

類似文の比較による省略可能な格要素の認定

篠原 直嗣*

増山 繁*

山本 和英**

{sinohara, masuyama}@smlab.tutkie.tut.ac.jp

yamamoto@slt.atr.co.jp

*豊橋技術科学大学 知識情報工学系

** ATR 音声言語通信研究所

概要

テキストの自動要約に関する研究へのアプローチのひとつに、一文ごとに重要でない箇所を削除して要約を作成するものがある。本稿では、この一文ごとの要約での利用を想定した、文中の省略可能な格要素を認定する手法について述べる。提案する手法では、文の述語、格要素の表層格、格要素の名詞の意味属性の3つの手がかりによりコーパスから類似文を検索し、類似文の格要素の差異から省略可能な格要素を認定する。新聞記事コーパスを対象に実験を行なったところ、省略可能とするのが妥当な認定結果を65.4%の精度で得ることができた。

Recognition of Omissible Case Elements by Comparing Similar Sentences

SHINOHARA Naotsugu*, MASUYAMA Shigeru*, YAMAMOTO Kazuhide**

* Dept. of Knowledge-based Info. Eng., Toyohashi Univ. of Tech.

** ATR Spoken Language Translation Research Laboratories

Abstract

In the research on automatic summarization, there is an approach that summarizes each sentence by deleting unimportant parts. This paper presents a method for recognizing omissible case elements considering application to such summarization. The method uses three kinds of clues: predicates, surface cases, and noun's attributes of cases, in order to retrieve similar sentences in corpora, and then recognizes omissible case elements from the difference of case elements of similar sentences. Experimental results for newspaper corpora shows that results of recognition are valid to be omissible at the rate of 65.4%.

1 はじめに

近年、テキストの自動要約に関する研究が盛んに行なわれている [4]。それらの中に、文中の重要でない箇所削除や、語句の言い換えによって一文ごとに文を短縮して要約を作成するというアプローチがあり、これまでいくつかの手法が提案されている。一文ごとの要約は、重要文抽出による要約に比べて、情報の欠落を少なくすることができるという利点がある。

山本ら [10] は、一文ごとの要約手法として、ヒューリスティックスに基づいた連体修飾節などの削除を提案している。この手法は、重要文抽出による要約結果をさらに要約するという位置付けで提案されているが、単独で用いることも可能である。若尾ら [11] や山崎ら [9] は、人手で作成された字幕とその元となったニュース原稿を人手で比較し、それによって作成した言い換え規則を用いた要約手法を提案している。また、加藤ら [5] は記事ごとに対応のとれたニュース原稿と字幕放送の原稿を用いて、言い換えに関する要約知識を自動獲得する研究を行なっている。

ところが、これらの手法には次のような問題点がある。まず、不要箇所の削除や言い換えに関する規則を人手で作成するには多大な労力が掛かり、網羅性などの問題も残ることが挙げられる。また、加藤らを使用したような原文と要約文の対応をとれたコーパスは一般には大量に存在しておらず、そのようなコーパスを人手で作成するには多大な労力が掛かるため、大規模な要約知識を得るのは困難である。

このような理由から、原文と要約文の対応がとれていないコーパスから、要約に利用できる言語知識を自動獲得するというタスクの実現には十分な意味があると考えた。ここでコーパスとして想定するのは、形態素情報などの付与されていない一般の電子化テキストである。したがって CD-ROM などで提供されている新聞記事のバックナンバーや電子辞書、WWW 上で公開されている文書などを利用することができ、コーパスの大規模化も比較的容易に実現可能である。

本稿では、一文ごとの要約への新しいアプローチである、文中の省略可能な格要素の削除による要約を目的とした、コーパスを用いた省略可能な格要素の認定法を提案する。提案する手法では、コーパスから検索した類似する文の対を使用し、それらの格要素の差異から省略可能な格要素を認定する。これ

は「類似する文のどちらか一方にしか出現しない格要素は省略可能である」という仮定に基づいている。以下、第 2 章では省略可能な格要素の削除による要約について述べ、既存の言語知識を用いても実現できない省略可能な格要素の認定に関する事例を取り上げる。第 3 章では第 2 章で取り上げた事例にも対応できる、省略可能な格要素の認定手法について述べ、第 4 章では提案する手法を組み入れて作成したシステムについて述べる。次に第 5 章で、作成したシステムを用いた実験について述べ、実験結果について考察する。

2 格要素の削除による要約

この章では、格要素の削除による要約について説明する。まず、文の構造に関するモデルの一つとして、格文法または結合価文法と呼ばれる文法理論がある。このモデルは述語を中心に文の構造をとらえるものである。そして述語にいくつかの格要素がむすびつくことによって、文が形成されると考える。述語となり得る語と、その語にむすびつく格要素による制約条件を記述したものを格フレームという。ここで、格フレームに存在する格要素に対して必須格と任意格の情報が与えられている場合、あるいは格フレームに存在する格要素を必須格とし、それ以外の格要素を任意格にするとした場合には、格フレームを用いることで、本稿の目的である「省略可能な格要素の削除による要約」が実現できる。図 1 に格フレームの例と、その格フレームを用いて可能な格要素の削除による要約の例を示す。

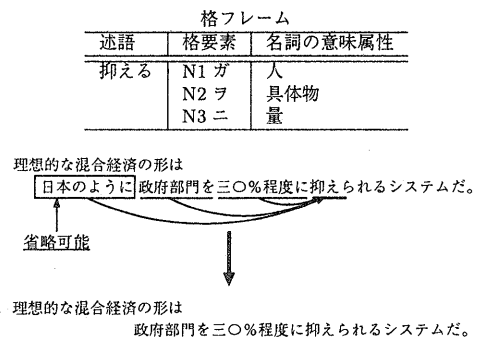


図 1: 格要素の削除による要約

この図では、述語「抑える」に係る3つの格要素のうち、格フレームと照合されない格要素「日本のように」が省略可能となる。このような処理は、IPAL[7]の基本動詞辞書や日本語語彙大系[1]の構文意味辞書といった、格要素と述語の関係を人手で収集したものをを用いることにより実現可能である。また現在までに、格フレームの獲得に関する研究は数多く行なわれてきており[3]、これらの研究で提案されている手法により獲得した格フレームを用いても実現可能であると思われる。

しかし、図2で示すように、格フレームを用いても省略可能と認定することができない格要素が存在する。

対策室には企画調整班と事業推進班を置き、
それぞれが各地方機関とともに施策の検討を進めていく方針。

図2: 格フレームでは認定できない例

図2において「進める」に係る格要素「それぞれが」は、人間ならば省略可能と認定することができる。ところが、このような認定は格フレームの情報を用いても行なうことが難しい。例えばIPALの基本動詞辞書において、動詞「進める」の格フレームに関する記述は表1のようになっている。このような格フレームから図2で示した省略可能な格要素の認定をすることは困難と思われる。

格フレーム	文例
N1ガN2ヲ(N3ニ／へ)	彼は 船を 沖へ 進めた。
N1ガN2ヲN3ニ	彼は 娘を 大学に 進めた。
N1ガN2ヲ	彼は 会の準備を 進めている。
N1ガN2ヲ	政府は 国の産業を 進めている。

表1: 「進める」の格フレーム

そこで本稿では図2で示したような、格フレームを用いても省略可能と認定できない事例が存在することに注目し、このような事例に対しても有効な、省略可能な格要素の認定手法について述べる。

3 提案する手法

本稿で提案する手法の基本的なアイデアは、コーパス中に存在する類似文を比較することで、何らかの要約知識が得られるのではないかと予想であ

る。そして2章で取り上げた事例についても、図3で示すように、文1に対する類似文として文2がコーパスから検索されたならば、「進める」に係る格要素の差異から文1の「進める」に係る格要素「それぞれが」を削除可能と認定することができる。

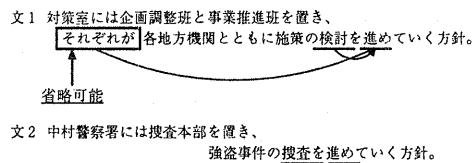


図3: 格要素の差異による認定

またこのようにして認定した結果を、一文ごとの要約で利用するために知識ベース化する際、動詞「進める」に係る「それぞれが」を省略可能とはせず、述語列「置く、進める」の述語「進める」に係る格要素「それぞれが」を省略可能とする。

4 システムの構成

3章で述べた手法を用いて省略可能な格要素を認定するためのシステムを作成した。システムの構成は図4に示すように、大きく分けて次の三つの部分で構成される。

1. コーパスからの文例データベースの構築
2. 文の述語列に基づいた類似文の検索
3. 類似文の比較から省略可能な格要素を認定

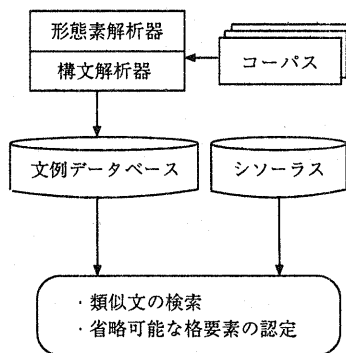


図4: システムの構成

4.1 文例データベースの構築

本稿で提案する手法において、検索したい類似文は、文の述語列が同じでなければならないという条件がある。そのため、あらかじめコーパス中の文に対し、文の述語列、および各述語に係る格要素を求めておく。そして、一つのレコードが記事番号、文番号、文、文の述語列、述語数、各述語に係る格要素の六つの属性から成る文例データベースを構築する。

ここで述語の抽出には、形態素解析器『茶釜』¹を使用した。『茶釜』の解析結果において、動詞、形容詞、形容動詞、サ変名詞に該当するものに対し、ヒューリスティックスを適用することにより述語を抽出する。また、述語に係る格要素は形態素解析器 JUMAN²、および構文解析器 KNP³の解析結果から抽出する。文例データベースのレコードの例を表2に示す。

記事番号	920130-0046
文番号	5
文	アンケートは会員企業九百五十二社を対象に 昨年夏に実施し、三百五十七社から回答を得た。
述語列	実施, 得る
述語数	2
格要素	(対象に, 夏に)(三百五十七社から, 回答を)

表 2: 文例データベースのレコードの例

4.2 類似文の検索

類似文検索については、用例に基づく機械翻訳や翻訳支援に関する研究の中で様々な手法が提案されている。隅田ら[8]は、検索入力文を語句の置き換えや削除により一般化し、データベース中の用例との照合から類似文を検索する手法を提案している。また、宇津呂[2]は類似文を効率的に検索する手法として、類似度テンプレートと検索入力文から検索質問を生成し、類似文を検索するという手法を提案している。本稿では、これら従来の類似文検索法を参考にして類似文検索を行なう。

一般に、文の類似度の指標には、構文構造の類似度と意味的な類似度があると考えられる。本稿での類似文検索では、構文構造の類似度を求めるために「文の述語列」と「述語に係る格要素の種類」を使用す

る、また意味的な類似度を求めるために「述語に係る格要素の名詞部分の意味属性」を使用する。

4.2.1 述語列による類似文集合の抽出

まず、文例データベースから述語列ごとに文例を取り出し類似文集合とする。次に類似文集合ごとに類似文集合中のある文(以下、対象文)とその他の文との類似度を計算し、類似度の順位付けを行なう。ここで類似度の順位付けには、以下で示す二つの類似度を用いる。

4.2.2 格要素の種類による類似度

まず一つ目として、各述語に係っている格要素の格助詞の重複から類似度 α を求める。

$$\alpha = \frac{2D_c}{C_1 + C_2} \quad (0 \leq \alpha \leq 1)$$

C_1 : 対象文の格要素数 C_2 : 類似文集合の文の格要素数
 D_c : 重複する格要素数

4.2.3 格要素の名詞部分の意味属性による類似度

各述語の格要素の種類による順位付けのみでは、同じ類似度の文例が複数存在してしまう。そこで二つ目の類似度として、格要素の名詞部分から類似度を求め、さらに順位付けを行なう。

本手法では EDR 電子化辞書⁴の「日本語単語辞書」に基づいて名詞に割り当てた意味属性と、「概念体系辞書」から得た上位概念の意味属性を用いて、名詞の類似度を求める。

名詞の類似度を測る方法としては、意味属性体系上での共通親属性の位置や、両意味属性間のパスの長さから類似度を求める方法が考えられる。しかし、一般に名詞には複数の意味属性を割り当てることができる。そのため、名詞の類似度を求めるために、その名詞がどの意味属性の名詞として使われているのかを、文脈情報などから一意に決定しなければならない。本手法では、この多義性の問題に立ち入らず、崔らの研究[6]で提案されている手法と同様の方針をとり、名詞に割り当てられた複数の意味属性から総合的に名詞の類似度を求める。手順を以下に示す。

求める類似度を β とする。まず同じ格助詞を伴う格要素の、名詞の意味属性リストを単語辞書との照合から得る。そして、それぞれの意味属性の上位概念の意味属性リストを概念体系辞書から得る。次に両

¹<http://cl.aist-nara.ac.jp/lab/nlt/chasen/index.html>

²<http://pine.kuee.kyoto-u.ac.jp/nl-resource/juman.html>

³<http://pine.kuee.kyoto-u.ac.jp/nl-resource/KNP.html>

⁴http://www.iiijnet.or.jp/edr/J_index.html

名詞のもつ意味属性から名詞間の関係を [6] で提案されているように、次の三つに分類する。例を図 5 に示す。

同一関係 両名詞の意味属性が同じ

同義関係 両名詞の意味属性のうち共通なものが一つ以上存在する

類似関係 両名詞の意味属性のうち共通なものがなく、上位概念に共通な概念が存在する

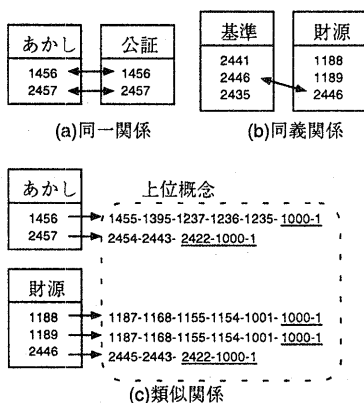


図 5: 名詞間の関係

この三つの関係に基づき、名詞間の類似度 β を求める。同一関係の場合は $\beta = 1$ とし、同一関係でない場合は 同義関係による類似度 a と、類似関係による類似度 b を次式により求める。

$$a = \frac{2D_a}{A_1 + A_2} \quad (0 \leq a \leq 1)$$

A_n : 名詞 n の意味属性数 ($n = 1, 2$)

D_a : 重複する意味属性数

$$b = \frac{1}{N_1 N_2} \sum_{i,j=0}^{N_1, N_2} \frac{2D_{ij}}{N_{1i} + N_{2j}} \quad (0 \leq b \leq 1)$$

N_n : 名詞 n の意味属性数

N_{ni} : 名詞 n の意味属性 i の上位概念数

D_{ij} : 意味属性 i, j の上位概念数の重複数

求めた類似度 a, b を使用し、次式により同一関係でない場合の名詞間の類似度 β を求める。

$$\beta = 1 - e^{-(a+b)} \quad (0 \leq \beta \leq 1)$$

以上の手法により求めた名詞の類似度の例を表 3 に示す。

表 3: 名詞の類似度

類似度	名詞 1	名詞 2
1.0000	イメージ	概念
1.0000	自由自在	意外
0.8647	%	倍
0.8111	政策	対策
0.7880	生涯	青春
:	:	:
0.1192	承認	市民
0.1175	論者	回答
0.1153	展開	ビール
0.1143	回答	市民
0.1110	支援	酒

4.2.4 類似文の順位付け

類似度 α, β から、次式により最終的な類似度 δ を求め、類似文の順位付けを行なう。

$$\delta = w\alpha + \beta \quad w: \text{重み}$$

ここで、意味的な類似度を表す β は、構文構造の類似度を表す α が等しい文例をさらに順位付けするために使用する。したがって、重みを $w = 100$ としている。

4.3 省略可能な格要素の認定

前節の処理により、対象文と対象文との類似度で順位付けされた類似文集合が得られる。次に、3 章で示したように対象文と類似文との格要素の差異から、省略可能な格要素を認定する。この時、対象文との類似度の高い類似文から順番に、省略可能な格要素を認定していき、認定が成功した場合には、その対象文に対する処理を終了する。認定に失敗した場合には、次に類似度の高い文との処理に移り、認定が成功するか、類似文がなくなるまで繰り返す。

ここで、図 6 で示すように照合できなかった格要素数が二つ以上の場合には、コーパスのスパース性により類似文がコーパス中に存在しなかったとみなし、認定を失敗とする。

大蔵省は来週にも金利調整審議会を開き、
 預貯金金利の改定を正式に決めた。
 同事業団はこのほど役員会を開き、
 来年度の事業計画を決めた。

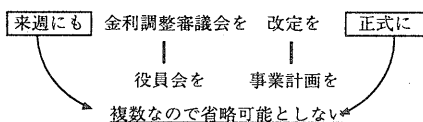


図 6: 認定を失敗とする例

5 実験と考察

5.1 実験

実験には、日本経済新聞 CD-ROM 版 92 年 1 月から 6 月の 96422 記事に対して、直接引用表現を含む文などを除外することによって得た 407353 文を使用した。まず最初に、文例データベースを構築したところ、述語列の異なり数は 226709 種類、出現回数が 2 以上の述語列数は 21408 種類であった。

次に、述語数が 2 以上の述語列 50 種、文例 2171 個を対象にして、提案手法を用いて省略可能な格要素の認定を行ない、得られた認定結果 1190 個について手作業による評価を行なった。ここで認定数 1190 個という数は、文中の格要素に冗長性がない文を対象文とした場合には省略可能な格要素が認定されないことを考慮すれば、妥当な数であるといえる。そして評価の結果、省略可能とするのが妥当であると判定されたものは 778 個となり、65.4%の精度で省略可能な格要素の認定ができたことを確認した。

5.2 妥当であると判定されたもの

妥当であると判定されたもの 778 個を表層格の種類により分類すると、デ格が 306 個、二格が 278 個、ト格が 71 個、それ以外が 123 個となった。以下では、それぞれの表層格の場合の例を挙げ、その特徴について考察する。

デ格、二格、ト格

認定結果のうち、デ格、二格、ト格のものは、値段、数、場所、手段の意味をもつものがほとんどであった。表 4 に例を示す。これらについては、シソーラスを用いた一般化が可能だと思われる。そのため一文ごとの要約での利用も比較的容易に可能である。

表 4: デ格、二格、ト格の例

述語列	係り先の述語	省略可能な格要素
比べる, 減少	減少	約三十兆円で
比べる, 増加	増加	実質で
開く, 決める	開く	札幌市内で
持つ, 実施	実施	アンケート方式で
比べる, 減少	減少	約四分の一に
受ける, 実施	実施	二十四年ぶりに
比べる, 減少	減少	十六億九千万円と
受ける, 上昇	上昇	一段と

それ以外

妥当であると判定されたものに対する割合は 15.8%と低いが、ここに分類されたものには、人手による獲得や、シソーラスによる一般化などが困難だと思われるものが含まれており、本稿で提案する手法で得られる言語知識の中で最も有用であると考えられる。表 5 に例を示す。

表 5: それ以外の例

述語列	係り先の述語	省略可能な格要素
実施, 発表	実施	投票が
着工, 完成	完成	施設が
みる, 調べる	調べる	事故原因を
受ける, 入る	受ける	警察庁人事課長から
受ける, 先行	先行	朝方から
変更, 発表	変更	現物受け渡しから

次に、一般に用言とのむすびつきが強いとされるガ格とヲ格の格要素について、省略可能とされた 2 つの例を以下に示す。

[例 1]
 タイ内務省は、全国百四十二の選挙区で **投票が**
 実施された総選挙の確定議席を正式発表した。

[例 2]
 現場にはスリップ跡がなかったため、
 同署は飲酒運転の疑いもあるとみて **事故原因を** 調べている。

投票が : 省略可能な格要素

図 7: ガ格とヲ格の例

ここで、これらの格要素が省略可能となる理由を考えてみても、文脈から省略可能という以外には、明確な理由付けが困難であることが分かる。したがって、実際に人間が書いた文との差異から、省略可能な格要素を認定するという本稿で提案したアプローチは、このような格要素を省略可能と認定する手法として有効であると考えられる。

5.3 妥当でないと判定されたもの

妥当でないと判定されたもの412個を表層格の種類により分類すると、ガ格が95個、ヲ格が58個、ニ格が100個、その他が159個となった。これらの認定の誤りを改善するための対策としては、省略可能と認定された格要素が、係り先の述語に係る唯一の格要素の場合には、その格要素を省略可能と認定しないということが考えられる。格要素と結合していない述語は、文中での役割を果たすことができず、文全体の意味が取れなくなることが多いためである。

また、妥当でないと判定された中にも、前後の文脈を考慮した場合には、省略可能であると判定できるものがあり、それらの例をいかに言語知識として利用可能なものにするのかも、今後の検討課題である。

表 6: 妥当でないと判定されたものの例

述語列	係り先の述語	省略可能な格要素
比べる, 減少 発表, 増加 実施, 発表 開く, 決める 受ける, 実施	減少 増加 実施 決める 受ける	輸送量が 一般会計規模が 機構改革を 候補者に 株主総会で

6 おわりに

本稿では省略可能な格要素の削除による一文ごとの要約を目的とした、省略可能な格要素を認定する手法について述べた。新聞記事コーパスを対象とした実験の結果、一文ごとの要約に利用可能な認定結果を比較的高い精度で得ることができた。そして実験結果から、格フレームを用いる手法では省略可能と認定するのが困難な事例についても、本手法が有効であることが確認された。

今後の課題としては、認定の精度を向上させる方法を検討し、認定結果の知識ベース化を行なうことで、当初の目的である一文ごとの要約への適用を試みることが挙げられる。また、今回取り上げた事例について言語学的な分析をすることで、別のアプローチによる省略可能な格要素の認定についても検討したい。

謝辞

言語データとして、日本経済新聞 CD-ROM 版の使用を許可して頂いた日本経済新聞社に深謝致します。

参考文献

- [1] 池原悟, 宮崎正弘, 白井諭, 横尾昭男, 中岩浩巳, 小倉健太郎, 大山芳史, 林良彦 (編) (1997). 日本語語彙大系. 岩波書店.
- [2] 宇津呂 武仁 (1994). “類似度テンプレートを利用した高速類似用例検索.” 情報処理学会研究報告 NL-103-5, pp.33-40.
- [3] 宇津呂 武仁, 松本 裕治 (1995). “コーパスを用いた言語知識の獲得.” 人工知能学会誌, vol.10, no.2, pp.197-204.
- [4] 奥村 学, 難波 英嗣 (1999). “テキスト自動要約に関する研究動向.” 自然言語処理, vol.6, no.5, pp.1-25.
- [5] 加藤直人, 浦谷則好 (1999). “局所的要約知識の自動獲得手法.” 自然言語処理, vol.6, no.7, pp.73-92.
- [6] 崔 進, 小松英二, 安原 宏 (1993). “EDR 電子化辞書を用いた単語類似度計算法.” 情報処理学会研究報告 NL-93-1, pp.1-6.
- [7] 情報処理振興事業協会 技術センター (1987). 計算機用日本語基本動詞辞書 IPAL (Basic Verbs), 一解説編一.
- [8] 隅田英一郎, 提 豊 (1991). “翻訳支援のための類似用例の実用的検索法.” 電子情報通信学会論文誌, vol.J-74-D-II, no.10, pp.1437-1447.
- [9] 山崎 邦子, 三上 真, 増山 繁, 中川聖一 (1998). “聴覚障害者用字幕生成のための言い換えによるニュース文要約.” 言語処理学会第4回年次大会発表論文集, pp.646-649.
- [10] 山本 和英, 増山 繁, 内藤 昭三 (1995). “文章内構造を複合的に利用した論説文要約システム GREEN.” 自然言語処理, vol.2, no.1, pp.39-55.
- [11] 若尾 孝博, 江原 暉将, 白井 克彦 (1997). “テレビニュース番組の字幕に見られる要約の手法.” 情報処理学会研究報告 NL-122-13, pp.83-89.

付録

以下に、実験結果の人手による分類の詳細を示す。

実験結果

述語列	文例数	認定数	妥当なもの					妥当でないもの					妥当な割合(%)
			ア格	ニ格	ト格	その他	計	カ格	ヲ格	ニ格	その他	計	
引き下げる, 発表	30	21	0	15	0	6	21	0	0	0	0	0	100.0
比べる, 終了	27	7	7	0	0	0	7	0	0	0	0	0	100.0
出す, 倒産	46	3	2	1	0	0	3	0	0	0	0	0	100.0
下げる, 安い	57	1	0	1	0	0	1	0	0	0	0	0	100.0
運ばれる, 死亡	43	27	17	8	0	1	26	0	0	0	1	1	96.3
発表, 増加	35	26	24	0	0	1	25	1	0	0	0	1	96.2
出す, 呼びかける	43	22	10	10	0	1	21	0	0	0	1	1	95.5
集める, 開く	29	17	9	4	0	3	16	0	0	0	1	1	94.1
招く, 開く	28	15	8	3	0	3	14	0	1	0	0	1	93.3
比べる, 減少	110	77	24	7	30	7	68	5	0	0	4	9	90.9
実施, まとめる	60	54	5	40	0	1	46	2	4	1	1	8	85.2
比べる, 増加	77	43	29	4	2	1	36	6	0	0	1	7	83.7
実施, 発表	64	49	9	24	1	7	41	0	6	1	1	8	83.7
発表, 比べる, 減少	24	23	9	0	9	1	19	4	0	0	0	4	82.6
発表, 減少	26	19	11	1	3	0	15	2	0	0	2	4	78.9
建設, 始める	28	18	1	5	0	8	14	2	1	0	1	4	77.8
開く, 決める	101	67	26	17	6	3	52	1	4	4	6	15	77.6
発表, 上昇	51	32	20	2	0	2	24	0	0	0	8	8	75.0
比べる, とどまる	58	23	9	2	3	3	17	1	0	0	5	6	73.9
開く, まとめる	31	19	4	6	0	4	14	1	2	0	2	5	73.7
比べる, 減る	25	14	2	4	0	4	10	2	0	0	2	4	71.4
比べる, 増える	52	30	8	9	0	4	21	2	0	2	5	9	70.0
開く, 決定	87	69	6	38	3	1	48	0	3	17	1	21	69.6
比べる, 下落	26	13	2	1	6	0	9	1	0	0	3	4	69.2
比べる, 上昇	58	38	19	2	0	5	26	5	0	1	6	12	68.4
受ける, 実施	26	14	0	1	0	8	9	0	3	0	2	5	64.3
比べる, 拡大	26	10	2	3	0	1	6	2	0	0	2	4	60.0
開く, 協議	59	39	13	6	2	2	23	0	14	0	2	16	59.0
向ける, 進める	34	12	1	2	3	1	7	2	0	0	3	5	58.3
比べる, 低下	54	29	9	2	0	5	16	4	0	3	6	13	55.2
進める, 目指す	28	14	2	4	0	1	7	0	0	2	5	7	50.0
入る, いう	24	8	0	4	0	0	4	0	0	2	2	4	50.0
受ける, 入る	27	6	0	0	0	3	3	1	0	0	2	3	50.0
出資, 設立	38	22	1	4	1	3	9	3	3	2	5	13	40.9
着工, 完成	46	37	0	12	0	3	15	1	1	17	3	22	40.5
着工, 目指す	74	65	1	16	0	7	24	0	0	37	4	41	36.9
実施, 得る	76	44	8	5	1	2	16	0	3	0	25	28	36.4
設立, 始める	32	17	1	4	0	2	7	3	1	3	3	10	35.3
かける, 建設	43	21	1	2	0	4	7	2	6	3	3	14	33.3
上昇, 低下	27	6	0	0	0	2	2	2	0	0	2	4	33.3
置く, 就任	29	3	1	0	0	0	1	0	0	1	1	2	33.3
みる, 調べる	51	22	1	0	0	6	7	11	0	1	3	15	31.8
下回る, とどまる	34	10	1	1	1	0	3	5	0	0	2	7	30.0
出る, いう	30	11	0	1	0	2	3	2	1	2	3	8	27.3
反発, 始まる	27	11	1	2	0	0	3	2	1	0	5	8	27.3
建設, 完成	30	24	0	3	0	3	6	15	2	1	0	18	25.0
続落, 始まる	29	9	0	1	0	1	2	2	0	0	5	7	22.2
続伸, 始まる	25	9	0	1	0	1	2	2	2	0	3	7	22.2
比べる, 終える	60	18	2	0	0	0	2	0	0	0	16	16	11.1
反落, 始まる	26	2	0	0	0	0	0	1	0	0	1	2	0.0
計	2171	1190	306	278	71	123	778	95	58	100	159	412	65.4