

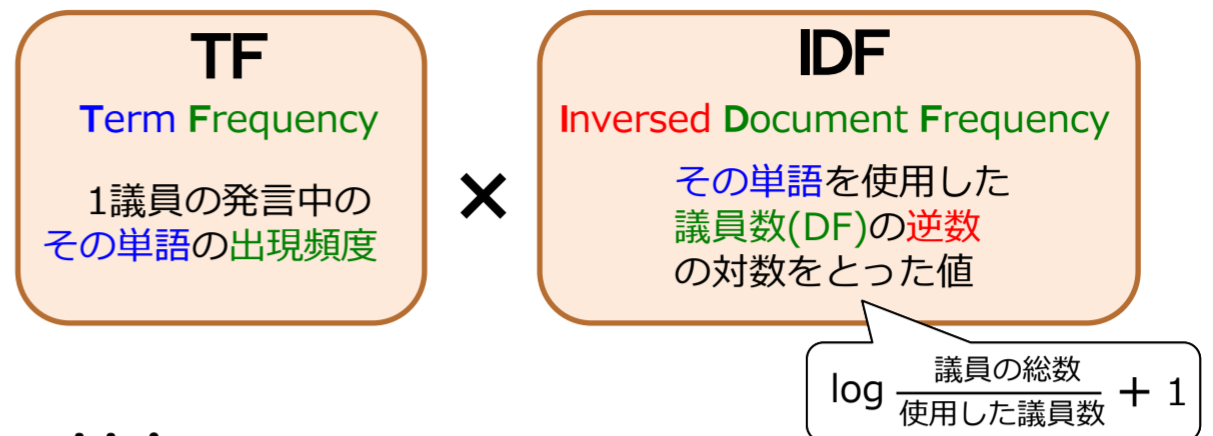
高丸研究室研究紹介 2015年度版

地方議会議員の 議会発言における 特徴語抽出

<http://www.takamaruzemi.com>
<http://www.local-politics.jp>

TF・IDF法による議会発言特徴語

- 議員発言中の**すべての単語**(名詞)に対して「TF値」と「IDF値」を求めて、掛け合わせます(TF・IDF)。
- 1議員の発言中の単語を「TF・IDF」が高い順に並べます。
- **TF・IDF** が高い単語がその議員の「**議会発言特徴語**」となります。



TF・IDFの基本

- **その議員が多く使った単語**のTF値は高くなります。
- **その単語を使った議員数が少ない**ときIDF値が高くなります。

例えば3人の議員が次のように発言したとします。



U市議会本会議における TF・IDFによる議会発言特徴語 (平成23~26年度の上位3単語)

議員氏名	議会発言の特徴語		
	1	2	3
S.S.	母子家庭	博物館	現況届け
E.K.	発生土	ランク	備蓄基地
E.A.	市競輪	あっせん	複式簿記
H.H.	備蓄避難所	給食残渣	防犯灯
M.H.①	生徒手帳	食物アレルギー	子ども
G.U.	まちなか	境界変更	単位自治会
K.M.	自転車	准看護師	景観づくり
K.S.①	取調べ	地上デジタル放送	追加計上
Y.G.	公共サービス基本法	センター	LRT
M.K.①	ネットワーク型	LRT	自転車
A.K.	無電柱化	追加計上	子ども消防団
M.T.	公明党	取り組み	取調べ
Y.K.	予算要望	高齢者	支援システム
K.K.①	有床診療所	地方消費者	追加計上
H.F.	ラジオ体操	土地区画整理事業費	土地区画整理事業
Y.O.①	市競輪	読書活動	バルーン
K.K.②	土地区画整理事業費	追加計上	会計補正予算
N.K.	再開発	食肉センター	請求代表者
N.S.	追加計上	地域自治制度	介護予防サービス
K.F.①	保育所	正社員	子ども
M.H.②	繰り入れ	空き家	償還払い
K.E.	セクター	景観的	官民較差
F.N.	市当局	自治会	自治会長
H.M.	公共建築物	太陽光	市民栄誉賞
M.W.①	議員案	AED	市介護保険
T.K.	認知症	ピロリ菌	胃がんリスク検診
N.T.	教科書	ペット	LRT
S.N.	公共交通	公共交通ネットワーク	地域内
K.K.③	総務常任委員会委員長	委員会付託	議員案
M.W.②	本委員会	追加計上	正副委員長
T.A.	LRT	共産党	保険料
K.F.②	教育委員	ペット	教育委員会
Y.I.	空き家	運営委員会	LRT
M.K.②	自校炊飯	老人クラブ	いじめ
M.Y.	暫時休憩	地域公共交通	LRT
H.T.	防災士	集会所	校務分掌
N.A.	小型家電	使用済み	創業支援
Z.A.	高架化	委員長	年長委員
Y.O.②	大震災	業務実績	建設改良事業
Y.H.	コンベンション協会	開設者	生産者
K.K.④	議員案	委員会付託	各常任委員会委員長
M.S.	委員長	年長委員	意見書
K.N.	副議長	市政発展	議会運営
K.S.②	認知症	委員長	年長委員

おとは議員の発言「**予算**」 ※自身が繰り返し使用しているが、他の議員も使っている単語

$$3 \times \left(\log \frac{3}{3} + 1 \right) = 3 \times (0 + 1) = 3$$

TF

IDF

TF・IDF

TF: 3回発言している IDF: 3議員中3人が使用している

たえ議員の発言「**子育て**」 ※自身が一定の頻度で使用し、他の議員があまり使っていない単語

$$2 \times \left(\log \frac{3}{1} + 1 \right) = 2 \times (1.1 + 1) = 4.2$$

TF

IDF

TF・IDF

TF: 2回発言している IDF: 3議員中1人が使用している

るい議員の発言「**教育**」 ※他の議員があまり使っていない単語だが、自身の使用頻度も低い単語

$$1 \times \left(\log \frac{3}{1} + 1 \right) = 1 \times (1.1 + 1) = 2.1$$

TF

IDF

TF・IDF

TF: 1回発言している IDF: 3議員中1人が使用している

- **たえ議員**の発言「**子育て**」→他の議員が発言していない単語を2回発言したため、TF・IDF値が**4.2**ととっても高くなります。
- **るい議員**の発言「**教育**」→これも他の議員は発言していない単語ですが、発言回数が1回だけのため、TF・IDF値が**2.1**と低くなりました。
- **おとは議員**の発言「**予算**」は、発言頻度はもっとも高いですが、他の議員も発言しているためIDF値が低く、TF・IDF値は**3**になりました

役職による特徴(議長等)もありますが、「どの議員でも同じ」ではなく、**議員ごとに取り組んでいる課題が異なっていることが分かります。**

本研究は科研費基盤(C) (No.26370498)「学際的応用を考慮した地方議会会議録コーパスの言語学的研究」(研究代表者・高丸圭一)の成果の一部である。