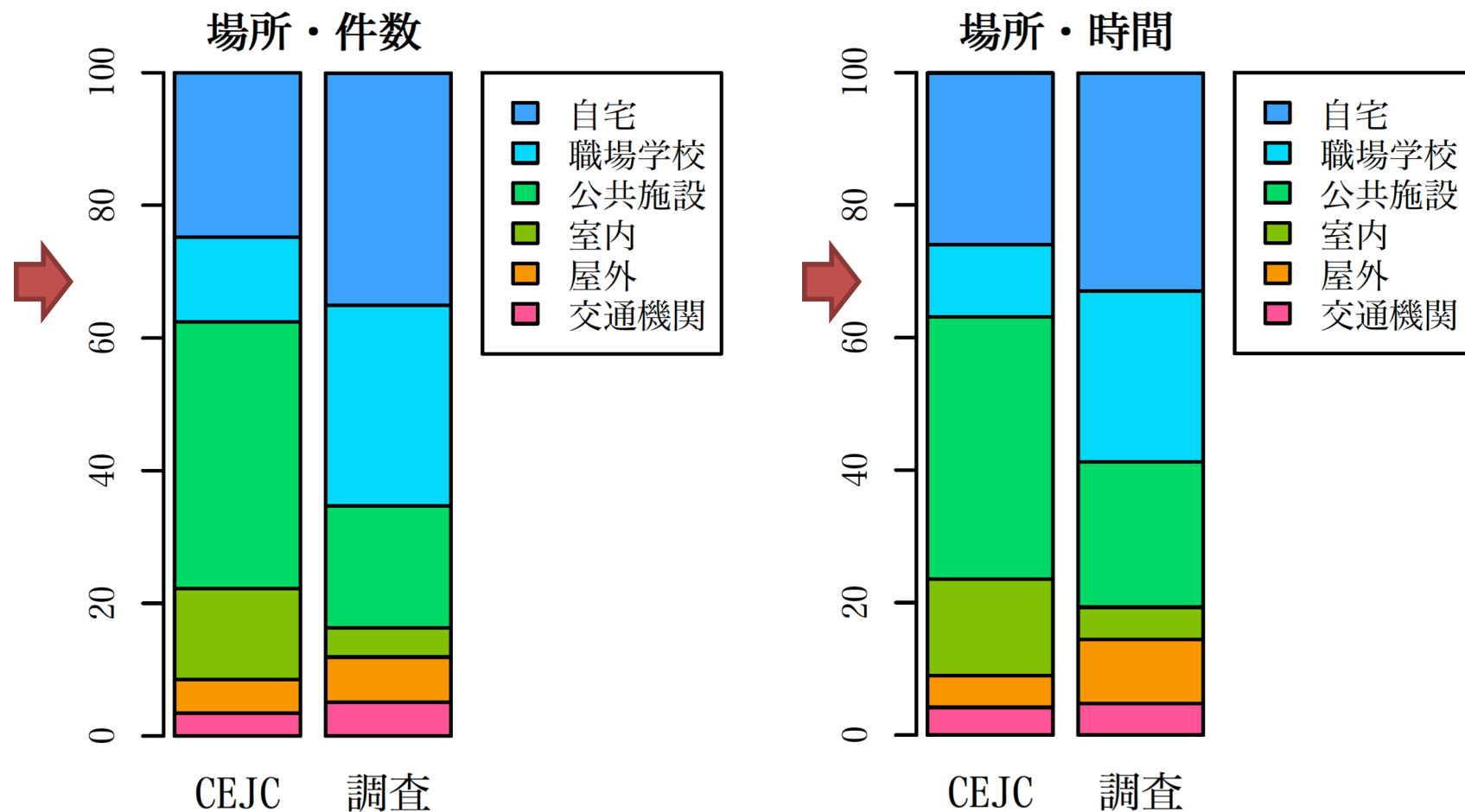
調査結果 -場所 × 会話形式-



調査結果 - 「職場」での会話形式 × 活動 -



これまでに収録された会話の場所の内訳



行動調査結果と比べて、職場で収録された会話の割合は少ない

職場データは本当に少ない？

- ✓ 床屋・美容院での理容師・美容師との会話
- ✓ 飲食店での店員との会話
- ✓ 歯科医院での歯医者との会話
- ✓ 取材先での店主との会話
- ✓ 習いごと教室での先生との会話

調査協力者にとっては職場ではないが相手にとっては職場

→ これらを合わせると職場データは少なからずある
(特に接客場面)

これまでに収録された職場の会話

- 会議・打合せ場面
- 日常業務場面
- 接客場面
- 教授場面
- 休憩場面、など

実例をご紹介します
(ハンドアウトは割愛)

公開情報

■ 2018年12月頃 50時間 モニター公開

- 協力者20人 × 平均2.5時間
- 117会話, 延べ話者数435人, 異なり話者数281人
- 映像・音声・転記テキスト・形態論情報(短単位)
- 公開方法
 - オンライン検索システム『中納言』での公開: 無償
 - HDDでの映像・音声・転記データ等の配布: 実費負担(1万円予定)

■ 2021年度末 200時間 本公開