

類似度を用いた可視化情報の提示による 問診票回答収集効果の検証

福島 拓^{1,a)} 前島 康宏² 甲斐 充彦¹

概要：現在、グローバル化による多言語間コミュニケーションの機会が増加している。しかし、多言語間での正確な情報共有は十分に行われていない。正確な多言語支援が求められる場では、用例対訳が多く用いられており、用例対訳の収集や、用例対訳を用いた自由な対話支援のために応答用例対と呼ばれる概念が提案されている。しかし、用例対訳や応答用例対の収集においては、作成者が既に収集された用例対訳や応答用例対を網羅的に把握することが難しく、用例対訳や応答用例対の収集の妨げになっていた。そこで本稿では、医療分野の問診票入力場面で使用される応答用例対の収集を目的とした、問診票回答収集システムの構築について述べる。本システムでは、収集された応答用例対をグラフを用いて可視化することで、利用者が収集済みの応答用例対の網羅的な把握の支援を目指す。

Verification of the Effect to Collect Responses of Interview-sheet by Presentation of Visualized Information using Similarity

Taku Fukushima^{1,a)} Yasuhiro Maejima² Atsuhiko Kai¹

1. はじめに

近年の世界的なグローバル化により多言語間コミュニケーションの機会が増加している。訪日外国人数も年々増加傾向にあり、2014年は過去最高の約1,415万人にのぼっている [1]。しかし、すべての訪日外国人が日本語を理解しているとは言い難い。また、一般に多言語を十分に習得することは非常に難しく、母語以外の言語によるコミュニケーションは困難なこともあり [2], [3], [4]、非母語によるコミュニケーションは十分に行うことができない。

非母語によるコミュニケーションで影響が顕著に現れる分野の1つに医療がある。医療分野では、わずかなコミュニケーション不足で医療ミスが発生する恐れがあるため、適切な支援が求められている。

そこで、多言語対応の医療支援システムの開発が多く行われている [5], [6], [7], [8]。これらのシステムでは、正確な多言語変換が可能な用例対訳が用いられている。用例対訳とは、用例を多言語に正確に翻訳したコーパスのことを

指し、「保険証はお持ちですか?」「はい」「いいえ」などの利用現場で使用される言葉を多言語で提供することができる。この用例対訳を用いて、利用者が適切な質問やその回答を使用することで、正確な多言語対話が可能となる。

我々は用例対訳の収集、共有を目的とした多言語用例対訳共有システム TackPad (タックパッド) の開発を行っている [9]。収集した用例対訳は、正確性評価を行った後、多言語対応医療支援システムへの提供を目指している。また、用例対訳を用いた質問と回答の対である「応答用例対」を用いることでより正確な多言語間コミュニケーション支援を目指している [10]。応答用例対は用例対訳を用いた質問文と用例対訳を用いた回答文、回答文と入れ替えても同じ意味となる文である類似文から構成されている。回答文は質問文に対する適切な回答を提示するために、類似文は回答文の翻訳コストを低減するために、それぞれ用いられている。

応答用例対の例を図 1 に示す。応答用例対は、質問文 (例:「どうしましたか?」) に対して適切な回答文群 (例:「頭が痛いです」など) を保持することで、利用者に適切な回答群を提示することができる。また、類似文群 (例:「頭

¹ 静岡大学大学院総合科学技術研究科

² 静岡大学工学部

^{a)} fukushima.taku@shizuoka.ac.jp

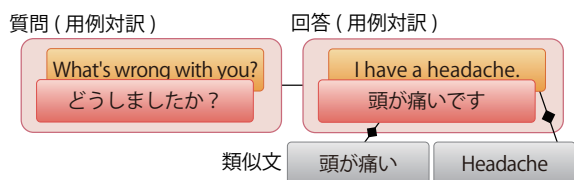


図 1 応答用例対の例

が痛い」など)も併せて保持している。類似文は回答文と入れ替え可能であるため、類似文の対訳を作成せずに多言語に変換することができる^{*1}。図 1 では 2 言語のみの例であるが、実際の用例対訳は多言語に翻訳されている。類似文を応答用例対に併せ持つことで、類似文の翻訳が不要になるため、翻訳コストの低減が可能となる。

しかし、用例対訳や、用例対訳を用いた応答用例対は十分に作成されていない。その原因の一つとして、用例対訳や応答用例対を収集するシステムにおいて、既に収集された用例対訳や応答用例対にどのような内容の文が含まれているかを把握することが困難であることが考えられる。従来システム [9] でもカテゴリ分けやタグによる分類が行われているが、すべての文に対して自動的に適切なカテゴリやタグを付与することは難しい。このことにより、利用者が収集済み用例対訳などの全体像をつかむことが困難となり、用例対訳などの収集の妨げになっていた。

そこで本稿では、医療分野の問診票入力場面で使用される応答用例対の収集を目的とした、回答文と類似文を収集するシステムについて述べる。本システムでは、収集された応答用例対を可視化することで、利用者が応答用例対の全体像を把握することを支援する。なお、問診票はある特定の質問文に対して患者が回答文を入力する。このため、本稿では応答用例対のうち質問文は固定とし、その質問文に対応する回答文と類似文を収集することで、多言語問診票の入力場面における応答用例対作成性能の向上を目指す。

2. 関連研究

入力された自由文を全て多言語に翻訳が可能な機械翻訳を用いて、子供向けの機械翻訳 [11] やビジネスでの多言語支援 [12] など、様々な分野で多言語間コミュニケーション支援が行われている。しかし、機械翻訳の精度は年々向上しているものの、正確性が求められる医療分野でそのまま利用可能な精度には達していない [13]。また、機械翻訳はルールや統計データに基づいて動的な翻訳を行うため [14]、すべての対訳の正確性を確保することはできない。

そこで現在、正確性が求められる分野では用例対訳による支援が多く行われている。医療分野で用例対訳を用いたシステムとして、宮部らが開発した多言語医療受付支援システム M^3 がある [5]。 M^3 はタッチパネルで操作可能とし

^{*1} 図 1 の場合、類似文「頭が痛い」を使用して回答文の翻訳である「I have a headache.」に翻訳することができる。

たシステムで、対話機能、外国人患者の受診支援機能(問診機能、受診科選択機能など)を有している。また、杉田らは用例対訳を用いて多言語問診を携帯電話上で実現している [6]。これらのシステムでは、用例対訳を用いて症状ごとにフローをあらかじめ作成している。このため、医療機関が想定した症状について正確な対応が可能である。

前述のシステムで使用する用例対訳の収集・共有に関する研究も行われている。Tatoeba Project では、Web 上で 130 言語の日常会話に関する用例対訳を収集している [15]。我々は医療分野で使用する用例対訳の収集を目的とした、多言語用例対訳共有システム TackPad [9] の開発を行っている。また、用例対訳の登録者や用例対訳の正確性評価者のモチベーション維持を目的としたシステムの開発を行っている [16], [17]。文献 [16] では楽しさや達成感などの要素を用いて用例対訳の正確性評価作業のモチベーション維持を、文献 [17] では携帯端末の通知機能を用いてシステムに登録した用例がどのような貢献を果たしたかを利用者に知らせることでモチベーション維持をそれぞれ行っている。しかし、これらの研究では登録された用例対訳をシステム利用者が網羅的に確認することができていない。本稿では収集済みの文の関係を可視化することで、利用者が収集済みの文を網羅的に確認可能とすることを旨とする。

3. 問診票回答収集システム

本章では、構築した問診票回答収集システムについて述べる。本システムは、回答文登録機能と回答文可視化機能から構成されている。また、本システムは Web 上で実装を行っている。以降の各節で各機能の詳細について述べる。

3.1 回答文登録機能

本節では、回答文登録機能について述べる。以下に、本機能の利用手順を示す。

- (1) 利用者は問診票の質問に対する回答をシステムに入力する。
- (2) システムは利用者が入力した文と文字列が類似している文を登録済みの回答文の中から検索する。
- (3) システムは利用者に (2) の類似している文を提示し、利用者に類似文判定を依頼する。利用者が類似と判断した場合は、質問文と提示した回答文の類似文として登録する。利用者が類似していないと判断した場合は、質問文に対する新たな回答文として登録する。

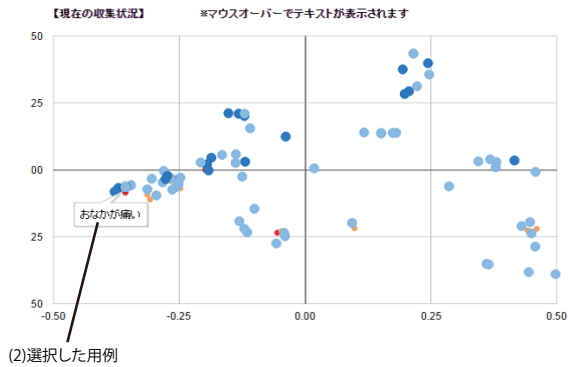
(1) では、一度に最大 5 文の回答文の入力を可能とした。

(2) では、N-gram に基づく用例対訳検索手法 [18] を利用している。文献 [18] では、用例を言語によって文字 2-gram もしくは文字 4-gram に分割し、検索文字列との共起を調べることで多言語の類似文検索を実現している。また、tf-idf を用いることで、重要単語への重み付けを行っている。本システムでは文字 4-gram を用いた。

Q.どのような症状がありますか？(外科での質問)

(1)質問文
上記の質問文に対する入力状況下のグラフで表しています。
青色が質問文に対する回答、赤色がその回答に類似している文章です。
文の配置は、文章間の類似度の高いほど近い配置となっています。

●入力されている回答です ●あなたが入力した回答です
●回答の類似文です ●あなたが入力した類似文



あなたが登録した回答数/全回答数 = 19/72
あなたが登録した類似文の数/全類似文の数 = 4/14

この質問の回答文を作成する | ホームに戻る

図 2 回答文の可視化例

(3)の類似文判定では、応答用例対の定義 [10] に基づき、提示された回答文と利用者が入力した文を入れ替えても、質問文に対する答えとして意味が変わらないかどうかの判定を利用者に依頼する。同じ意味と判定された場合は既存の応答用例対の類似文として、別の意味と判定された場合は新たな応答用例対(回答文)として、それぞれ登録する。なお、本システムでは単言語での用例の収集を目的としている。本来であれば、類似している用例群(回答群)の中から代表となるものを選定する必要があるが*2、本システムでは簡便化のために、類似している用例のうち最初に登録されたものを回答文、2番目以降に登録されたものを類似文とした。

3.2 回答文可視化機能

本節では、回答文可視化機能について述べる。本機能には、TOP ページからアクセスすることができる。本機能では、回答文間の関係をグラフ化することで可視化を行う。また、類似文については該当する回答文の周りに提示することとした。このことで、回答文間の関係性や回答文に付随する類似文の数を容易に把握できるようにしている。

回答文の可視化では、まず、質問文ごとに登録された回答文群を取得し、各回答文を形態素解析器で単語に分解する。次に、単語の生起回数を要素としたベクトルに変換し、各回答文間のコサイン類似度 $\cos(X, Y)$ を求める。その後、(1)式で距離 $d_{(X, Y)}$ に変換し、全回答文間の距離行列を求める。

$$d_{(X, Y)} = 1 - \cos(X, Y) \quad (1)$$

最後に、距離情報を保って低次元の座標値に変換すること

*2 回答文は他の言語に翻訳する必要があるため、翻訳を行うべき(翻訳を行いやすい)文を選定する必要がある。

Q.どのような症状がありますか？(内科での質問)

上記の質問文に対する入力状況を表しています。
赤色で表示されている文はあなたが登録した文です。

回答文	類似文
頭が痛いです	頭が痛い
喉が痛い	-----
鼻水かたまりません	-----
頭がズキズキします	-----
胸の真ん中あたりが痛いです	胸のあたりが痛いです
呼吸をするとヒューヒューいいます	-----
腹が痛い	-----
風邪気味	-----
胃がキリキリします	-----
胸が苦しい	-----

図 3 従来手法の画面例

ができる多次元尺度法を用いて 2 次元の座標値を得た。このことで、同じ単語を多く使用している回答文を近くに、異なる単語を多く使用している回答文を遠くに配置した可視化を行っている。なお、形態素解析器は ChaSen を、ベクトルの計算は R 言語をそれぞれ用いた。

可視化例を図 2 に示す。図 2 は、図 2(1)の質問文に対して登録された回答文と類似文を可視化したものである。前述の計算で得られた座標値をもとに散布図で可視化を行っている。また、図 2(2)のように、グラフ上の点をマウスオーバーすることで、どのような用例に対応しているかを表示している。グラフ上の点は 4 色で色分けを行っており、利用者自身が登録した回答文を青色、その他の利用者が登録した回答文を水色、利用者自身が登録した類似文を赤色、その他の利用者が登録した類似文をオレンジ色でそれぞれ提示している。このことで、利用者が応答用例対作成にどの程度貢献したかを把握しやすくした。あわせて、質問文に対して登録されている回答文数と類似文数、および、利用者自身が登録した回答文数と類似文数を提示している。

4. 実験

本章では、問診票回答収集システムの有用性確認実験について述べる。本実験では収集用例を可視化することで、システム利用者へ与える影響について検証する。

本稿では、前章で述べたものを提案手法、前章のシステムでグラフ提示箇所を表に変更したものを従来手法とする。従来手法は、多言語用例対訳共有システム [9] で利用されている提示方法である。提案手法の画面例を図 3 に示す。従来手法で利用者自身が作成した文は赤字で提示する形を取った。また、従来手法と提案手法のデータベースは共通とした。本実験は、2015 年 1 月 29 日から 2 月 5 日までの 8 日間行った。被験者は静岡大学工学部の学生 10 名であり、各手法に 5 名ずつ割り振っている。なお、収集する言語は日本語のみとした。問診票の質問文は各システム共通の 5 文用意した。以下にその質問文を示す。

質問文 1: どのような症状がありますか？(外科での

表 1 手法別の回答文および類似文の登録数

	回答文	類似文	合計
提案手法	178	17	195
従来手法	95	8	103
合計	273	25	298

・単位は文である。

質問)

質問文 2: どのような症状がありますか? (内科での質問)

質問文 3: 今までにかかった病気や現在治療中の病気は何ですか?

質問文 4: アレルギー (かゆみ・発疹など) を起こした食べ物は何ですか?

質問文 5: お酒は何を飲みますか?

実験協力者には、実験を開始する前に、システムの利用方法について説明した。問診票の質問文に対して回答するときは、実際に自分が医療機関にかかったときのことを想定して回答する旨を伝えている。また、メールと Twitter を利用し 1 日 3 回の通知をそれぞれ行っている。通知内容には、メールと Twitter からシステムに直接アクセスできるように URL を記載した。メールについては、8 時、12 時、20 時に通知を行い、Twitter については、10 時、16 時、23 時に通知を行った。さらに、実験の協力者には毎日最低 5 文の追加を依頼した。また、実験後にアンケートへの回答を依頼した。

5. 実験結果と考察

5.1 登録された回答文と類似文

表 1 に手法別の回答文および類似文の登録数を表す。表 1 より、提案手法は従来手法よりも回答文、類似文ともに多くの文が登録される傾向にあることがわかる。ただし、有意な差は見られなかった。

このうち、提案手法の実験協力者の一人は、実験期間中に 75 文の回答文の登録 (全体の 27%)、11 文の類似文の登録 (全体の 44%) をそれぞれ行っていた。また、合計 24 回のグラフの閲覧を行っており、回答文の登録とグラフの閲覧を交互に行っている様子が確認できた。このことから、グラフを提示することで回答文や類似文の登録を促進できる可能性があると考えられる。

ただし、他の提案手法の実験協力者は、日がたつにつれてグラフを閲覧しない傾向が見られた。実験で可視化された例を図 4 に示す。アンケートの自由記述から、「グラフ上の点が密集していたり、少ないスペースにはどんな言葉が入るのか気になった」という意見があった。しかし、図 4 に存在するような、グラフ上の点がまばらな箇所に入る言葉を推測することは難しいことなど、本手法で提示したグラフの閲覧方法が十分に伝わっていなかったため、日がたつにつれてグラフの閲覧意欲が低下していったと考えられ

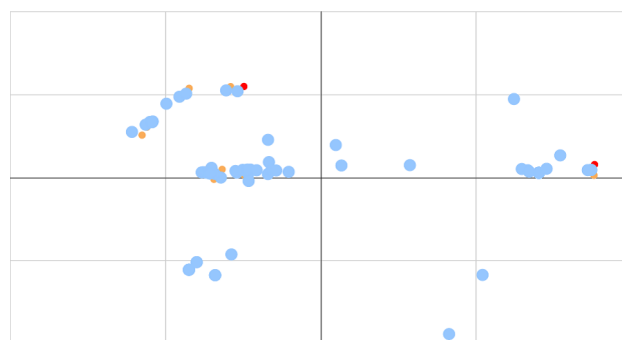


図 4 作成されたグラフの例 (質問文 2)

る。今後は、不足している用例がわかりやすい可視化手法に改良する必要があると考えられる。

また、全体で 21 文^{*3}の既存回答文と重複した用例が入力されていた。これは、全体の入力数の 6.6% (=21/(298+21)) にあたり、比較的大きな値となっている。実験協力者が既存用例を入力した場合は、システムに登録されない旨がシステムから示される。このことが実験協力者の回答文入力へのモチベーション低下につながっていた可能性がある。

また、実際に登録された回答文を確認したところ、「膝が曲がらない」「肘が曲がらない」「腰が曲がらない」など、「が曲がらない」という、一部の単語が共通した回答文が入力されている様子が観察された。特に提案手法の方がこのような回答文を登録する傾向が強かった。類似した回答文が近くに表示される可視化により、上記のような回答文の想起が行われたと考えられる。なお、このことは類似文の入力が少なかった要因の一つでもあると考えられる^{*4}。

5.2 他の実験協力者への意識

他の実験協力者の追加状況が気になるかどうかを聞いたアンケート結果を表 2 に示す。表 2 より、提案手法の方が従来手法よりも他の実験協力者の追加状況が気になるという傾向が得られた。実験後のアンケートから、提案手法を利用した実験協力者からは「全体で自分がどれぐらい登録しているのかを気にしていた」「他の人がどのぐらい協力しているのか気になった」などの意見が得られた。これは、可視化することにより自身の貢献が視覚的に確認しやすくなったことが影響していると考えられる。また、従来手法では他の人が回答文を追加しても表の行が増えるだけで、視覚的にあまり変化がないことに対して、提案手法の可視化では回答文を追加すると少しずつグラフの形が変化することも影響していると考えられる。

5.3 今後の課題

本システムでは、収集された回答文のグラフ化を行う際

^{*3} 表 1 の外数。

^{*4} 本稿の類似文は「回答文と入れ替えても意味が変わらない文」であるため、前述の「が曲がらない」など、一部の品詞が異なっている文は類似文ではない。

表 2 アンケート結果

		質問文	評価の分布					中央値	最頻値
			1	2	3	4	5		
(1)	従来手法	他のシステム利用者の追加状況が気になる	0	2	2	1	0	3	2,3
(2)	提案手法		1	0	1	2	1	4	4

・評価項目は、1:強く同意しない、2:同意しない、3:どちらともいえない、4:同意する、5:強く同意する、である。
 ・評価の分布の単位は人である。

のベクトル変換時に、品詞ごとに重み付けを行っていなかった。このため、名詞や動詞、形容詞などと助詞や助動詞が同じ重み付けとなっていた。例として、「頭/が/痛い/です」と「頭/に/痛み/が/あり/ます」の場合、頭に痛みがあることは共通している。また、「頭/が/痛い/です」と「腕/が/重い/です」の場合、意味が異なっている。しかしながら、本手法では「頭/が/痛い/です」に近い文章は「が」「です」が共通して使用されている後者となる。このため、品詞ごとに重みを変えて、適切な類似関係を得る必要があると考えられる。

また、「頭/が/痛い/です」と「頭痛/です」は意味的には同じであるが、一致する品詞がないため、本手法では類似文の比較の場面においても、可視化の場面においても検出することができない。このため、今後はシソーラスなどを用いた手法を検討する必要があると考えられる。

また、本実験では、実験協力者による回答文登録数が少ない傾向にあった。Twitter やメールによる依頼を行っていたが、今後、ほかのモチベーション維持手法を取り入れる必要があると考えられる。

6. おわりに

本稿では、医療分野の多言語問診票入力場面で使用される応答用例対の収集を目的とした、問診票回答収集システムについて述べた。本システムでは、収集された応答用例対を可視化することで、利用者が収集済み応答用例対の全体像の把握支援を目的としている。

本稿では、応答用例対の可視化手法とその実装を行った。また、実験結果から、可視化を行った提案システムは表による提示を行った従来システムよりも、多くの回答文が登録される傾向があることを示した。

今後は、本システムに挙げた課題点を解決した後、長期の実験を行う。

謝辞 本研究の一部は JSPS 科研費 (26730105) による。

参考文献

- [1] 法務省：平成 26 年における外国人入国者数及び日本人出国者数について (確定値)、法務省 (オンライン)、入手先 (http://www.moj.go.jp/nyuukokukanri/kouhou/nyuukokukanri04_00046.html) (参照 2015-05-15)。
- [2] Takano, Y. and Noda, A.: A temporary decline of thinking ability during foreign language processing, *Journal of Cross-Cultural Psychology*, Vol. 24, pp. 445-462 (1993).
- [3] Aiken, M., Hwang, C., Paolillo, J. and Lu, L.: A group

- decision support system for the Asian Pacific rim, *Journal of International Information Management*, Vol. 3, No. 2, pp. 1-13 (1994).
- [4] Kim, K. J. and Bonk, C. J.: Cross-Cultural Comparisons of Online Collaboration, *Journal of Computer Mediated Communication*, Vol. 8, No. 1 (2002).
- [5] 宮部真衣, 吉野 孝, 重野亜久里: 外国人患者のための用例対訳を用いた多言語医療受付支援システムの構築, 電子情報通信学会論文誌, Vol. J92-D, No. 6, pp. 708-718 (2009).
- [6] 杉田奈未穂, 丸田洋輔, 長谷川旭, 長谷川聡, 宮尾 克: ケータイ多言語対話システムとその応用, シンポジウム「モバイル'09」, pp. 63-66 (2009).
- [7] 福島 拓, 吉野 孝, 重野亜久里: 用例対訳と機械翻訳を併用した多言語問診票入力手法の提案と評価, 情報処理学会論文誌, Vol. 54, No. 1, pp. 256-265 (2013).
- [8] 尾崎 俊, 松延拓生, 吉野 孝, 重野亜久里: 携帯型多言語問診票対話支援システムの開発と評価, 電子情報通信学会技術研究報告, Vol. AI2010-47, pp. 19-24 (2011).
- [9] 福島 拓, 吉野 孝, 重野亜久里: 正確な情報共有のための多言語用例対訳共有システム, 情報処理学会論文誌. コンシューマ・デバイス&システム, Vol. 2, No. 3, pp. 23-33 (2012).
- [10] 福島 拓, 吉野 孝: 正確かつ自由度を高めた多言語問診対話支援を目的とした応答用例対構築モデル, 情報処理学会論文誌, Vol. 56, No. 1, pp. 219-227 (2015).
- [11] Matsuda, M. and Kitamura, Y.: Development of Machine Translation System for Japanese Children, *Proceedings of IWIC'09*, pp. 269-271 (2009).
- [12] 照井賢治, 菱山玲子: 多言語ケースメソッドシステムによる異文化分析, 電子情報通信学会技術研究報告, Vol. AI2012-35, pp. 79-84 (2013).
- [13] 林田尚子, 石田 亨: 翻訳エージェントによる自己主導型リベア支援の性能予測, 電子情報通信学会論文誌, Vol. J88-D1, No. 9, pp. 1459-1466 (2005).
- [14] 塚田 元, 渡辺太郎, 鈴木 潤, 永田昌明, 磯崎秀樹: 統計的機械翻訳, NTT 技術ジャーナル, Vol. 19, No. 6, pp. 23-25 (2007).
- [15] Bond, F., Nichols, E., Appling, D. S. and Paul, M.: Improving Statistical Machine Translation by Paraphrasing the Training Data, *Proceedings of IWSLT 2008*, pp. 150-157 (2008).
- [16] 狩野 翔, 福島 拓, 吉野 孝: 用例評価のモチベーション維持支援システム「用例の森」の開発と評価, 情報処理学会論文誌, Vol. 53, No. 1, pp. 138-148 (2012).
- [17] 吉野 孝, 西田祥子, 福島 拓: 携帯端末の通知機能を活用した用例対訳登録における利用者への貢献情報の提示による動機付けの検証, 情報処理学会論文誌, Vol. 56, No. 1, pp. 118-125 (2015).
- [18] 坂本 廣, 北村泰彦, 福島 拓, 吉野 孝: N-gram に基づく多言語用例検索手法の評価, 電子情報通信学会技術研究報告, 人工知能と知識処理, Vol. 110, No. 428, pp. 51-56 (2011).