

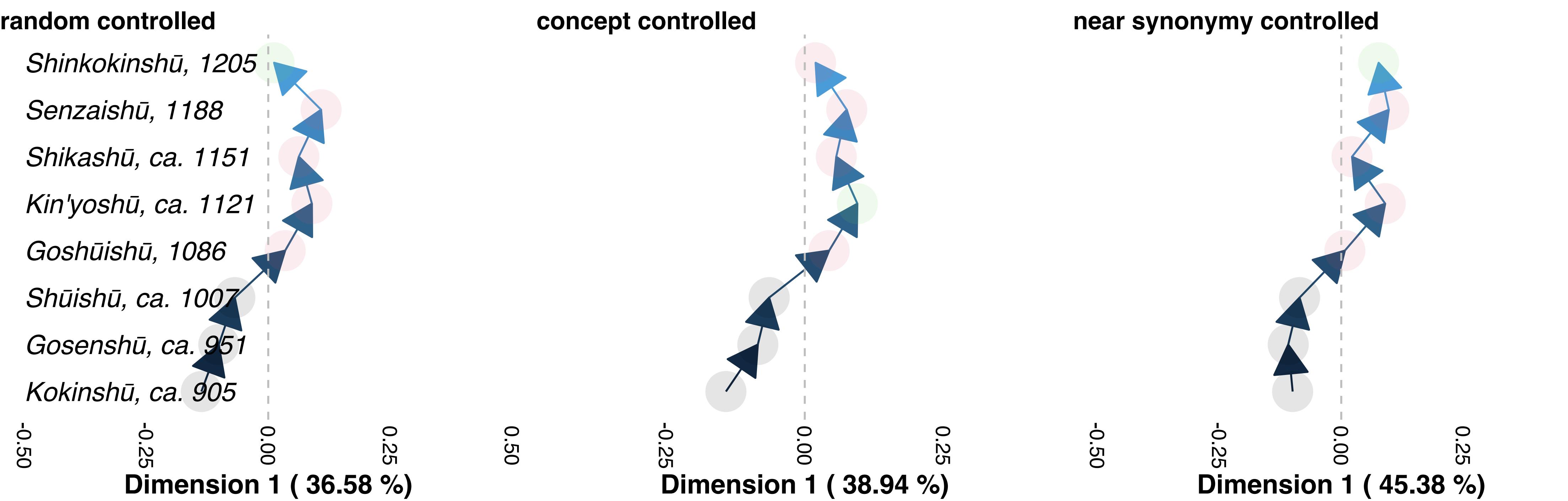
語彙プロファイルにおける変化

じんもんこん 2025, 2025 年 12 月 12 日

陳旭東 † ホドシチエクボル ‡ 山元啓史 †



結果 1.



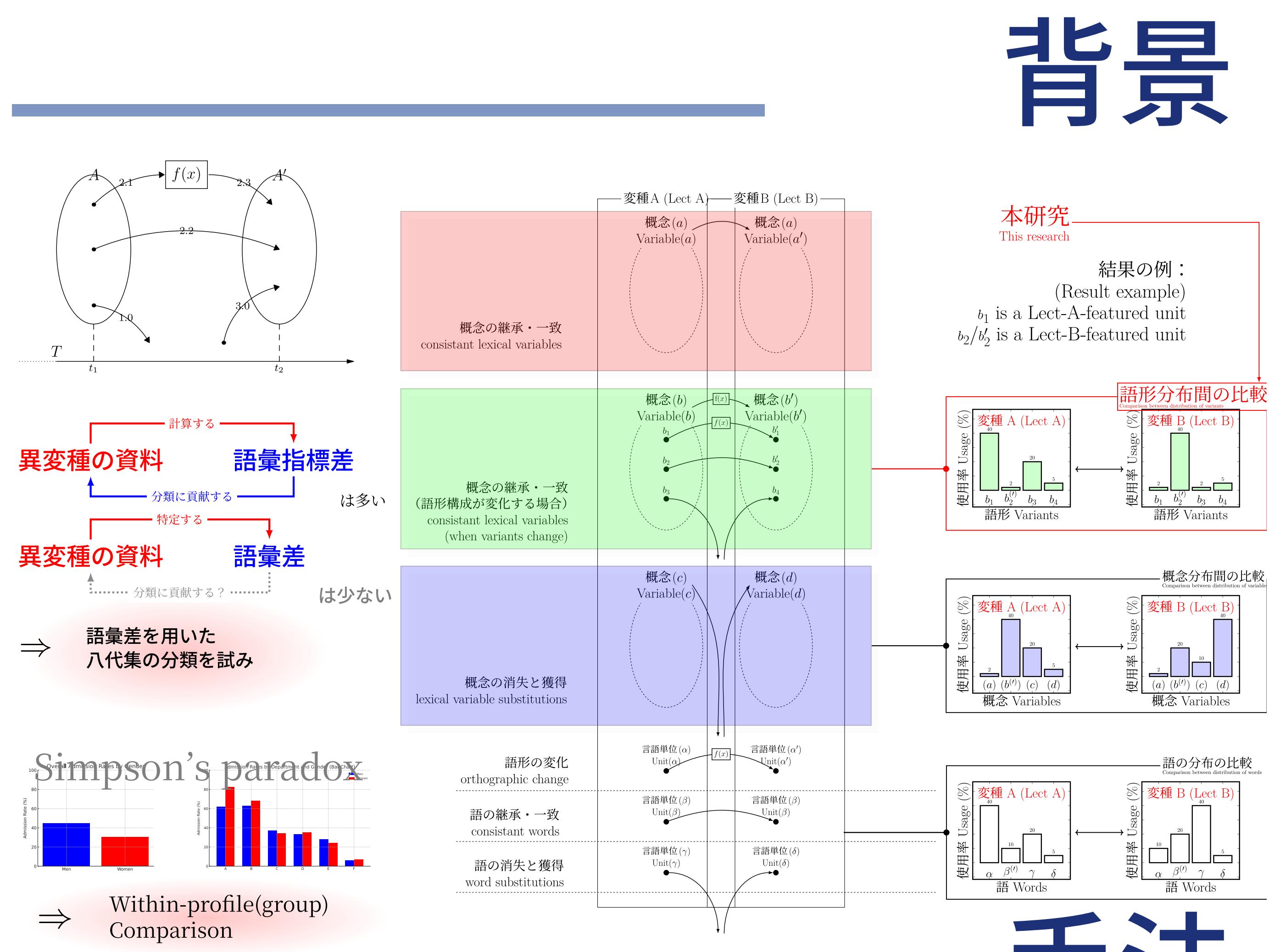
- 異なる水準のパターンに大きな相違が見られず
- 後拾遺で正負転換 (0 値横断)
- クラスタリング :
 - 「古今・後撰・拾遺」同クラスター
 - 「後拾遺・詞花・千載」同クラスター
 - 「新古今」「金葉」は中間水準で不一致
- 金葉以降の往還 :
 - △ 金葉 (同時代重視)
 - ▼ 詞花 (後拾遺集歌人重視)
 - ▽ 千載 (同時代重視)
 - ▼ 新古今 (本歌取り隆盛)

結果 2.



三水準共通点 :

- 「金葉→詞花」「拾遺→後拾遺」の変化が「古今→後撰」の変化より大きい



背景

手法

比較対			
「ムグラ」語群	葎	八重葎	
歌集 A	100%	0%	100%
歌集 B	62.5%	37.5%	100%

$$\text{Profile-based dissimilarity}(A, B | \text{ムグラ}) = [1.00 - 0.625] + [0.00 - 0.375] = 0.75$$

Comparison by likelihood ratio test; return p value

古今	後撰	…	新古今	桜花
28	18	…	12	梅
23	16	…	25	權
:	:	…	:	:
0	0	…	1	

(a) Absolute frequency matrix*

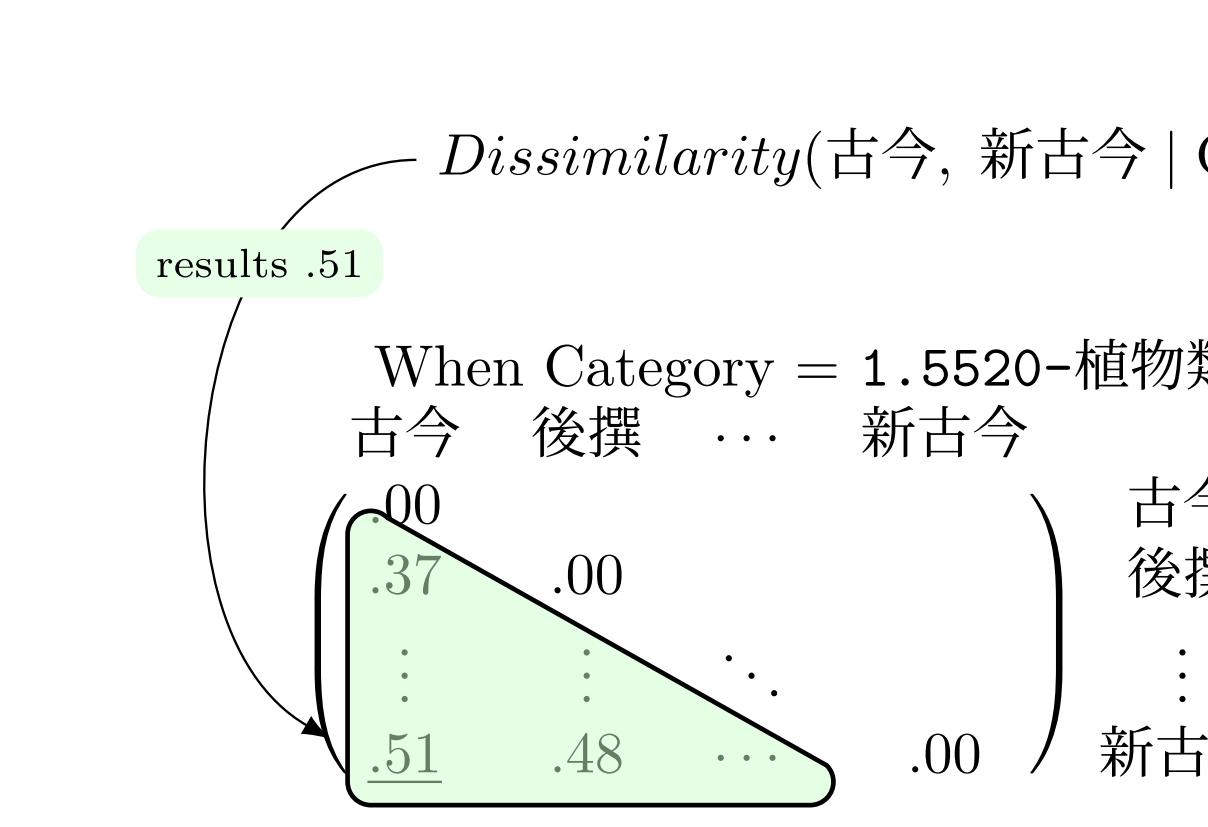
* Each column indicates a categorical vector of an anthology

Dissimilarity Calculation with city-block distance; return D_{CB}

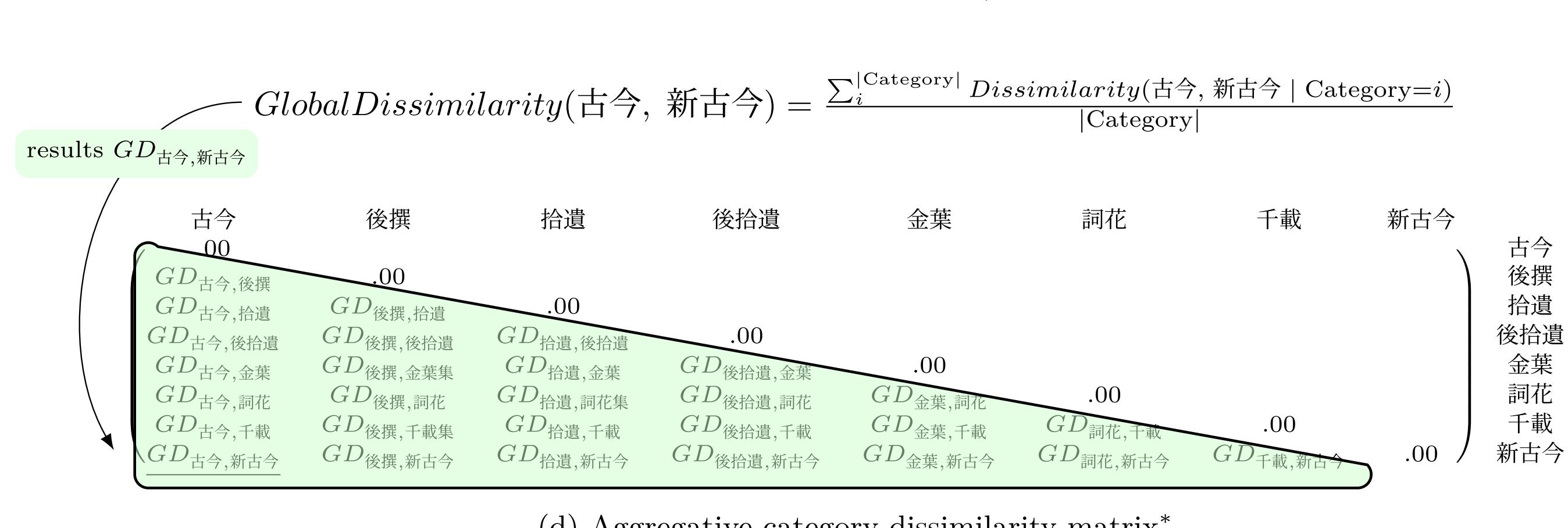
古今	後撰	…	新古今	桜花
.14	.07	…	.03	梅
.11	.07	…	.07	權
:	:	…	:	:
.00	0	…	.00	

(b) Relative frequency matrix*

* Each column indicates a categorical vector of an anthology

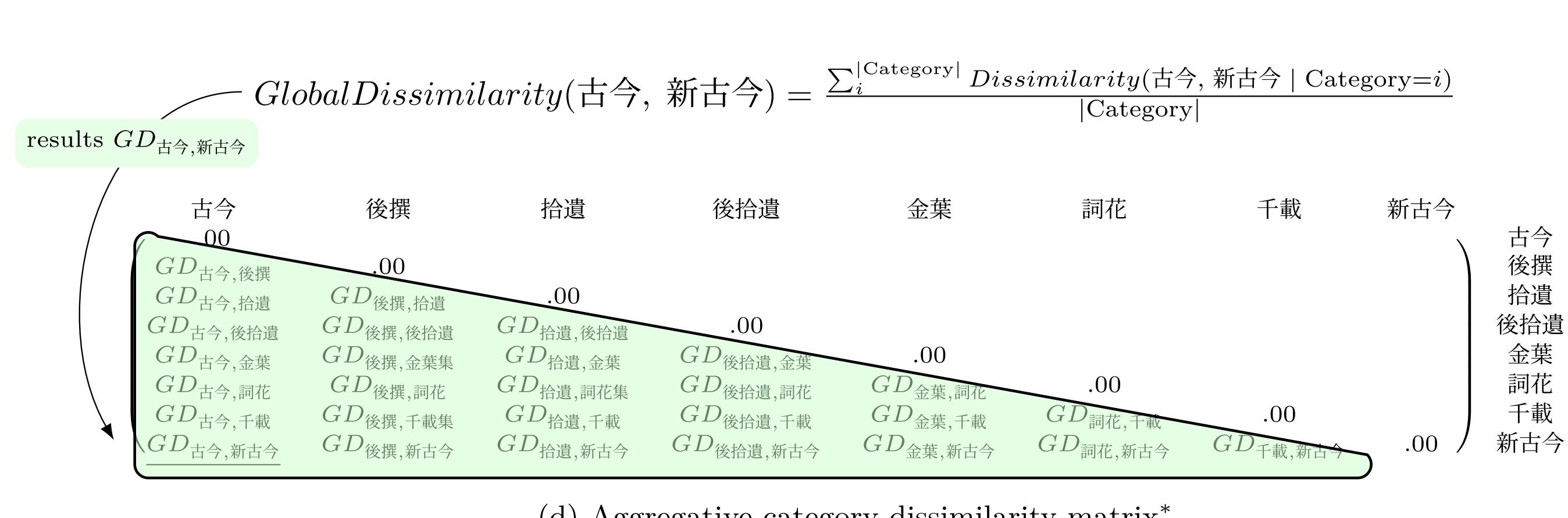


(c) Single-category (flora) dissimilarity matrix



(d) Single-category dissimilarity matrix*

* $D_{古今,後撰}^i = \text{Dissimilarity}(\text{古今}, \text{新古今} | \text{Category} = i)$, and so on



(d) Aggregative-category dissimilarity matrix*

* $GD_{古今,新古今} = \sum_i^{|\text{Category}|} \text{Dissimilarity}(\text{古今}, \text{新古今} | \text{Category} = i)$

古今
後撰
拾遺
後拾遺
金葉
詞花
千載
新古今