

発話の時間推移を考慮した自由対談文のセグメント分割手法

西本一志[†] 間瀬健二[†] 中津良平[†]

組織の円滑な運営や新たな発想の獲得などの面においてインフォーマルコミュニケーションは重要な役割を果たすため、我々はインフォーマルコミュニケーションの支援手法の研究を進めている。本論文ではその一環として、不特定分野の自由対話を書き起こし、ある程度の編集を加えた対談文を、話題に応じてセグメント分割する手法について検討する。従来の話題構造抽出手法の多くは、目的指向の対話かあるいは新聞や雑誌の記事ないし論文などの一般的文章を処理対象としており、インフォーマルコミュニケーションのような自由対話を対象とするものは少ない。我々は本論文において、従来から利用されている手がかり句や結束性に加え、特に対話において重要な意味を持つ時間推移情報を利用した広域結束性を定義し、これを利用した、話題に応じた対談文のセグメント分割手法を提案する。提案した手法を実装した実験システムを用いての実験の結果、人手によるセグメント分割に対して、再現率 73.2%、適合率 66.7%の精度で一致する結果を得た。

Segmentation of Transcribed Free Conversation Taking Utterance Dynamics into Account

KAZUSHI NISHIMOTO,[†] KENJI MASE[†] and RYOHEI NAKATSU[†]

Informal communication plays an important role in many conversations such as for creating ideas and managing an organization smoothly. For this purpose, we have been studying methods to support informal communication. Currently, typical data for structuring conversations in the literature has come from goal-oriented conversations or general texts, i.e. scientific papers, articles of newspapers or magazines. In this paper, we propose a segmentation method for free conversation based on topics. We obtain free conversation in text form from transcribed round-table discussions. In addition to segmenting using standard cohesion and cue-phrases, our segmentation method also uses global-cohesion. Global-cohesion is a measure of regularity of keywords and topic transitions over an entire conversation. Among other factors, we use time transition information to obtain our global-cohesion measurement. In our experiments, we obtained a 73.2% recall factor and a 66.7% precision factor indicating that our segmentation scheme using a mixture of cohesion, cue-phrases and global-cohesion is useful for structuring conversations.

1. はじめに

絶えず多様化・複雑化しながら知的生産性を求められる現代社会においては、人と人のコミュニケーションにおける相互理解や創造的共同作業などの支援が非常に重要なタスクの1つとして認識されている^{1),2)}。筆者らは、特にインフォーマルコミュニケーションと呼ばれる対話の支援に注目している。インフォーマルコミュニケーションの意義としては、たとえば、休憩時間の雑談のような、自然発生した個人的関係を基礎とするインフォーマルコミュニケーションの存在によって、はじめて組織が効果的に運用されることが

指摘されている³⁾。さらに、そういった雑談における同僚や友人の発言の中に、思いがけない視点からの情報や、日頃気になっていた問題の解決の糸口を発見することは、我々自身しばしば経験することである。このようにインフォーマルコミュニケーションは、組織の円滑な運営や創造性の触発など、様々な点で重要な役割を果たす。

このため、従来から遠隔地間でのインフォーマルコミュニケーションを支援するシステムの研究などが行われている^{4),5)}。しかし、これらは基本的にインフォーマルコミュニケーションのための電子的な場を提供するにとどまり、インフォーマルコミュニケーションにおける創造性の積極的な触発や支援を意図したものではない。すなわち、Young⁶⁾の発想支援システムの分類手法を適用すると、従来のインフォーマルコミュニ

[†] 株式会社エイ・ティ・アール知能映像通信研究所
ATR Media Integration & Communications Research
Laboratories

ケーション支援システムの多くは、創造性の支援レベルとしてはたかだか秘書レベルあるいは枠組みレベルにとどまっており、生成レベルの機能を備えるものは存在しないといえる[★]。我々が目標としているのは、生成レベルの機能を備えたインフォーマルコミュニケーション支援の実現であり、その1つとして、人同士の創造的な対話に介在して、人の思考を触発し創造性を高めることを目的とし、適切なタイミングで適切な情報を自律的に検索提供するエージェントの研究を進めている^{7),8)}。

このようなエージェントの実現には、インフォーマルコミュニケーションを対象とする話題処理技術が必要であるため、筆者らは現在その研究を進めている⁹⁾。将来的には口頭で発話された音声対話を認識し処理することが望まれるのはいうまでもないが、現状の音声認識技術ではインフォーマルコミュニケーションのような自由対話の発話を高精度に認識処理することはできない。のみならず、自由対話を人手によって忠実に書き起こした文章の処理すら、多数の文法的不適格文や省略の存在ゆえにままならないのが実情である。

そこで本論文では、上記目標への第1段階として、現状の自然言語処理技術で機械処理（形態素解析）可能な程度に人手による修正を施した、3人以上の話者による対話を書き起こした文章を処理対象とし、これを話題に基づきセグメント分割する手法を検討する。テキストベースでのオンラインディスカッションのような対話を想定すれば、上記に準ずる程度の品質の文（章）による対話は実現可能である^{★★}。したがって、このような文章を処理対象とすることは実用的にも価値があるといえる。

以下、2章では、まず本研究で処理対象とする「対談文」の定義と特徴を示す。次いで従来の話題構造抽出手法について概観し、それを我々が処理対象としている対談文に適用する場合の問題点をあげ、インフォーマルコミュニケーションにおける対話を話題分割する手法の技術的要件を考察する。その結果、このような対談文のセグメント分割手法として、発話の表層情報に基づき、隣接する発話どうしが連結している程度を評価する局所結束性に加えて、発話の時間推移に基づ

き、複数の発話において話題が維持される程度を評価する広域結束性を用いた手法を提案する。3章では局所結束性を求めるための要因と、それらに基づく局所結束性の程度の評価手法について、4章では広域結束性を求めるための要因と、それらに基づく広域結束性の程度の評価手法について述べ、5章ではこれらの結束性を用いたセグメント分割手法を説明する。6章では、対談記事を用いた学習によるシステムパラメータの決定ならびにシステムの評価実験について述べ、本手法による不特定分野の自由対談文におけるセグメント分割の可能性を検討する。7章では、本論文で提案した手法についてまとめるとともに、今後本手法をよりラフな対話や口頭での対話に適用するために必要な検討項目について簡単に述べる。

2. 関連研究と予備的考察

2.1 対談文の特徴と処理の技術的要件

2.1.1 対談文の定義

本論文では対談文を処理対象とする。まず、対談文とその構造ならびに構成要素について、以下のとおり定義する。

発話 ある1人の話者が、ある1回の発言機会において発した発言内容の全体。したがって、一般的には1つの発話は複数の文で構成される文章となる。ただし、非常に短い発話の場合は、1つの発話が1つの文で構成される場合もありうる。なお、発話を複数の文に分割する際、どこからどこまでを1つの文とするかの判断は曖昧であり、一意に決定できない場合が多い。このため本論文では、「文」という単位への分割の判断は、すべて書き起こしを行う対談文編集者の判断に委ねるものとする^{★★★}。

対談 複数の話者が互いに協調して順次なす、内容的に関連がある発話の時系列的連鎖。

対談文 対話を書き起こした文章に対し、さらに形態素解析や構文解析などの機械処理が可能な程度に人手による修正を施した文章のこと。個々の発話にはその発話の話者を識別するための識別子を付与する。

2.1.2 処理対象とする対談文の特徴

対談文には、通常の新聞記事や論文などの文章と比べて、以下の2点の特徴がある。

- 複数の話者が交替になす発話の連続によって構成

★ なお、支援の度合や洗練度は、秘書レベルが最も低く、生成レベルが最も高いのだが、それは支援の効果の高さと必ずしも比例するわけではない⁹⁾。それぞれのレベルは、それぞれに特長ある支援を行うので、状況に応じて適切なレベルの支援を利用することが望ましい。

★★ たとえば、文献¹⁰⁾で示した、オンラインディスカッションシステムを用いての2つの対談例における発言は、すべて対談文に準ずる品質の文で構成されていた。

★★★ 機械処理により発話を文に分割する処理手法の研究は今後の課題とし、本論文では取り扱わない。

される。

- 最初の発話から最後の発話に向けて、一方向に流れる時間が重要な役割を果たす。この結果、対談文は次の2つの性質を持つ。

変容性 時間の経過とともに、過去の発話に含まれる情報の一部は徐々に失われる一方、新たな情報が順次追加されていく。この結果、対談内容の一貫性 (coherence) は対談全体を通じては必ずしも維持されない。

先行不能性 ある時点までの対談に対する各種判断・処理には、その時点以後の未来の発話に現れる情報は使用できない。たとえば、発話をまたがる後方照応はありえないこと[☆]などはその一例。

また、本論文では特にインフォーマルコミュニケーションにおける対談を処理対象とする。このような対談には、話題にかかわる次の2つの特徴がある³⁾。

- 明確な到達目標がなく、話題は不特定分野に広がる。
- 特定の構造を持たない自由対談である。

2.1.3 処理の制約と技術的要件

本論文で提案する手法では、以下の制約と要件を満足することを目標とする。

話者数 本研究では比較的多数の話者が気ままに話し合う状況における対談の処理を目指す。このため、対談に参加する話者の人数は3人以上の場合を想定する。

セグメント境界の位置 話題セグメントの境界は、発話と発話の間に存在するものとする。現実には1つの発話の途中において話題転換が起こったように見える場合もありうるが、このような場合でも1つの発話の途中にセグメント境界を置くことはなく、その発話全体として話題転換があったかどうかを判断し、転換があったと判断できる場合は、その発話と前発話との間に境界を置く。なお、セグメント境界という言葉と話題転換点 (箇所) という言葉はほぼ同じ意味であるが、厳密には、話題転換点とは新たな話題が開始された発話のことであり、セグメント境界とは、話題転換点となった発話とその直前の発話との間のことを指す。

明瞭な話題転換点の検出 話題転換点には、多くの人々が一致してその存在を認識できる明瞭な話題転換点と、各人の判断がバラバラとなり、1人かせい

ぜい2人程度だけが認識する曖昧な話題転換点があると思われる^{☆☆}。曖昧な話題転換点の判断は主観的な根拠によると思われるため、機械的に検出することは難しいと予想される。したがって、本研究では明瞭な話題転換点のみを検出対象とする。

判断処理タイミングと判断材料 本研究では、実際に実時間的になされる対談に1章で述べたようなエージェントが介入し、セグメントの内容に依存した各種対談支援処理を、対談の進行に対し遅延なく行えるようにすることを目標としている。そこで話題転換点の検出には、現在話題転換の有無を判断しようとしている発話 (以下これを現発話と呼ぶ) と、それ以前の発話から得られる情報を判断材料とし、現発話より未来になされる発話を使用しない。

システムの持つ知識 インフォーマルな自由対談を処理可能とするため、一般的な国語辞書的知識以外の、特定のドメイン知識や話題構造知識を用いない。

2.2 従来研究とその問題点

従来から話題や話題構造の認識処理の研究は多数なされている。これらの研究が処理対象とする文章は、次の2種類に大別できる。1つは、本来時間要因が重要な意味を持つ文章で、2人以上の話者による対話がこのあたると。もう1つは時間要因があまり意味を持たない文章で、たとえば新聞や雑誌の一般的記事などの、発話を書き起こしたものではない文章や、あるいは1人の話者によるモノログなどがこれにあたる。モノログは講演の記録などの、発話で構成された文章ではあるが、このような発話は一般に事前に内容がかなり詳細に設計されているのが普通である。したがって、上記の変容性や先行不能性は必ずしもあてはまらないため、時間要因があまり意味を持たない文章と見なされる。

2.2.1 時間要因が重要な意味を持つ文章の処理

時間要因が重要な意味を持つタイプの文章を処理する研究のほとんどは、たとえば会議の受付での参加申し込み時の対話のような、目的指向型の対話を対象としている。このような対話ではあらかじめどのような発話がやりとりされるかがかなりの程度想定可能であるため、plan-goal モデルに基づく知識や、語彙的な話題階層知識をシステムに持たせて、次発話をあらかじめ予測することによる話題処理が成されることが多い (たとえば文献11))。またこのような対話は一般に

☆ 同一発話内のきわめて近傍に対する後方照応はこの限りではない。

☆☆ これについては6.3.1項で検証する。

質問・応答によって構成されるため、小文末表現を利用した発話間の結束性を用いて発話対を決定することによる対話の構造化¹²⁾なども行われる。

これらの手法は、目的指向を前提とすることで対話全体での一貫性を仮定して変容性の問題を回避し、また特定の対話内容分野を想定した知識を持ち、種々の推論を行うことによって先行不能性の問題を回避している。つまり、実際には時間推移の影響を無視できる状況を想定しているといえる。したがって、これらの手法の多くは、時間推移の影響を回避できない自由対話には適用できない。発話間の結束性の利用は、目的指向型以外の対話についても有効と考えられるが、自由対話においては発話対は必ずしも成立しないという指摘もあり¹³⁾、目的指向型の対話処理における結束性の求め方とは異なる結束性の獲得手法を利用することも必要であろう。

一方、自由対話を扱う手法としては、cue phraseと呼ばれる特定の発話表現や意図構造、注意状況を利用した話題構造抽出手法¹⁴⁾、分野同定 FIFO や対話分野に応じたドメイン知識の使用による話題同定手法¹⁵⁾、因果関係をもとにした発話対抽出手法¹⁶⁾などが試みられている。

しかし、表層的情報だけから意図構造や注意状況を獲得することは困難である。また、分野同定 FIFO を用いる手法は、ある時点での話題転換を検出するために、その時点より未来になされる数発話からの情報を必要とするので、先行不能性を侵害する。さらに、インフォーマルな自由対談の処理にドメイン知識を用いることは現実的でない。ただし、cue phrase は非常に明解に話題構造を提示する情報であり、しかも一般に変容性と先行不能性を侵さないため、その利用は様々な対話（さらには一般的な文章）において有効であると考えられる。

2.2.2 時間要因があまり意味を持たない文章の処理

時間要因があまり意味を持たないタイプの文章を処理する研究においても、やはり結束性を用いる手法と、新たな話題が導入される際の特定の言語現象を利用する手法が一般的である。

結束性を用いる手法としては、たとえば隣接する2つの文塊（1つの文塊は複数の文で構成される）どうしの類似性を、それらの文塊が共有するキーワードとそれら個々の重みに基づいて求め、その結果文塊どうしが内容的に連続かどうかを判定することにより、文章全体をセグメント分割する手法が提案されている¹⁷⁾。この類似性はある種の結束性と見なすことができ、しかも基本的にキーワードの共有関係だけで求め得るも

のであるので、一般的な文章のみならず、我々が対象とする対談文にも適用可能である。ただし、文献17)で使用されている *tf·idf* と呼ばれるキーワードの重みづけ手法¹⁸⁾は時間要因を考慮しておらず、処理対象の文章全体から得られる統計量に基づく重みづけを行っているため、対談文における先行不能性に対応していない。また、結束性とは文中に含まれる要素間の関係によって生じるものであるが、結束性が認められた文どうしが必ずしも内容的に連続しているとは限らず、また逆に結束性が認められない文どうしが内容的に連続している可能性もあることが指摘されている¹⁹⁾。この問題に対処するには一貫性を考慮する必要があるのだが、一貫性を取り扱うには表層的情報だけでは通常不十分で、深い意味理解に立ち入らねばならないという問題がある。

話題導入時の言語現象を利用する手法では、cue phrase や話題マーカを基本として話題分割を行い、さらに文長などを考慮して大局話題と局所話題の判定を行っている²⁰⁾。この手法は我々の対象とする対談文にも適用可能である。ただし、現実には話題マーカは非常に多くの発話に出現するため、これを単純に使用すると話題転換が非常に多数検出される傾向が見られることが、我々の予備実験で分かった。これは実際の対談の中で人間が認識する話題の粒度よりもはるかに細かい粒度での話題を検出する結果をもたらす。

2.3 話題構造検出の方針

以上の考察、ならびに従来研究の概観から、我々は以下のような手段を採用する。

- (1) 対談データから採取する情報は表層的な情報のみ、すなわち形態素情報、時間情報および話者情報とする。
- (2) 特定分野に依存するドメイン知識や話題構造知識を使用しない。
- (3) 発話間の結束性を利用してセグメント境界を検出する。結束性の判定には、以下の情報を用いる。
 - (a) cue phrase などの話題転換/継続にかかわる特定の言語現象
 - (b) キーワードの共有、あるいは質問応答・照応などの、ある発話とその直前発話とを関連づける表層的特徴
 - (c) 対談の進行による時間推移情報に基づく、一貫性に準じる情報

★ 詳細は4.1.2 項参照

☆☆ cue phrase については頻度はそう高くなく、対談の中で人間が認識する話題の粒度程度か、それよりやや大きい程度である。

なお、従来から一般に使用される発話間の結束性は、現発話そのもの、あるいは現発話とその直前の発話とから得られる局所的な情報に基づいているので、これを本論文では局所結束性と呼ぶ。上述した結束性の(3)aと(3)bの判定方法は、このような局所情報のみに基づいている。したがってこれらによって得られる結束性は局所結束性である。

一貫性は、「意味内容」という深層情報によって文どうし(あるいは発話どうし)を結び付けるものである。したがってこれは隣接する2文(2発話)のみではなく、通常いくつかの文(発話)にまたがる広い範囲に対する結束力を産み出す。本研究では、すでに述べた理由により表層情報のみを取り扱うため、意味内容に深く立ち入ることはできない。しかし、話題候補語が表層的な情報から獲得可能であることは、すでに2.2節で示した従来研究の多くで示されており、それによって各発話の意味内容を浅いレベルで扱うことが可能となる。以上から、厳密な意味での一貫性を取り扱うことはできないが、話題あるいは個々の話題候補語の、対談の進行にともなう変容の可能性を取り扱うことによって、一貫性に準じる情報を得ることが可能であると考えられる。

そこで、本研究では、現発話以前のすべての発話から得られる情報の時間推移に基づく結束性を定義する(上述した(3)cの判定方法によって得られる結束性にあたる)。これを広域結束性と呼ぶ。これによって、たとえば現発話とその直前の発話とが局所結束性を判定するための要因をいっさい持たない場合であっても、広域結束性によって現発話が直前発話とどの程度連結されているかを判定できるようになる。

3. 局所結束性

局所結束性は、現発話のみ、あるいは現発話とその直前発話に含まれるいくつかの表層的要因から得られる、現発話とその直前発話との連結の程度を示すものである。本章では、局所結束性を求めるための表層的要因と、それらの要因に基づく結束性の強さの評価方法について述べる。

3.1 局所結束性の要因

局所結束性を求める手がかりとして、以下に示す表層的要因を各発話から抽出し、これらの要因を用いて局所結束性の強さを評価する。これらの表層的要因は、すべて現発話とその直前発話のみから得られるものである。本研究では、局所結束性の要因を断絶要因と連結要因の2つに分類する。断絶要因とは、隣接する発話の間の連続性を積極的に断ち、多くの場合新しい話

表1 局所結束性を求める手がかりとなる表層的要因
Table 1 Surface factors to evaluate local-cohesion.

断絶要因	連結要因
Cue phrase	共有語
相手特定表現	反対語
	指示語
	接続詞
	前疑問文
	応答文
	コメント文

表2 使用した cue phrase
Table 2 Cue phrases.

ところで、じゃあ、話が逸れますが、もうひとつ、 それにしても、～といえば、話は変わりますが、 話を戻しますが、話は違いますが。

題を導入する機能を持つと見なされる要因のことであり、連結要因はこの逆で、隣接する発話を選び付ける機能を持つと見なされる要因のことであり、表1に、局所結束性を求める手がかりとなる表層的要因の一覧を、断絶要因、連結要因のそれぞれについて示す。

3.1.1 断絶要因

以下の2つを断絶要因と見なす。

Cue phrase 本研究でいう Cue phrase は、一般に新しい話題を導入する際に使用される特定の表現のことを指す。したがって、これらは通常、話題転換点を明示的に示す。プロトタイプシステムで使用している Cue phrase を表2に示す。

相手特定表現 対談に参加している人物Aに対し、「Aさん、～についてどう思いますか」のように、その人の名前を呼びかけて尋ねる表現で、3人以上の会話特有の表現である。予備調査において、明示的に対談相手を指定するのは、その相手がよく知っていそうな話題を持ち出す場合であることが多かった。したがって、このような表現が使用される際に話題の転換が起る可能性が高いと考えられるため、相手特定表現も断絶要因として扱うことにした。

3.1.2 連結要因

以下の要因は、すべて直前発話に対するなんらかの参照を行うものである可能性が高いため、連結要因と見なす。

共有語 前発話に含まれる名詞の同一語が現発話にあるかどうか。これは、文献17)における文間で共有されるキーワードにあたる。なお、共有語については、ごく一部同義語も共有語として扱っている。

反対語 前発話に出現したある名詞の反対語が現発話にあるかどうか。

指示語 現発話の第1文に指示代名詞があるかどうか。

接続詞 現発話の第1文に接続詞があるかどうか。

前疑問文 現発話の前の発話が疑問文かどうか。

応答文 現発話が、「なるほど」「おっしゃるとおりです」などだけのごく短いあいづち的な発話かどうか。応答文の検出は、あらかじめ「なるほど」「おっしゃるとおり」などの相づち表現を集めた相づち表現辞書を準備しておき、ごく短い発話においてこの辞書に含まれる表現が現れる発話を抽出することによって行う。

コメント文 前発話に対する確認的内容を持つ発話かどうか。第1文が「～ね。」で終わるものをコメント文と見なす。これは1種の小文末表現による結束性と見なすことができる。

なお、同義語・反対語は手動で作成した辞書を使用して検出する。この辞書には現在、実験に使用した全対談文データから収集した57の項目が登録されている。また、現発話の直前の発話が応答文である場合には、共有語ならびに反対語は現発話と応答文である前発話とを比べて求めるのではなく、現発話と応答文の前の発話（つまり現発話の前々発話）とを比べて求めるものとする。また、実際の対談では直前発話の一部をそのまま繰り返す現象がしばしば見られ、これは有力な連結要因となると思われる。しかし、どの部分がどの程度の長さ（割合）で一致する場合にそれを連結要因として採用するか判断が難しいため、現在の手法では繰り返し表現を局所結束性要因として扱っていない。

3.2 局所結束性の程度の評価

以上の局所結束性の要因を用いて、各発話を以下の3つのカテゴリに分類する。表3に、各カテゴリと各要因との関係を示す。

断絶群 いずれかの断絶要因を1つ以上持つ発話（連結要因の有無については不問）。局所結束性は強い断絶を示す。したがって、断絶群に含まれる発話において話題が転換する可能性は高い。

不安定群 局所結束性の要因をまったく含まない発話。局所結束性は断絶性・連結性とも弱い。不安定群に含まれる発話において話題が転換する可能性は中程度である。

表3 局所結束性要因に基づく発話のカテゴリ分類法
Table 3 Categorization of utterances based on the local-cohesion factors.

カテゴリ	断絶要因	連結要因
断絶群	あり	不問
不安定群	なし	なし
連結群	なし	あり

連結群 断絶要因は含まず、1つ以上の連結の要因を持つ発話。局所結束性は強い連結性を示す。したがって、連結群に含まれる発話において話題が転換する可能性は低い。

4. 広域結束性

本章では広域結束性を求めるための要因として**話題維持指数**と**名詞の話題維持寄与指数**を定義し、これらを用いて広域結束性の強さを評価する方法について述べる。これらの指数は、すべて現発話以前の発話における表層情報と、対談の時間推移情報から得られるため、対談の変容性と先行不能性を侵さない。しかもこれらの獲得には、特定の分野知識や対談構造、planなどを必要としない。なお、これらの指数はいずれも本質的に連結の程度を示すものであり、断絶の程度を示すものではない。すなわち、ある要因が最も話題維持に寄与しない状態にあるとしても、それはたかだかそこまでの話題が維持されてもされなくてもよいことを示すにとどまり、話題が転換することを積極的に示すものではない。

4.1 広域結束性の要因

4.1.1 話題維持指数

話題維持指数は、話題転換が発生した直後は話題の転換が発生しにくい、話題転換後ある程度対談が進行すると次の話題転換が発生してもよい（しかし、必ずしも転換しなければならないわけでもない）状況となるというヒューリスティクスに基づき、その傾向を定量化するものである。

第n発話の話題維持指数 C_n を次式で定義する。

$$\begin{aligned} \text{if } O_{n-1} = 0 \text{ then } C_n = 1, t = 0 & \quad (1) \\ \text{if } O_{n-1} = 1 \text{ then } C_n = \frac{1}{1 + e^{rc(t-t_c)}} & \end{aligned}$$

ここに、 t は前の話題転換点からの経過時間、 rc と t_c は定数である。 O_{n-1} は第 $(n-1)$ 発話においてシステムが判断した話題の転換の有無（0：転換，1：話題維持）である。

図1に、システムが判断した話題転換の有無 O_n

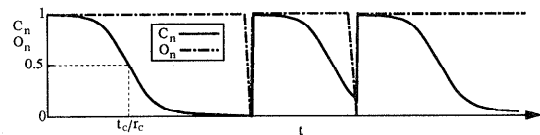


図1 プロトタイプシステムによる話題転換点検出例とその場合の話題維持指数の変化の様子

Fig. 1 Example of detection of topic transition by the prototype system and temporal change of topic maintain index under the example.

の経過例と、その場合の話題維持指数 C_n の変化の様子を示す。図から分かるとおり、話題維持指数 C_n は話題転換検出直後に値 1 にリセットされ、その後 $t = t_C/r_C$ 経過時に $C_n = 0.5$ となった後、次第に 0 に漸近する。ここに、話題維持指数 C_n は、1 に近いほど話題の維持傾向が強く、0 に近いほど話題の維持傾向が弱いことを示す。したがって、この話題維持指数の変化は、先に述べたヒュースティクスを記述するものである。

4.1.2 名詞の話題維持寄与指数

名詞の話題維持寄与指数は、対談のある時点において、その直前までになされてきた話題の維持に対して、ある名詞が寄与する程度を示すものである。これは、一般的にいうキーワード重みと類似している。ただし、一般的の意味でのキーワード重みとは、あるキーワードの、そのキーワードを含む発話の内容にとっての重要度のことであるが、名詞の話題維持寄与指数は、以下に説明するように、それとは無関係ではないがやや異なった意味を持つものである。

名詞 w の第 n 発話における名詞の話題維持寄与指数 $N_{w,n}$ を次式で定義する。

$$N_{w,n} = \frac{1}{1 + e^{\sqrt{\left(\frac{t_w}{t_N}\right)^2 + \left(\frac{f_{w,n}}{f_N}\right)^2} - r_N}} \quad (2)$$

ここに、 $f_{w,n}$ は名詞 w の第 $(n-1)$ 発話までの出現回数、 t_w は名詞 w が前回使用されてから今回使用されるまでの経過時間である。ただし、名詞 w が初発話である場合は、 t_w の値は非常に大きな値とする。また、 t_N 、 f_N 、 r_N は定数である。

図 2 に名詞 w の話題維持寄与指数 $N_{w,n}$ の変化の様子を示す。図から分かるとおり、話題維持寄与指数

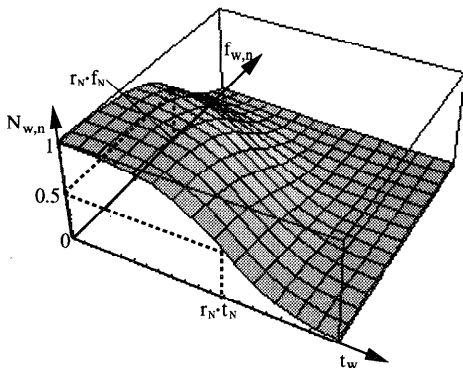


図 2 名詞の話題維持寄与指数の変化の様子

Fig. 2 Index of contribution for each noun to maintain the current topic.

$N_{w,n}$ は、語 w の累積出現頻度 $f_{w,n}$ と出現間隔 t_w がともに小さい場合にのみ 1 に近い値をとり、 $f_{w,n} = 0$ の場合 $t_w = r_N \cdot t_N$ で $N_{w,n} = 0.5$ 、あるいは $t_w = 0$ の場合 $f_{w,n} = r_N \cdot f_N$ で $N_{w,n} = 0.5$ となり、その後 $f_{w,n}$ と t_w のいずれか、あるいは両方の値が大きくなるに従って 0 に漸近する。ここに、話題維持寄与指数 $N_{w,n}$ は、値が 1 に近いほど話題の維持への寄与度が高いことを示し、0 に近いほどその名詞の話題維持への寄与度が低いことを示している。したがって、このような話題維持寄与指数 $N_{w,n}$ の変化は、累積出現頻度 $f_{w,n}$ が低くしかも出現間隔 t_w が短い語 w は、その直前までの話題の維持に対して強く寄与するものとシステムが評価することを示している。

このように評価する根拠について以下説明する。語の出現頻度によって重みづけを行っている点で、式 (2) の寄与指数算出方法は $tf \cdot idf$ (term frequency · inverse document frequency)¹⁸⁾ と同じ考え方に基づいている。 $tf \cdot idf$ では、テキスト集合 \mathbf{D} に含まれるあるテキスト d_i における語 w_j の重み W_{d_i, w_j} を、以下によって算出する。

$$W_{d_i, w_j} = tf_{d_i, w_j} \cdot idf_{D, w_j}$$

ここに、 tf_{d_i, w_j} は、テキスト d_i における語 w_j の出現頻度である。また、 idf_{D, w_j} は、以下の式で与えられる。

$$idf_{D, w_j} = \log \left(\frac{N_D}{df_{D, w_j}} \right)$$

ここに、 N_D はテキスト集合 \mathbf{D} の大きさであり、 df_{D, w_j} は語 w_j を含むテキストの数である。すなわち、 $tf \cdot idf$ では、

- ある 1 つのテキストの中で高頻度に出現する語は、そのテキストにとって重要である。
- テキスト集合全体を通して出現頻度が高い語は、
 - そのテキスト集合全体を通しての大局的なテーマにかかわる語
 - どんな種類のテキストでも使用される一般的な語

のいずれかと考えられるため、特定テキストにおける重要性（そのテキストを他のテキストと差別化する機能）は低い。

という考え方にに基づき、重みづけを行っている。

一方、式 (2) では、以下の点で $tf \cdot idf$ と異なっている。第 1 に、 $tf \cdot idf$ ではテキスト集合全体における語の出現頻度を使用しているのに対し、式 (2) では現発話以前の発話のみにおける頻度を使用している。これは対談の先行不能性を考慮しているためである。

第 2 に、式 (2) では、ある語が前回使用されてから

の出現間隔を考慮している。これは対談の変容性を考慮しているものである。文献17)などでは、時間的推移性のない一般的なテキスト群をセグメント分割の対象としている。しかし、対談文の中の各発話のような時間推移のある状況では、各語の出現間隔が話題の維持と転換に対して重要な寄与を果たす。たとえば、対談全体を通じて使用頻度の低い語は、それを含む発話にとってはやはり重要な語と見なせるが、

- 第 $(n-1)$ 発話と第 n 発話がその語を共有する場合
- 第 $(n-k)$ 発話 ($k \gg 1$) と第 n 発話がその語を共有するが、第 $(n-k+1)$ 発話から第 $(n-1)$ 発話までの発話はその語を含まない場合

を比べると、第 $(n-1)$ 発話における話題が第 n 発話でも維持されるかどうかの判断に対する、その語の寄与のあり方は、明らかに異なる。前者の場合、その語は、話題の維持に強力に寄与するであろう。しかし、後者の場合は話題の維持に対する寄与は低い。むしろ、その語は新たな話題を導入するもの、あるいはかつて話題となっていたがいったん消えた話題を再導入するものとなる可能性もある。対談全体を通じて高頻度な語であっても、ある程度長期間使用されなかった後に再出現した場合は、やはりかつて話題となっていたがいったん消えた話題を再導入するものとなる可能性が考えられる。以上が、式(2)と $tf \cdot idf$ との相違点である。

ここで、長期間使用されなかった語や初出語(特に初出語)は、新しい話題をもたらすものとなりうるので断絶要因と見なすべきではないかという議論がありうる。しかし、後で述べる実験用対談データなどについて調査した結果、初出語は対談中にかなり多数出現することが分かった。したがって初出語をすべて積極的課題転換の指標として扱おうと、話題転換でない箇所を話題転換点として多数誤認識する傾向が見られた。そこで本研究では、初出語についてもその出現は話題転換を促すものではなく、それらの語個々による話題の維持傾向は弱いと判断するにとどめることにした。

なお、 tf_{d_i, w_j} にあたる項は、式(2)には含まれない。これについては、4.2節で示すように、現発話に含まれる複数キーワードのうちから代表値を決定する際に考慮する。

4.2 広域結束性の程度の評価

以上定義した指数を用いて、広域結束性の強さを以下に示す3通りの方法で評価する。

4.2.1 話題転換後の経過時間に依存する話題維持傾向の評価

話題維持指数 C_n をそのまま適用し、この評価結果

とする。すなわち、話題転換検出直後に再び話題が転換する可能性は低く、そこまでの話題が維持される傾向が強いと評価する。

4.2.2 直前発話との共有語による話題維持力の評価

局所結束性で連結性の指標と見なしている、連続する発話の共有語について、さらに各語の話題維持寄与指数を用いてその結束性の程度を詳細に評価する。第 n 発話が第 $(n-1)$ 発話との共有語を複数個含む場合、第 n 発話に含まれるすべての共有語の話題維持寄与指数 $N_{w,n}$ の最大値を、第 n 発話における共有語による話題維持寄与指数の代表値 S_n と定義する。代表値を最大値としているのは、話題維持寄与指数の大きい語、すなわち低頻度だが連続的に出現している語はその区間での話題語である可能性が非常に高く、したがって、その語が現発話における共有語による話題維持傾向を支配するだろうとの予想に基づく。なお、共有語を含まない発話の S_n は0とする。

4.2.3 現発話を含む名詞による話題維持力の評価

第 n 発話に含まれるすべての名詞の話題維持寄与指数によって、第 $(n-1)$ 発話までになされている話題をその発話が維持する傾向を評価する。話題維持寄与指数の値域を幅0.1の10階級に分割し、現発話に出現した全名詞の話題維持寄与度が最も多く含まれる階級(mode)を上位2つ求め、その2階級の加重平均をその第 n 発話の代表値 N_n とする。ここで代表値を全評価値の平均でなくモードで決定するのは、

- (1) 度数の高い階級がその発話の話題維持傾向を特徴づけるであろう、
- (2) 評価値の分布は正規分布ではなく一般に歪んだ分布となるであろう、

という予想から、モードの方が発話の特徴をより良く表現できるであろうとの考えに基づく。この処理によって、ある単一発話内での各語の出現頻度が考慮される。この結果、 N_n は第 n 発話に出現するすべての名詞による全体的な話題維持傾向を示す値となる。

5. セグメント分割手法

本章では、局所結束性と広域結束性を組み合わせ、最終的に対談文をセグメント分割する方法について説明する。基本的な方針は、話題維持力がきわめて弱い状態にある発話を検出することにより、その発話で話題転換が起こると見なすものである。

対談文のセグメント分割手法のアルゴリズムをまとめたフローチャートを、図3に示す。まず、入力された発話文を形態素解析し、名詞や各種局所結束性要因を抽出し、さらにそれまでの対談における名詞の出

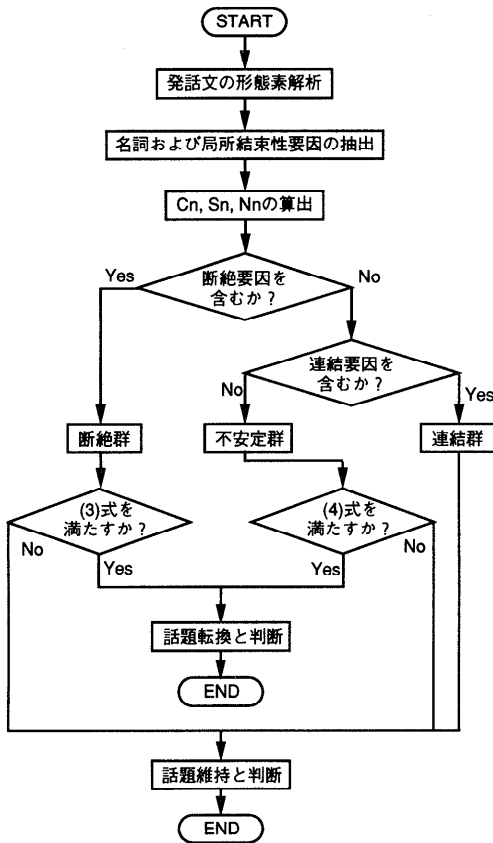


図3 セグメント分割の処理の流れ

Fig. 3 Flow chart of segmentation process.

現頻度や時間経過などの履歴情報を用いて、 C_n 、 S_n 、 N_n を求める。次いで局所結束性を用いて隣接発話間の関係を求める。その結果、連結群に含まれる発話については強い結束力があるものと見なし、その発話では話題転換は起こらないものとする。これは基本的に従来の結束性に基づく話題構造化手法における仮定と同様であるし、また実際に我々が行った被験者実験においても、連結群に含まれる発話で話題転換が起こる確率は非常に低いものであった。一方、不安定群と断絶群に含まれる発話については、さらに広域結束性を用いて以下のようにしてセグメント分割の有無を判定する。

5.1 断絶群に含まれる発話についての判定

従来の cue phrase を基本とする話題構造化手法では、断絶群に含まれる発話では話題の転換が起こるものとして処理されていた。しかし、実際に被験者実験を行うと、cue phrase を含むにもかかわらず話題の転換はなかったと判断されるような場合が若干存在する(表8参照)。これはその発話を含む連結要因の影響が非常に強く、それが cue phrase による断絶力を上回

るような場合に起こるものと考えられる。そこで、以下のような判定を行うことによって、断絶群に含まれる発話から広域結束性が極端に強くはない発話を抽出し、それらの発話で話題の転換が起こるものと見なすことにする。すなわち、第 n 発話における広域結束性の各要因が以下の条件を満たす発話を抽出し、該当する発話で話題転換が起こるものと判断する。

$$(C_n < \alpha_C \wedge N_n < \alpha_N) \quad (3)$$

$$\vee (N_n < \alpha_N \wedge S_n < \alpha_S)$$

$$\vee (S_n < \alpha_S \wedge C_n < \alpha_C)$$

ただし、 α_C 、 α_N 、 α_S は 0 以上 1 以下の定数である。

5.2 不安定群に含まれる発話についての判定

従来、不安定群に含まれる文(発話)についてはどの手法においても特に積極的には考慮されておらず、つねに「その他の文(発話)」として扱われていた。しかし、実際には cue phrase を含まない発話においても人間は話題転換を認識する場合があります、しかもそれはこの不安定群に属する発話においてである確率が比較的高いことが我々の行った実験で分かっている(表5, 表7参照)。そこで、不安定群に含まれる発話について、第 n 発話における広域結束性の各要因が以下の条件を満たすような、特に話題維持傾向の弱い発話を抽出しそれらの発話で話題の転換が起こるものと判断する。

$$(C_n < \beta_C \vee N_n < \beta_N) \quad (4)$$

ただし、 β_C 、 β_N は 0 以上 1 以下の定数である。

6. 実験

以上述べた手法に従ってプロトタイプシステムを実装した。本章では、このプロトタイプシステムを用いて行った学習用データによる実験と、それに基づく各種パラメータの決定、および評価用データに基づく評価実験結果を示す。

6.1 学習用データによるパラメータの決定

まず学習用データを用いた実験により、式(1)~(4)の各定数の決定を行った。学習用対談データとして、雑誌「婦人公論」に掲載された対談記事4編を用いた。それぞれの記事について4~6人の被験者に提供し読んでもらった。被験者は、筆者らの研究所内の研究員と職員、ならびにその家族であり、のべ13人である。このうちから、実験ごとに4~6人の被験者を無作為に割り当てた。

実験では、対談文を印刷したものを被験者に渡し、これを気軽に流し読みして、読んでいて「話題が変わった」と思った箇所にチェックするよう指示した。チェックは、必ずそう思った時点で読んでいる箇所につける

表4 学習用データに対する被験者による話題転換点抽出結果
Table 4 Results of topic transition point extraction by subjects for training data.

対談文	発話数	転換点数	一話題あたりの 平均発話数	標準 偏差	最短話題 構成発話数	最長話題 構成発話数
1	136	11	11.3	8.2	4	34
2	167	14	11.1	7.7	3	33
3	127	6	18.1	5.1	10	26
4	121	9	12.1	8.2	5	34
合計	551	40	12.5	7.9	3	34
一発話あたりの平均文字数						
被験者が対談文を読む平均的速さ			10.5 文字/sec			

ものとし、決して過去に遡ってチェックをつけないように要求した。これはたとえば、ある時点で読んでいる箇所から得た情報に基づき、その箇所よりも以前のすでに通過した箇所では話題が変わっていたと感じた場合でも、そこに遡ってチェックをするのではなく、そう感じた時点で読んでいる箇所にチェックをつけるようにするということである。この実験結果から、60%以上の被験者が話題転換と判断した発話を話題転換があった発話として抽出した。以下このような発話を転換点、それ以外の発話を非転換点と呼ぶ。

なお、「気軽に流し読み」することを指示したのは、セグメント分割を意識し過ぎて記事を読むことによって対談の変容性が無効化することを極力抑えたかったことによる。また、遡ってのチェックを禁じたのは、先行不能性の無効化を避けるためである。被験者実験に先行不能性を求める理由は、2.1.3 項で示した判断処理タイミングと判断材料の要請、すなわち遅延なき話題転換点検出手法を実現するためである。被験者に先行不能性を課さずに自由な先読みを許した場合、ある箇所における判断の結果は、先行不能性を課した場合と異なった結果となる可能性がある。あらかじめ完了された対談をセグメント分割する場合には、先行不能性を課さない方がより適切な分割結果を与えることはいうまでもない。しかし、進行中の対談をセグメント分割する場合には、このような先読み情報を利用する判断手法は適切でないし、そもそも適用できない。現実人間は進行中の対談において、ある時点における話題転換の有無を、その時点より未来になされる発話からの情報は用いず、その時点までになされた発話のみから得られる情報に基づき判断し、その結果に基づき次の自分の行為を決定している。このような、人による進行中の対談における判断と、できるだけ同等の判断を実現するために、被験者実験に先行不能性の制約を課した。

被験者による転換点抽出結果を表4に示す^{*}。この結果に基づき、まず式(1)と式(2)の各定数を決定す

る。図1と図2に示すように、話題維持指数と名詞の話題維持寄与指数はいずれも1から0に漸近する関数である。そこで、両指数の値が1から0へ明瞭に変化する領域(すなわち値0.5を中心とする領域)を、被験者実験の結果から得られた、話題の転換が生じやすい発話数の領域におよそ一致させるように各定数を設定する。これによって、両指数による話題の維持ないし転換の判別の分離性能が良くなる。なお、これらの式には時間項が含まれているが、対象データが対談文であるので実際の時間は使えない。そこで、話題維持指数の算出においては1発話で t が1増加するものとし、名詞の話題維持寄与指数の算出においては t_w は文字数を時間の代わりとして使用する。

話題維持指数 C_n は、話題転換直後の話題転換の検出を抑制するために用いられるので、以下の考えに基づき各定数を決定する。

- 表4から、話題転換は最短で3発話、平均12.5発話で発生し、標準偏差 σ が7.9である。
- そこで、ほとんど話題転換が起こらないと思われる、話題転換後3発話までの領域の C_n の値域を1~0.5とし、およそ(平均 $-\sigma$)発話にあたる5~6発話で $C_n \approx 0$ となるようにする。

以上から、式(1)の定数をそれぞれ $r_C = 2.0$ 、 $t_C = 6$ とした。

名詞の話題維持寄与指数 $N_{w,n}$ については以下のようにして定数を決定する。まず、語の現発話までの出現頻度に関係して、以下のように考える。

- 表4から、1つの話題は平均12.5発話の間維持される。
- そこで、ある話題語は12回出現すると話題語としての機能が不安定になる($N_{w,n} = 0.5$ となる)と仮定する。

^{*} たとえば対談文1の場合、転換点は11か所あるので、話題セグメントは12あることに注意。このため対談文1における1話題あたりの平均発話数は $136/11 \approx 12.4$ ではなく、 $136/12 \approx 11.3$ となる。

表5 学習用データに対する本手法の適用実験結果
Table 5 Recognition results for training data.

対談文	発話数	転換点数	局所結束性のみによる判別結果						最終判別成績		
			断絶群 (転/非)	不安定群 (転/非)	連結群 (転/非)	断絶群+不安定群 (転/非) 再現率 適合率		最終抽出箇所 (転/非)	再現率	適合率	
1	136	11	7 (7/0)	4 (2/2)	125 (2/123)	11 (9/2)	81.8%	81.8%	10 (9/1)	81.8%	90.0%
2	167	14	10 (8/2)	5 (2/3)	152 (4/148)	15 (10/5)	71.4%	66.7%	12 (10/2)	71.4%	83.3%
3	127	6	6 (5/1)	3 (1/2)	118 (0/118)	9 (6/3)	100.0%	66.7%	8 (6/2)	100.0%	75.0%
4	121	9	3 (3/0)	7 (4/3)	111 (2/109)	10 (7/3)	77.8%	70.0%	9 (7/2)	77.8%	77.8%
計	551	40	26 (23/3)	19 (9/10)	506 (8/498)	45 (32/13)	80.0%	71.1%	39 (32/7)	80.0%	82.1%
再現率			57.5%	22.5%	20.0%	—	—	—	—	—	—
適合率			88.5%	47.4%	1.6%	—	—	—	—	—	—

次に、語の出現間隔に関係して、以下のように考える。

- 人間の記憶の特性に基づき、語 w が前回使用された後、
 - 作動記憶の中から消えないうちに再使用されれば、語 w は話題の維持に寄与し、
 - 作動記憶から消えた後に再使用されれば、語 w は話題の維持に寄与しないと仮定する。

- Murdockら²¹⁾の実験によれば、再生率50%となる減衰時間は、作動記憶の使用量に応じて以下のとおりである。
 - 1チャンク*で73秒程度。
 - 3チャンクで7秒程度。

ただしこの実験は、維持リハーサル効果がない条件で行われている。

- 対談文を読む作業の場合にはチャンクは3以上使用されていると考えられる。また、ある語 w が使用されていない場合であっても、読んで得られる文の内容による、ある種の維持リハーサル効果があると考えられる。
- そこで、粗い仮定ではあるが、維持リハーサル効果を考慮して、語 w の使用後、20秒程度で再生率が50%程度となるものとする。実験での観察から、被験者はほぼ毎秒10.5文字程度を読む。つまり、20秒は210文字に相当する。

以上から、式(2)の定数をそれぞれ $t_N = 35$, $f_N = 2$, $r_N = 6$ と設定した。

以上の定数を用いて、被験者実験に用いた対談文の各発話について広域結束性の強さを求め、それらを被験者実験による各発話における話題の転換・維持の結果と比較したところ、式(3), (4)の各定数を以下の値にするとかなり高い精度で転換点・非転換点を分離できることが分かった。

$$\alpha_C = 0.935, \alpha_N = 0.625, \alpha_S = 0.935$$

$$\beta_C = 0.750, \beta_N = 0.535$$

α_C と β_C 、および α_N と β_N を比べると、 α_C , α_N の方がいずれも大きい値となっている。これは、断絶群の発話の方が話題が転換しやすいため、より高い話題維持傾向がないと話題が維持できないだろうという直観と合致する。

表5に、学習用データに対する処理結果を示す。表中(転/非)は、各群に含まれる転換点数と非転換点数を示している。最終判別成績とは、広域結束性も使用しての、最終的な判別成績である。認識成績の評価には再現率と適合率を用いた。すなわち、被験者が抽出した転換点数の数を T_H (表中の、転換点数)、本システムが抽出した話題転換点数の数を T_S (表中の、ある群に含まれる発話数)、システムが抽出した話題転換点のうち、被験者が抽出したものと一致しているものの数(表中の、ある群に含まれる転換点数)を T_C とすると、再現率は T_C/T_H 、適合率は T_C/T_S である。

表5において、転換点が含まれる割合(すなわち適合率)は、断絶群では非常に高く、連結群では非常に低く、不安定群では両者の中間程度であることが分かる。話題転換は全発話のうちの7%程度でしか発生しない比較的まれな現象である。このため、転換点の認識数をわずかに増やす(つまり再現率を向上させる)と、一方で誤認識数が大幅に増加し、適合率が悪化するという現象が一般に発生する。実際に表5から、断絶要因だけを用いた話題転換抽出処理を行う(つまり断絶群なら転換、それ以外は非転換と見なす)と、適合率は88.5%(断絶群の欄下部の適合率)と高いが、一方で40ある転換点のうち17もが認識もれとなるため再現率が57.5%(断絶群の欄下部の再現率)と低い。そこでこの再現率を向上させるために不安定群の発話をも話題転換点と見なすことにすると、17の認識もれのうち9か所を認識可能となり再現率は80.0%(断絶群+不安定群の欄の再現率の計)と高くなるが、一方

* チャンクとは人間の作動記憶の保持容量の単位である²²⁾。

で不安定群には非転換点が10含まれるため、適合率が71.1%（断絶群 + 不安定群の欄の適合率の計）と低くなってしまふ。しかし、最終判別成績に示すとおり、本手法で提案した広域結束性の併用によって、再現率を80.0%（最終判別成績の欄の再現率の計）と維持したまま、同時に適合率も82.1%（最終判別成績の欄の適合率の計）と、11%ほど高くすることが可能となる。

6.2 評価実験

6.1節で得たシステムパラメータを適用したプロトタイプシステムを用いて、評価用データについて実験を行った。実験では、6.1節で用いた対談文とは別の対談文5つを用い、まず各対談文について6.1節と同様の被験者実験を行い転換点を抽出した。対談文は、評価実験の場合と同じく、婦人公論の対談記事を用いた。被験者の総数はのべ8人であり、1実験あたりの被験者数は4人である。これらの被験者は、すべて学習用データにおける被験者とは異なる。表6に、被験者による転換点抽出結果を示す。一方、これらの対談文をプロトタイプシステムで処理し、その結果と被験者実験結果を比較した。結果を表7に示す。また表8には、評価用データに対する判定の詳細を示す。この表において、転換正答、誤答、非転換正答、検出洩れとは以下のような意味である。

転換正答 転換点であり、かつシステムが話題転換と判定した発話。

誤答 非転換点であるにもかかわらず、システムが誤って話題転換と判定した発話。

非転換正答 非転換点であり、かつシステムが話題転換はないと判定した発話（ただし、連結群に属する発話については記載を省略している）。

検出洩れ 転換点であるにもかかわらず、システムが誤って話題転換がないと判定した発話。

表8には、以上4つのそれぞれのカテゴリに属すると判別された発話の数を示す。なお、断絶群に分類されたものについては、さらにcue phraseを持つものか（表中では“cue”と表示）、相手特定表現を持つものか（表中では“相手特定”と表示）もあわせて示す。また、“比”として示しているのは、各カテゴリ内における、cue phrase、相手特定表現、不安定群、連結群のそれぞれに属する発話の数の、そのカテゴリに属する発話の総数に対する比である。また、cue phrase、相手特定表現、不安定群、連結群のそれぞれに含まれる転換点数と非転換点数もあわせて示す。ここに、転換点数は転換正答数と検出洩れ数との和で与えられ、また非転換点数は誤答数と非転換正答数との和で与えられる。

表6 評価用データに対する被験者による話題転換点抽出結果

Table 6 Results of topic transition point extraction by subjects for evaluation data.

対談文	発話数	転換点数	1 話題あたりの 平均発話数	標準 偏差	最短話題 構成発話数	最長話題 構成発話数
A	113	7	14.1	5.9	4	23
B	131	9	13.1	4.6	5	22
C	114	9	11.4	10.3	2	34
D	85	6	12.1	4.8	4	19
E	105	10	9.5	6.7	2	24
合計	548	41	12.2	7.1	2	34
1 発話あたりの平均文字数						69

表7 評価用データに対する本手法の適用実験結果

Table 7 Recognition results for evaluation data.

対 談 文	発 話 数	転換 点数	局所結束性のみによる判別結果						最終判別成績		
			断絶群 (転/非)	不安定群 (転/非)	連結群 (転/非)	断絶群+不安定群 (転/非) 再現率 適合率		最終抽出箇所 (転/非)	再現率	適合率	
A	113	7	9 (6/3)	8 (0/8)	96 (1/95)	17 (6/11)	85.7%	35.3%	11 (5/6)	71.4%	45.5%
B	131	9	6 (5/1)	7 (1/6)	118 (3/115)	13 (6/7)	66.7%	46.2%	9 (6/3)	66.7%	66.7%
C	114	9	6 (5/1)	6 (2/4)	102 (2/100)	12 (7/5)	77.8%	58.3%	8 (6/2)	66.7%	75.0%
D	85	6	8 (4/4)	5 (0/5)	72 (2/70)	13 (4/9)	66.7%	30.8%	7 (4/3)	66.7%	57.1%
E	105	10	7 (6/1)	5 (3/2)	93 (1/92)	12 (9/3)	90.0%	75.0%	10 (9/1)	90.0%	90.0%
計	548	41	36 (26/10)	31 (6/25)	481 (9/472)	67 (32/35)	78.5%	47.8%	45 (30/15)	73.2%	66.7%
再現率			63.4%	14.6%	21.9%	—	—	—	—	—	—
適合率			72.2%	19.4%	1.9%	—	—	—	—	—	—

表 8 評価用データに対する判定の詳細
Table 8 Detail of decision for evaluation data.

対談文	転換正答			誤答			非転換正答			検出洩れ			
	断絶群		不安定群	断絶群		不安定群	断絶群		不安定群	断絶群		不安定群	連結群
	cue	相手特定		cue	相手特定		cue	相手特定		cue	相手特定		
A	3	2	0	0	2	4	1	0	4	1	0	0	1
B	3†	3†	1	0	1	2	0	0	4	0	0	0	3
C	3	2	1	0	1	1	0	0	3	0	0	1	2
D	4	0	0	0	2	1	2	0	4	0	0	0	2
E	2	4	3	0	1	0	0	0	2	0	0	0	1
計	15	11	5	0	7	8	3	0	17	1	0	1	9
比	50.0%	36.7%	16.7%	0%	46.7%	53.3%	15.0%	0%	85.0%	9.1%	0%	9.1%	81.8%
転換点数 (転換正答数 + 検出洩れ数)			非転換点数 (誤答数 + 非転換正答数)			合計							
cue phrase	16 (15+1)			3 (0+3)			19						
相手特定	11 (11+0)			7 (7+0)			18						
不安定群	6 (5+1)			25 (8+17)			31						
連結群	9 (0+9)			472 (0+472)			31						

†対談文 B には転換点に cue phrase と相手特定表現を同時に持つ発話が 1 つ存在していたので、これを重複カウントしている。

6.3 考 察

6.3.1 明瞭な話題転換点の存在の検証

本項では、2.1.3 項で述べた、話題転換点は明瞭な話題転換点と曖昧な話題転換点に 2 分されるという予測と、明瞭な転換点抽出のための被験者数の閾値の妥当性について検証する。まず、実験における話題転換点検出傾向について見る。6.1 節と 6.2 節で示した学習用ならびに評価用の被験者実験において、1 人以上の被験者によって話題転換が起きたと判断されたすべての箇所について、何人の被験者が各箇所を話題転換ありと判断したかを表 9 に示す。1 人の被験者が話題転換を検出する平均間隔 (発話数) もあわせて示す。このように、1 人だけが転換点と判断した箇所は 94 か所 (全体の 44.5%) と最も多く、2 人が同時に転換点と判断した箇所は 36 か所 (17.1%) に減少するが、3 人以上が同時に転換点と判断する箇所は 81 か所 (38.4%) と再び増加する。

次に、被験者による話題転換点の検出傾向について理論的な予測を行う。今、ある被験者がある発話において話題転換があると判断する確率は、すべての発話について均等であると仮定する。これは、すべてが曖昧な話題転換点であると仮定することにあたる。この仮定のもとで、1 人の被験者が平均 n 発話につき 1 回話題転換があると判断する場合、ある被験者がある発話で話題転換があると判断する確率は $1/n$ である。 m 人の被験者がいる場合、ある同一発話を i 人 ($0 \leq i \leq m$) が同時に話題転換点と判断する確率 $P_i^{(m)}$ は、

$$P_i^{(m)} = {}_m C_i \left(\frac{1}{n}\right)^i \left(\frac{n-1}{n}\right)^{m-i}$$

表 9 ある箇所について話題転換ありと指示した被験者の数
Table 9 Number of subjects who regarded a certain point as a topic transition point in the experiments.

1 人だけが転換点と判断した箇所	94
2 人の被験者が転換点と判断した箇所	36
3 人以上の被験者が転換点と判断した箇所	81
1 人の被験者が話題転換を検出する平均発話数	10.56

と与えられる。この確率分布は二項分布となるが、6.1 節と 6.2 節で示した学習用ならびに評価用の被験者実験における被験者数 $m = 4 \sim 6$ と、話題転換検出間隔 $n = 2 \sim 34$ 程度では、 n がごく小さい場合を除いて、 $P_0^{(m)}$ または $P_1^{(m)}$ を極大として i の増加とともに単調減少する。特に表 9 に示した、実験における 1 人の被験者の平均話題検出間隔 $n = 10 \sim 11$ 程度では、 m にかかわらず i の増加ともなつて確率は急激に低下し、 $P_3^{(m)}$ 以上はほぼ 0 に等しくなる。

ところが、表 9 に示すとおり、実験結果では 3 人以上の被験者が一致して話題転換ありと判断する箇所が、以上の予測に比べて非常に多い。このことは、この判断の一致が偶然の一致によるものでないことを示しているといえる。つまり、話題転換点のすべてを曖昧な転換点として扱うことは誤りであり、現実には多くの人が一致して話題転換点と判断しうる発話、すなわち明瞭な話題転換点が存在し、しかもその数は少なくないことが示唆される。またこの予測と実験結果の比較から、明瞭な話題転換点を抽出するには、 n の値が小さい場合などには $P_2^{(m)}$ は無視できない程度に大きい場合があり、2 人によって一致して抽出された発話が曖昧な転換点の偶然の一致かあるいは明瞭な転換点なのかを判別しにくい可能性が残るが、 $P_3^{(m)}$ 以上

は十分に小さくなるため、3人以上の被験者が一致して話題転換と判断する箇所を抽出すれば、ほぼ偶然の一致の影響を受けずに明瞭な転換点を抽出可能となることが分かる。

以上から、被験者数に幅があっても、基本的に3人以上が一致する話題転換点を検出することで、検出精度の問題は生じないといえる。なお、実験ではより確実に明瞭な転換点のみを抽出できるよう、人数の閾値を被験者総数の60%以上とした。しかし閾値を「3人以上」としても、抽出結果に差はなかった。

6.3.2 提案した手法の処理性能

表7から分かるように、評価用データにおいても、局所結束性のみによる判別（断絶群と不安定群に属する発話ですべて話題転換と見なす）では再現率は78.5%と高いが、適合率は47.8%と低い結果となっている。一方、広域結束性をあわせて使用すると、再現率は全体として73.2%となり5.3%ほど下がるが（しかし対談文B, D, Eについては低下しない）、適合率はすべての対談文について改善され、全体で66.7%と18.9%も向上することが分かる。

次に、判別の詳細について検討する。表8から、転換点の多くは、やはり断絶群に含まれていることが分かる。ただし、従来から一般的に使用されているcue phraseを含む転換点は15点で、転換正答のうちの50.0%、全転換点の36.6%にとどまる。一方、新たに導入した断絶要因である、「相手特定表現」を含む転換点は11点あり、転換正答のうちの36.7%、全転換点の26.8%を占めており、この両者を合算することによって、断絶要因による判別成績は再現率63.4%、適合率72.2%とかなり良好な値となる。このことから、相手特定表現も基本的に断絶要因として機能することが示されたといえる。

ただし、相手特定表現を含む18の発話のうち、転換点は11しかなく、cue phraseに比べ話題転換発生の確率は低い。そこで広域結束性を用いることによって非転換点を分別すべきであるが、表8から分かる通り、相手特定表現を持つ7つの非転換点はすべて誤答となっており、分別はうまくいっていない。

誤答となった相手特定表現を含む7つの発話の内容について調べたところ、うち6例が以下のような短い発話であった。

「Aさんはどう思われますか？」

このような発話では、明らかに「～について」の部分が省略されている。したがって、共有語や名詞による維持力の判定が機能できないため、相手特定表現の存在のみが判断材料となり誤った判定を導いていると

いえる。この問題の解決のためには、本来はゼロ照応の解析技術が必要である。しかし、上記例のような相手特定表現を含むごく短い発話であり、かつ「を」格などにゼロ照応が存在することだけが検出できれば、その照応先を具体的に解決するまでもなく、そこには明確な話題語が共有語として存在するものと見なせる可能性が高い。したがってこのような場合には非転換と判断するアルゴリズムを追加することによって、誤答を減らし精度（適合率）を向上させることが可能と思われる。

相手特定表現を含む発話については、誤答における上記の例のみならず、学習用データならびに評価用データに出現する24例のうち、22発話において同一語評価値 S_n が0となっており、しかも17発話には同一語が存在しないという顕著な特徴があった。このような状況では式(3)における S_n を用いての判別は無意味であるのみならず有害であろう。また、 N_n も全般に低い値をとり、平均0.06、最大値が0.3であった。一方、 C_n についてはCue phraseの場合と分布にあまり大きな差はない。Cue phraseと相手特定表現とは、上に述べたように話題転換発生の確率に差があり、そのような2つの要因に対して、同じ式(3)によって分別を行っていることには無理がある可能性がある。相手特定表現に特化した判別を行う必要があるか、必要があれば閾値をどのような値にすれば良いかは、今後の検討課題である。

また、相手特定表現を含む発話における誤答の残り1例は次のような発話であった。

「Bさん、やっぱり通販は安いですか？」

これは先の例同様ごく短い発話であるが、共有話題語「通販」が陽に含まれていた。ただし、これは実は対談全体を通しての大局的な話題語であったため、話題維持力が低く評価され、その結果誤って転換と判断されていた。この例の場合、本当の局所的话题の話題語は、述部に含まれている「安い」であった。この発話の数発話前から局所的话题として「通信販売は安い」という話題が展開していたため、被験者はここを非転換と判断したものと思われる。このように、話題語が体言ではなく用言として出現する場合、現状の手法では判別できない。話題語としての用言をどう扱うかも今後の課題である。

一方、cue phraseを含む発話に対しては広域結束性による発話の分別はうまく機能しており、検出洩れが1つあるが、誤答はない。これは、cue phraseが使用される発話では、一般に話題語が陽に提示されるため、広域結束性がうまく機能しやすいためと考えられ

る。特に、cue phraseがあるにもかかわらず話題転換がなかったと見なされるような、従来は例外的と思われるような発話を3か所非転換正答として分別できているのは、広域結束性を導入したことの明確な効果であるといえる。このうちの1つを例示する。

- 1) 「終末観ですか。我々の周りにも終末論者はたくさんいますね。それは精神分裂病患者なんです。精神分裂病の初期の症状に、世界没落感というのがあります。」
- 2) 「終末観は僕にもあるけど、じゃあ僕も精神分裂病かな。」
- 3) 「我々の中にもあるからこそ、彼らも感じるわけですよ。(以下略)」

この例の発話2)でcue phrase「じゃあ」が現れている。単純なcue phrase検出では、「じゃあ」をcue phraseとして登録している限り、ここを話題転換点として検出する。しかし現実には、被験者によって発話2)は転換点として検出されていないし、実際この部分を読むだけでも、この「じゃあ」は話題を変える機能は弱いことが意味的には把握できるだろう。本システムによる処理では、「終末観」と「精神分裂病」、「世界没落感」の3つのキーワードが発話1)で初出であり、続く発話2)において前2者が共有語となっているため、この両語の名詞の話題維持寄与指数 $N_{w,n}$ が非常に大きくなる(ともに1.0である)。したがって、これらに基づく共有語による発話2)の話題維持寄与指数 S_n は1.0となり、しかも、発話2)にはこの2つしかキーワードがないため、名詞による話題維持寄与指数 N_n も1.0となる。この結果、広域結束性による連結力がcue phraseによる断絶力を上回るため、システムは発話2)は話題転換点ではないという正しい結論を導いている。

cue phraseを含む転換点で唯一の検出漏れは、その2発話前の発話をシステムが転換点と判断していたため(これは誤答であった)、話題維持指数 C_n が大きくなってしまっていたことが原因となっている。このように C_n はシステムの過去の誤った判断の影響を後に及ぼしてしまう点で、両刃の剣となる可能性は否定できない。しかし、断絶群と不安定群あわせて検出漏れが2か所しかないことから、 C_n によるこのような悪影響は、比較的小さいということが示唆されている。

不安定群の発話に対する広域結束性による分別もほぼ有効に機能している。不安定群には25の非転換点が含まれているが、このうちの68%にあたる17発話が広域結束性によって非転換正答として分別されている。しかも広域結束性による分別によって生じる検出

漏れは、1つ発生しているのみである。

なお、不安定群において誤答となった8つの発話の内容を調査してみたが、明確な表層の特徴を得ることはできなかった。また現在はすべて検出洩れとなっている、連結群に含まれる転換点を分別する方法も今のところ得られていない。このような発話を処理できるようにするための対策として、先に述べた用言を話題候補語として扱うことや、ゼロ照応を含めた各種の照応関係を解析することなどにより、局所結束性と広域結束性の両方を拡張する手段が考えられる。これらの拡張は、今後の課題である。

なお、以上において広域結束性による判別には、式(3)、(4)に示す単純な閾値処理を用いてきたが、他の判別分析法の導入により判別精度が向上する可能性もある。これについても今後の検討課題である。

7. おわりに

本論文では、発話の時間推移が重要な役割を持つ不特定分野の自由対談文について、局所結束性に加え時間推移情報を考慮した広域結束性を用いることにより、話題の維持傾向の強さが弱まる箇所において話題転換が生じやすいことに着目したセグメント分割手法を提案した。つまり、まず局所結束性によるカテゴリ分類によって、前発話に対する連結力の弱い発話を抽出する。さらにそれらについて、疑似的な一貫性としての広域結束性を用いて話題の維持力が弱い発話を選び出し、これらを話題転換点と見なすのである。提案したアルゴリズムを雑誌に掲載された対談文に対して適用した評価実験を行った結果、以下のことが分かった。

- 対談文においては、従来から話題転換の明示的要因として使用されるcue phraseの他に、相手特定表現も話題転換を示す要因と見なすことができる。
- 広域結束性による分別アルゴリズムは、cue phraseを含む発話と、いっさいの局所結束性の要因を含まない不安定群発話に対して有効に作用し、再現率を維持しつつ適合率を向上させることが可能である。
- 相手特定表現についても、簡単なゼロ照応検出機能を追加することによって、適合率を向上させる可能性がある。

本手法で用いている局所結束性および広域結束性は、ともに現発話とそこまでの対談でなされた発話における形態素情報および時間情報のみから得られ、現発話以後の発話についての情報をいっさい必要としない。またそれらの結束性の獲得のために特定分野に依存した知識やplan知識なども必要としない。したがって、

本手法は対談の変容性と先行不能性を侵さず、しかも目的指向型対話や特定分野の対談に対しても適用することが可能である*。

現状のプロトタイプシステムでは、本論文で示したような、書き起こされて編集を加えた対談文、あるいはオンラインディスカッションシステムを用いてのフォーマルな雰囲気での対談のような、テキストベースでのかなり正しい日本語による対談を対象としているが、今後はさらにより現実的な対談にも適用可能なものとしていきたい。しかし、現実の対談には非常に多くの文法的不適格文や省略が存在する。特に省略については、明白な話題語は省略されやすい傾向があるため、現状のアルゴリズムをそのまま適用することができない。従来の一般的な照応解析手法では通常深い意味理解が必要であり、不特定分野の自由対談には適用困難である。このための対策として、照応/省略現象に関する基本的な性質²³⁾を制約とし、体言と用言との共起関係を用いた局所的な状況把握に基づく照応解析手法を現在考案中である。

さらに、口頭での対話への適用のために、文字数や発話数による疑似的な時間情報ではなく、実時間に基づく広域結束性パラメータの決定を行いたい。また、現実の対話において重要な役割を果たしているといわれる、身振りや表情などの非言語情報の利用による判定精度の向上も試みたい。

謝辞 本研究の機会を与えてくださった(株)エイ・ティ・アール知能映像通信研究所の酒井保良会長に深く感謝いたします。また、実験やシステム構築にあたって多大なご協力をいただいた株式会社東洋情報システムの高橋誠氏ならびに三井直貴氏、また被験者実験にあたって快くご協力いただいた多くの方々に厚く御礼申し上げます。

参考文献

- 1) Boden, M.A.: Agents and Creativity, *Comm. ACM*, Vol.37, No.7, pp.117-121 (1994).
- 2) Edmonds, E.A., Candy, L., Jones, R. and Soufi, B.: Support for Collaborative Design: Agents and Emergence, *Comm. ACM*, Vol.37, No.7, pp.41-47 (1994).
- 3) 岡田謙一, 市村 哲, 松浦宣彦: グループウェアにおけるコミュニケーション支援, *情報処理*, Vol.34, No.8, pp.1028-1036 (1993).

- 4) Fish, R.S., et al.: The Video Window System in Informal Communications, *Proc. CSCW '90* (1990).
- 5) Heath, C. and Luff, P.: Disembodied Conduct: Communication through Video in a Multi-Media Office Environment, *Proc. CHI '91*, pp.25-38 (1991).
- 6) Young, L.F.: The Metaphor Machine: A Database Method for Creativity Support, *Decision Support Systems*, Vol.3, No.4, pp.309-317 (1987).
- 7) 西本一志, 角 康之, 間瀬健二: 一参加者として対話に加わる対話活性化エージェント, *信学技法*, TL96-7, pp.1-12 (1996).
- 8) 門林理恵子, 間瀬健二: 新しいコミュニケーション環境としての MetaMuseum, マルチメディア通信と分散処理ワークショップ論文集, pp.71-78, 情報処理学会 (1995).
- 9) 西本一志, 安部伸治, 間瀬健二: 不特定分野の自由展開型対話における話題転換のリアルタイム検出, 自然言語処理の応用に関するシンポジウム論文集, pp.41-48, 情報処理学会 (1995).
- 10) 西本一志, 角 康之, 間瀬健二: Augmented Informative Discussion Environment "AIDE", 第2回知能情報メディアシンポジウム予稿論文集, pp.259-266, 電子情報通信学会 (1996).
- 11) 大田雅彰, 山下洋一, 溝口理一郎: 音声対話理解のための話題の決定について, *信学技報*, SP-93-129, pp.9-16 (1994).
- 12) 加藤直人, 森元 暉: 統計的手法による談話構造解析, *人工知能学会研究会資料*, SIG-SLUD-9502-3, pp.16-23 (1995).
- 13) 宮地泰造, 伊草ひとみ, 近藤省造, 太細 孝, 古川康一: 話題管理機能を持つ対話システムの試作, *情報処理学会研究報告知識工学と人工知能*, 38-7, pp.1-8 (1985).
- 14) Grosz, B.J. and Sidner, C.L.: Attention, Intentions, And Structure of Discourse, *Computational Linguistics*, Vol.12, No.3, pp.175-204 (1986).
- 15) 竹下 敦: 対話のインタラクション構造を用いた話題の認識, *情報処理学会研究報告自然言語処理*, 87-10, pp.75-82 (1992).
- 16) 西澤信一郎, 中川裕志: 日本語の自由会話における談話構造の推定~因果関係の場合について, *情報処理学会研究報告*, 96-NL-113, pp.89-95 (1996).
- 17) Hearst, M.A. and Plaunt, C.: Subtopic Structuring for Full-Length Document Access, *Proc. SIGIR*, pp.59-68 (1993).
- 18) Salton, G.: *Automatic Text Processing: The Transformation, Analysis, and Retrieval of Information by Computer*, Addison-Wesley (1989).
- 19) Morris, J. and Hirst, G.: Lexical Cohesion

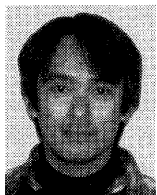
* ただし、本手法を、特定の話題分野の対話や目的指向型対話を対象とした従来手法と比較した場合、従来手法が処理対象として想定している種類の対話に対する処理精度は、一般に従来手法の方が高いであろう。

Computed by Thesaural Relations as an Indicator of the Structure of Text, *Computational Linguistics*, Vol.17, No.1, pp.21-48 (1991).

- 20) 竹下 敦, 井上孝史, 田中一男: モノログに対するブラウジング支援のための話題構造抽出, 情報処理学会論文誌, Vol.37, No.11, pp.1919-1927 (1996).
- 21) Murdock, B.B., Jr.: Short-term retention of single paired-associates, *Psychological Reports*, Vol.8, p.280 (1961).
- 22) 淵 一博, 古川康一, 溝口文雄: 知識情報処理シリーズ5 インタフェースの科学, 共立出版 (1987).
- 23) 藤澤伸二, 増山 繁, 内藤昭三: 日本語文章における照応・省略現象の基本的検討, 情報処理学会論文誌, Vol.34, No.9, pp.1909-1918 (1993).

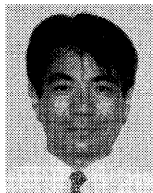
(平成 9 年 4 月 14 日受付)

(平成 10 年 3 月 6 日採録)



西本 一志 (正会員)

1987年, 京都大学大学院工学研究科機械工学専攻修士課程修了。同年松下電器産業(株)入社。1992年より(株)ATR 通信システム研究所に出向。現在(株)ATR 知能映像通信研究所第二研究室客員研究員。エージェントによる人の創造的活動の支援の研究に従事。人工知能学会, 言語処理学会各会員。



間瀬 健二 (正会員)

1979年名古屋大学工学部電気工学科卒業。1981年同大学大学院修士(情報)課程修了。工学博士。同年日本電信電話公社(現在NTT)入社。以来NTT電気通信研究所にて画像情報システムコンピュータグラフィックス画像処理とそのヒューマンインタフェースの研究に従事。1988~1989年米国MITメディア研究所客員研究員。1995年2月より(株)国際電気通信基礎技術研究所(ATR)。現在(株)ATR 知能映像通信研究所第二研究室室長。コミュニケーション支援のためのインタフェースエージェントの研究を推進している。IEEE, 電子情報通信学会各会員。



中津 良平 (正会員)

1971年京都大学大学院電子工学専攻修士課程修了。同年日本電信電話公社入社。1985年NTT複合通信研究所主幹研究員。1987年同ヒューマンインターフェイス研究所グループリーダー。1990年同基礎研究所研究企画部長。1991年同情報科学研究部長。1994年ATR経営企画部担当部長。1995年よりATR 知能映像通信研究所代表取締役。