

説得対話コーパスの構築と分析

平岡 拓也¹ Graham Neubig¹ Sakriani Sakti¹ 戸田 智基¹ 中村 哲¹

概要: 近年の対話研究において、対話参加者の行動を変化させるように能動的に働きかける対話システム（説得対話システム）が、注目を集めるようになってきた。より説得力のあるシステムを構築するうえで、人間が行う説得を理解することは重要である。本研究では、人同士の説得対話コーパスの構築と、話者の説得力に貢献する要因の分析について述べる。具体的には、3人の販売員が19人の客に対して、特定の製品を購入させようとする対話を収集し、対話行為を表わすタグを付与した。そして、このタグ付きコーパスに基づいて、1) 対話を構成する主たる対話行為についての分析と2) 被説得者の満足度と説得者の説得力に影響を与える要因について回帰分析を行った。その結果、対話中に占めるほとんどの対話行為は情報交換を表わすものであり、説得者の発話の30%が特定の製品購入のための論証であることが分かった。また、一般的な対話行為から算出した特徴量が満足度推定に、フレーミングや論証を表わす対話行為から算出した特徴量が説得力の推定に、それぞれ特に有効な予測因子であることが分かった。

キーワード: 説得対話, コーパス構築, コーパス分析

Construction and Analysis of a Persuasive Dialogue Corpus

Abstract: Persuasive dialogue systems, systems which are not passive actors, but actually try to change the thoughts or actions of dialogue participants, have gained some interest in recent dialogue literature. In order to construct more effective persuasive dialogue systems, it is important to understand how the system's human counterparts perform persuasion. In this paper, we describe the construction of a corpus of persuasive dialogues between real humans, and an analysis of the factors that contribute to the persuasiveness of the speaker. Specifically, we collect dialogue between 3 professional salespeople and 19 subjects, where the salesperson is trying to convince a customer to buy a particular product. We annotate dialogue acts of the collected corpus, and based on this annotated corpus, perform an analysis of factors that influence persuasion. The results of the analysis indicate that most common dialog acts are information exchange, and about 30% of the persuader's utterances are argumentation with framing aiming at making listener select a particular alternative. Finally, we perform a regression analysis of factors contributing to the satisfaction of the customer and persuasive power of the salesperson. We find that factors derived from dialogue acts are particularly effective predictor of satisfaction, and factors regarding framing are particularly effective predictors of persuasive power.

Keywords: Persuasive dialogue, Corpus construction, Corpus Analysis

1. はじめに

従来の対話システム研究では、システムがユーザを満足させることを抽象的な目標とする様々なタスクが取り組まれてきた。例として、ユーザの作業補助のための情報提示 [1], ニーズが明確でないユーザに有益な情報推薦 [2], ユーザの楽しませることを目的とするチャット [3] が挙げ

られる。他方、計算機がユーザの習慣や思想などを変化させようとして働きかける状況を想定した説得や騙しの研究が近年注目されている [4], [5]。これらの技術の応用例として、ユーザの悪習の改善 [6], 商品販売やインタラクティブ広告等 [4] が挙げられる。又、他の関連研究として、論証のための対話システムの対話戦略の最適化 [7], ユーザとシステムの両目標を満足する意思決定を行うようユーザを説得するシステム [8] に関する研究が挙げられる。

現在の説得対話システムは、説得の技能やユーザを満足

¹ 奈良先端科学技術大学院大学
Nara Institute of Science and Technology

させる能力において、人に及ばない。そこで、本論文では人間の説得対話における説得力と被説得者の満足に影響する要因を明確にすることを研究目的とする。説得が行われる状況としてセールス会話を想定する。この状況では、販売員（説得者）は客（被説得者）を満足させつつ、特定の商品を購入させるように対話を行う。このような説得対話に対して分析を行い、説得者と被説得者が満足する対話の特徴を解明する。

本研究の主な貢献を以下に挙げる。

- 特定のカメラを購入させることを意図した販売員と客の模擬対話コーパスの収集とタグ付けを行う。収集したコーパスと2種類のタグ付けについては3節で述べる。
- 説得対話中の説得者と被説得者の主な対話行為について分析を行う。
- 説得の成功、被説得者の満足度と対話の特徴との関係を明らかにするため回帰分析を行う。

この分析で得られた知見は、対話システムの構築だけでなく、人間の説得法に考案を行ううえでも役立つと考えられる。

2. 関連研究

説得対話コーパス収集とタグ付けに関する先行研究はいくつか存在する。例えば、Georgilaら[9]は説得及び論証対話のためのタグセットを提案している。この研究で提案されたタグセットでは、論証タグはそれらが持つ役割によってより詳細に分類されている。例えば、「論証の無効」や「論証の受諾」等の役割を持った論証タグが存在する。一方で、我々は、被説得者の嗜好に関する情報と説得者のフレーミング理論[10]に着目した論証タグを提案する。これらの情報は説得に有効であることが知られており、単に論証の役割をタグ付けすることでは捕えられない。

Nguyenら[11]は被説得者の退屈、説得の成功の要因について分析を行っている。この研究はエージェントと人間の対話を分析している、一方で我々は人と人の対話を分析している。また、この研究は、説得者の発話スタイル、被説得者の参加（エージェントと対話を行う、またはエージェントの会話を聞く）とエージェントの数の効果に着目して、分析を行っている。一方で、我々は被説得者の嗜好情報と説得者のフレーミング理論に基づいた特徴の効果に着目して分析を行っている。

更に、我々の研究はユーザの満足度と説得の成功に関して予測を行うモデルを提案している。説得対話にこのモデルを適用することで半自動的にその対話の評価をすることが出来る。

3. カメラ販売コーパスの収集

3.1 データ収集

本研究では、説得対話が行われる場面の一例として、家電販売店でのカメラ販売における販売員（説得者）と客（被説得者）の対話を想定する。販売員は客に対して、複数のカメラ（意思決定候補）の中から特定のカメラを購入（意思決定）させることを目的とする。具体的には、販売員と客の2者による対話を通して、客が5つのカメラの中から、最終的に購入するカメラを決定する。収録に先立ち、販売員には、客に特定のカメラ（説得目標）を購入させるように指示する。また、客には事前に配布したカメラのカタログの中から、どのようなカメラが欲しいかを定めるよう指示する。収録中、販売員と客はカメラのカタログを資料として適宜参照できる。いつ対話を打ち切るかについては客が決定する。

家電販売店での製品販売経験を持つ30代から40代の男女計3名を販売員役とし、20代から40代の男女計19名を客役として、対話収録を行う。収録した総対話数は34対話であり、総時間は約340分である。収録したコーパスに関する情報を表1、対話の開始部の書き起こし文の一例を表2に示す。

3.2 対話行為のタグ付け

3.2.1 対話行為タグについて

収録した対話に対して詳細な分析を行うために、複数の種類のタグを各発話文に付与した。1種類目のタグは一般的な対話行為、2種類目は特に説得行為（論証とフレーミング）をそれぞれ捕えるために利用する。形式的には収録した対話と付与したタグは以下のように定義される:

$$U = \{u_1, u_2, \dots, u_K\} \quad (1)$$

$$u_k = \langle r, g, A, F \rangle \quad (2)$$

U は対話を表わしており、発話 u_k の系列から構成される。 u_k には4種類のタグが付与されており、役割タグ r は話者が販売員の場合 SALES、客の場合 CUST の値をとる。また、後述の対話行為タグ g 、論証タグ A とフレーミングタグ F も付与されている。

一般的な対話行為を表わすタグセットとして、国際標準化機構により定められた一般目的機能 (general-purpose functions:GPF)[12] を利用する。各発話に対して一つのGPFタグを付与する。GPFタグが一意に定まらない発話は、単一のタグのみが付与された複数の発話単位に分割する。表2にGPFタグの付与例を示す。例えば、“PROPOSITIONALQ”は事実や意見が正しいかを確認する発話に対して付与される。

対話中の説得行為をとらえるために、説得、態度心理学

表 1 収録したコーパスの情報

販売員	経験	年代	対話数	総時間 (分)	語数 (販売員)	語数 (客)
A	4 年	40 歳代	10	127	33,330	6,451
B	3 年	30 歳代	12	106	32,835	7,544
C	2 年	30 歳代	12	104	24,821	7,675
合計			34	337	90,986	22,626

表 2 収録コーパスにおける対話の開始部の書き起こし文と GPF タグの付与例

話者	書き起こし文	GPF タグ
客	えーっと、カメラが欲しいんですけど、 B カメラってありますか？	PROPOSITIONALQ
販売員	あ、B カメラございますよ	CONFIRM
販売員	え、何かでもう調べて来られたんですか？	PROPOSITIONALQ
客	あー、そうですね、	CONFIRM
販売員	ネットでちょっと調べて来て	CHECKQ
客	はい、そうですね、はい	AGREEMENT

研究における知見に基づいたタグセットを定義する。人間は複数の決定要因を自身の嗜好で重み付けることで、意思決定候補を評価していることが示唆されている [13]。また、フレーミング方法によって説得力の増加に効果があることが示唆されている。本研究では、ネガティブ/ポジティブフレーミング [10], [14] に着目する。これらのフレーミングでは、感情極性を持った語で意思決定候補を修飾する。具体的には、ネガティブフレーミングはネガティブな感情極性を、ポジティブフレーミングはポジティブな感情極性を持つ語で意思決定候補を修飾する。予備実験を通して、我々は被説得者の嗜好やフレーミングが説得を行う上で重要であると仮定し、これらを考慮したタグを提案する。

付与する論証タグ A は以下に定義される:

$$A = \{a_1, a_2, \dots, a_j\} \quad (3)$$

$$a_j \in ALT \quad (4)$$

a_j は対話で扱う意思決定候補集合 ALT (本研究では5つのカメラ) の要素である。論証タグは XML に類似する以下の形式を用いて、付与対象のコーパスに記述される:

$$\langle \arg \ alt=a_j \rangle \dots \langle / \arg \rangle \quad (5)$$

また、付与されたフレーミングタグ F は以下に定義される:

$$F = \{f_1, f_2, \dots, f_i\} \quad (6)$$

$$f_i = \langle a_i, p_i, r_i \rangle \quad (7)$$

$$a_i \in ALT \quad (8)$$

$$p_i \in \{POS, NEG\} \quad (9)$$

$$r_i \in \{YES, NO\} \quad (10)$$

a_i は論証の対象である意思決定候補を表している。 p_i はフレーミングがネガティブの場合 NEG, ポジティブの場合 POS の値をとる。 r_i は論証中に被説得者の嗜好に合致した

表 3 販売員の発話に対する論証タグの付与例

$\langle \arg \ alt=A \rangle \langle \fra \ alt=A, polarity=POS, pref=NO \rangle$ (カメラ A は) ポケットに入る大きさで一眼並みの性能で撮っていただける $\langle / \fra \rangle$ っていうことが今回のポイントなんですけれども。 $\langle / \arg \rangle$

決定要因への言及が存在するかを表している。 r_i は言及が存在する場合 TRUE, 存在しない場合 FALSE の値をとる。被説得者の嗜好に合致する決定要因はアンケート結果に基づいて決定する。 f_i は XML に類似する以下の形式を用いて、付与対象のコーパスに記述される:

$$\langle \fra \ alt=a_i, polarity=p_i, pref=r_i \rangle \dots \langle / \fra \rangle \quad (11)$$

表 3 はカメラ A ($a=A$) を対象としたポジティブフレーミング ($p=POS$) を表すタグの付与例である。この例では、客は、カメラの価格をカメラ選択の決め手として重視することを、アンケートに回答しているとする。まず、販売の発話中に価格に関する記述が含まれていないため、 $r=NO$ が付与される。そして、この発話に対して一つ以上の \fra が付与されているため、発話全体に $\langle \arg \ alt=A \rangle$ を付与する。

3.2.2 タグ付けの信頼性評価

収集した全対話の 10% をランダムに選択し、付与したタグの信頼性を評価を行った。GPF と論証タグはアノテータ間の一致率を評価する。また \fra タグの値、 \fra タグの記述区間と被説得者の嗜好情報は妥当性を評価する。妥当性とは第 1 アノテータが付与したタグの内、第 2 アノテータが妥当であると判断したタグの割合である。

当初付与された 18 種類の GPF タグの一致率は 30% となり、分析を行うには低い値であった。そこで、GPF の階層的なタグの定義を考慮して、一致率の低いタグを上位タグに統合した。その結果、統合後の 6 種類のタグ (表 4) の一致率は 76% となった。この一致率は異なるタスクにおける先行研究 [15] の一致率と同程度である。これらの統合された GPF タグを後の節の分析で利用する。論証タグの

表 4 GPF タグ統合の結果

GPF tag	PROPOSITIONALQ, SETQ, INFORM, ANSWER, COMMISSIVE, DIRECTIVE
---------	---

一致率は 94%となり、妥当性については fra タグの値が 100%, fra タグの範囲が 94%, 嗜好情報が 82%であった。

4. 説得対話の評価尺度と対話の特徴

本節では説得対話の達成を評価する尺度と、それらと関係すると考えられる対話の特徴を定義する。

4.1 説得対話の評価尺度

まず説得対話の評価尺度について、対話参加者ごとの目標に着目して議論する。本対話は 2 人の参加者で構成されるので、両者の目標達成を対話の成功として定義することは自然である。そこで、本研究では説得者と被説得者それぞれの対話評の価尺度を定義する。従来の対話システムと同様に、満足度を被説得者の対話の評価尺度として利用することは適切であると考えられる [16]。一方で、説得技術の分野において、決まった評価尺度は我々の知る限り存在しない。そのため、1) 被説得者が対話の最後に説得目標を選択したか (説得の成功)、2) 対話後の被説得者の説得目標に対する意見の変化度合い (意見変化) という 2 種類の評価尺度を提案する。

被説得者に対してアンケートを行い、被説得者の満足度、意見変化と説得の成功を数値化する。

満足度 (Sat): 被説得者の対話に対する満足度。本研究では、Sat を客の満足度を表す 5 段階のスコアとして定義する (1:満足でない, 3:どちらとも言えない, 5:満足である)。

意見変化 (ΔIn): 対話後の説得目標に対する被説得者の意見の変化度合い。 ΔIn は以下に定義される:

$$\Delta In = In_{after} - In_{before} \quad (12)$$

対話の前後に説得目標の購買意欲に関して被説得者にアンケート (1:購入したくない, 3:どちらとも言えない, 5:購入したい) を実施した。 In_{before} は対話前のアンケート結果を表し、 In_{after} は対話後のアンケート結果を表す。

説得の成功 (Suc): Suc は対話の最後に被説得者が説得目標を選択した場合に 1、それ以外は 0 の値をとる。

4.2 対話の特徴

この節では、前節で定義した説得力や被説得者の満足度に影響すると考えられる対話の特徴とその定量化について述べる。これらの特徴はネガティブ/ポジティブフレーミング、被説得者の嗜好や対話行為などに基づいて計算される。

4.2.1 ネガティブ/ポジティブフレーミングに基づく特徴

ネガティブ/ポジティブフレーミングに基づく特徴を以下に定義する:

非説得目標のネガティブフレーミング比率 ($R_{neg,a \neq t}$):

説得目標 t 以外の意思決定候補に対してネガティブな事実を述べた説得者の発話の比率を表わす。

$$R_{NEG,a \neq t} = \frac{\sum_{k=1}^K \delta(\exists f \in u_k.F(f.a \neq t \wedge f.p = NEG))}{K} \quad (13)$$

δ はクロネッカーのデルタを表しており、条件式が成り立つときに 1、それ以外は 0 の値をとる。

説得目標のポジティブフレーミング比率 ($R_{pos,a=t}$): 説得目標に対してポジティブな事実を述べた説得者の発話の比率を表す。

$$R_{POS,a=t} = \frac{\sum_{k=1}^K \delta(\exists f \in u_k.F(f.a = t \wedge f.p = POS))}{K} \quad (14)$$

4.2.2 被説得者の嗜好に基づく特徴

被説得者の嗜好に基づいた 3 種類の対話特徴量を以下に定義する。

嗜好に合致する決定要因の通知 (CPD_a): 被説得者の嗜好に合致する決定要因を意思決定候補 a が満たすことを、被説得者が説得者から通知されたかを表す。

$$CPD_a = \delta(\exists f f = \langle a, POS, YES \rangle) \quad (15)$$

意思決定候補の事前評価 (PCE_a): 対話開始時点での意思決定候補 a に対する被説得者の評価。本研究では、各カメラに対して 2 値変数を用意する。客が対話の前に欲しいカメラに対応する変数は 1、そうでない変数は 0 として計算する。

説得目標の事前評価 ($PPTA$): 事前アンケートの回答結果に基づいた被説得者の説得目標に対する評価。

4.2.3 その他の特徴

上記特徴に加え、対話の総時間、論証と対話行為の頻度に基づいた特徴量を以下に定義する。

論証の回数 (I): 対話における論証タグの総出現回数 I 。

一般伝達機能の頻度 ($R_{r,g}$): 対話における各役割ごとの GPF タグの比率。

$$R_{r,g} = \frac{\sum_i \delta(u_{i=1} = \langle r, g, \bullet, \bullet \rangle)}{\sum_i \delta(u_{i=1} = \langle r, \bullet, \bullet, \bullet \rangle)} \quad (16)$$

総時間 (TT): 秒単位での総対話時間。

5. 分析

この節では、コーパスに含まれる対話行為の分析と説得に寄与する要因の回帰分析について述べる。

表 5 論証タグと一般目的機能 (GPF) の分布

	GPF						論証タグ		
	PropQ	SetQ	Commissive	Directive	Answer	Inform	Tar	NonTar	Both
販売員	14%	4%	6%	8%	16%	45%	25%	3%	3%
客	21%	2%	9%	5%	17%	37%	-	-	-

表 6 対話の前後でそれぞれ客に選ばれた意思決定候補に基づく客の満足度の分布

		最終選択		
		PT	Not PT	None
初期選択	PT	4.0	-	5.0
	Not PT	2.0	2.0	4.4
	None	4.0	2.0	3.4

5.1 対話行為の分析

まず、説得対話を構成する主な対話行為に関する一般的な分析を行う。表 5 に販売員と客の GPF タグの分布と販売員の全発話中の論証タグの割合を示す。この結果から、情報提示 (Answer, Inform) タグが販売員と客の発話の半数以上を占めていることが分かる。更に情報探査 (PropQ, SetQ) タグを考慮すると比率は約 80%となる。

販売員の全発話の 31%は論証である。詳しくは、25%が説得目標のみ、3%が説得目標以外の意思決定候補のみ、3%が説得目標と説得目標以外の意思決定候補の両方を対象とした論証となる。この結果から、説得者は説得目標以外の意思決定候補を対象とした論証をほとんど行わないことが分かる。

表 6 に、被説得者の意思決定候補の初期と最終選択により分類した平均満足度を示す。初期選択が説得目標“PT”か何もない場合“NONE”で最終選択が“PT”の場合の平均満足度が 4.0 となり、被説得者はある程度対話に満足している。一方で、初期選択が説得目標以外の意思決定候補“Not PT”で最終選択が“PT”の場合の平均満足度は 2.0 となり、被説得者は対話に十分に満足していない。この結果は、客の初期選択が説得目標か何もない場合に、説得と被説得者を満足させることが同時に可能であることを示唆している。

5.2 説得要因の回帰分析

4.1 節で述べた対話目標と 4.2 節で述べた対話特徴との関係を分析するために回帰分析を行う。初めに、ステップワイズ多項線形回帰 [17] に基づいた特徴選択を行う。この方法では、信頼性の高い特徴のみで回帰モデルが構成されるまで、回帰モデルの学習と信頼性の低い特徴の排除を反復する。本研究では、各反復で p 値が .25 以下の特徴を排除する。最終的に得られた予測因子はいずれも統計的に有意 ($p < .05$) な特徴のみで構成される。そして、特徴選択後のモデルの予測精度を 1 個抜き交差検定法で評価する。

表 7 に評価結果を示す。まず、満足度 Sat に寄与する

要因に注目すると、モデルは満足度の全変動の約 40%を説明していることが分かる。そして、モデルに含まれる 2 つの特徴量両方が販売員の GPF タグから由来している。 $R_{SALES,PROPQ}$ の重みが高いことから、販売員は多くの質問を行うことで客に満足感を感じさせていることが示唆される。また、 $R_{SALES,COMMISSIVE}$ の重みが負の値をとる理由は、 $R_{SALES,COMMISSIVE}$ は客の質問に対する回答を失敗した度合いを表しているためである。例えば、「すみません。分かりません、調べておきます。」といった発話に COMMISSIVE が付与されている。

次に、意見変化 ΔIn の回帰モデルの重みに着目する。論証タグ由来の特徴の重みが全重みの 46%を占めており最も予測に貢献している。最も重みの高い特徴は I であり、説得目標の論証を多く行うと被説得者の意見変化をより引き起こすことが示唆される。他方で、 $PPTA$ の重みは大きな負の値である。このことは説得目標を初期選択した被説得者の意見変化を引き起こすことは困難であることを示唆している。また、GPF タグ由来の特徴量の重みは全重みの 33%を占めている。特に情報交換 ($R_{CUST,ANSWER}$, $R_{CUST,INFO-PROV}$) の重みが高いため、客が多くの発話をするほど意見変化を引き起こりやすいことが示唆される。

最後に、説得成功に対するロジスティック回帰結果に着目すると、常に説得失敗を予測した場合の 68%のチャンスレートと比較して、全データの 80%が正しく分類されている。ロジスティック回帰に用いた特徴の重みに着目すると、 PCE_B の重みが相対的に高く、もし客が対話の前にカメラ B を選択していた場合、説得がより困難になることが示唆される。 CPD_B は正の重みをもつ唯一の変数であり、説得目標以外の被説得者の嗜好に合致する代替案を挙げることが、結果的に説得目標の説得力を向上させることを示唆している。なお、カメラ B のみが特徴として選択された理由として、カメラ B が他のカメラよりも多く被説得者に初期選択され、対話中の話題として登場したことが挙げられる。

これらの結果全体に着目すると、説得者は洗練された対話戦略が必要となることが分かる。表 7 から、被説得者の満足度と説得の成功の両方に影響する予測因子は存在しない。従って、説得者は両方の目標を同時に達成するように対話する必要がある。例えば、説得者は説得を成功させるために多くの論証を行いつつ、ユーザの満足度を高めるために多くの質問を行う。しかし、 TT の負の重みから分かるように、対話が長引くほど説得の成功の見込みが低くな

表 7 被説得者の満足度と意見変容の線形回帰と説得成功のロジスティック回帰と変数選択結果. 各変数は正規化されている. また, *Bias* はバイアス項であり, 常に 1 の値をとる.

	$w_0 + w_1 x_1 + \dots + w_n x_n$						R^2
<i>Sat</i>	+3.56	<i>Bias</i>	+ .501	$R_{SALES,PROPQ}$	-.509	$R_{SALES,COMMISSIVE}$.396
ΔIn	+ .920	<i>Bias</i>	-.475	$R_{NEG,a \neq t}$	+ .625	<i>I</i>	.640
	-.303	CPD_E	+ .429	$PPTA$	+ .295	PCE_C	
	+ .422	$R_{CUST,ANSWER}$	+ .464	$R_{CUST,INFO-PROV}$	+ .276	$R_{CUST,COMMISSIVE}$	
	-.368	TT					
	$w_0 + w_1 x_1 + \dots + w_n x_n$						正解率
<i>Suc</i>	-4.349	<i>Bias</i>	+2.00	CPD_B	-8.14	PCE_B	80%
	-2.12	TT					

る. ゆえに, 両方の目標を上手く達成するために, 説得者は被説得者の各代替案の嗜好を効率的かつ正確に予測する必要がある.

6. 結論

本稿では, 人同士の説得対話に対して, 説得者の説得力と被説得者の満足度に貢献する要因に着目して分析を行った. また分析を行うために, 販売員と客のコーパスを収集し, 論証タグと対話の達成を予測するための対話の特徴を定義した.

実験結果から, 対話を構成する主な対話行為は情報交換であり, 特に販売の発話の多くは論証であることが分かった. また, 論証タグ由来の特徴は説得の成功, G P F 由来の特徴は満足度の予測にそれぞれ効果的に寄与することが分かった.

我々の研究の次のステップはこれらの知見を [8] の説得対話の枠組みに組み込むことである. 加えて, この実験結果はまだ本研究で収集したコーパスに限定されているため, 他の説得タスクにおいて提案したタグと特徴の汎用性を調査する必要がある.

謝辞

本研究の一部は I2R と ATR-Trek との共同研究により助成を受けたものである.

参考文献

[1] Ward, W.: The CMU Air Travel Information Service: Understanding Spontaneous Speech, *Proceedings of the Workshop on Speech and Natural Language* (1990).
 [2] Misu, T., Sugiura, K., Ohtake, K., Hori, C., Kashioka, H., Kawai, H. and Nakamura, S.: Modeling Spoken Decision Support Dialogue and Optimization of its Dialogue Strategy, *ACM Transactions on Speech and Language Processing* (2011).
 [3] Weizenbaum, J.: ELIZA – A Computer Program for the Study of Natural Language Communication between Man and Machine, *Communications of the Association for Computing Machinery* (1966).
 [4] B.J.Fogg: *Persuasive Technology*, Morgan Kaufman (2003).
 [5] Nijholt, A., Arkin, R. C., Brault, S., Kulpa, R., Mul-

ton, F., Bideau, B., Traum, D., Hung, H., Jr, E. S., Li, D., Yu, F., Zhou, L. and Zhang, D.: Computational Deception and Noncooperation, *IEEE Intelligent Systems* (2012).
 [6] Purpura, S., Schwanda, V., Williams, K., Stubler, W. and Sengers, P.: Fit4life: the design of a persuasive technology promoting healthy behavior and ideal weight, *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (2011).
 [7] Georgila, K. and Traum, D.: Reinforcement Learning of Argumentation Dialogue Policies in Negotiation, *Proceedings of INTERSPEECH* (2011).
 [8] Hiraoka, T., Yamauchi, Y., Neubig, G., Sakti, S., Toda, T. and Nakamura, S.: Dialogue Management for Leading the Conversation in Persuasive Dialogue Systems, *Proceedings of ASRU* (2013).
 [9] Georgila, K., Arstein, R., Nazarian, A., Rushforth, M., Traum, D. and Sycara, K.: An annotation scheme for cross-cultural argumentation and persuasion dialogues, *Proceedings of the SIGDIAL* (2011).
 [10] Levin Irwin, S. L. S. and Gaeth, G. J.: All frames are not created equal: A typology and critical analysis of framing effects, *Organizational behavior and human decision processes* 76.2 (2013).
 [11] Nguyen, H., Masthoff, J. and Edwards, P.: Persuasive effects of embodied conversational agent teams, *Proceedings of Human-Computer Interaction* (2007).
 [12] 24617-2, I.: *Language resource management-Semantic annotation frame work (SemAF), Part2: Dialogue acts. ISO* (2010).
 [13] Fishbein, M.: An investigation of the relationship between beliefs about an object and the attitude toward that object, *Human relations* (1963).
 [14] Mazzotta, I. and de Rosis, F.: Artifices for Persuading to Improve Eating Habits, *AAAI Spring Symposium: Argumentation for Consumers of Healthcare* (2006).
 [15] Ohtake, K., Misu, T., Hori, C., Kashioka, H. and Nakamura, S.: Annotating dialogue acts to construct dialogue systems for consulting, *Proceedings of the 7th Workshop on Asian Language Resources. Association for Computational Linguistics* (2009).
 [16] Walker, M. A., Litman, D. J., Kamm, C. A. and Abella, A.: PARADISE: A framework for evaluating spoken dialogue agents, *Proceedings of the eighth conference on European chapter of the Association for Computational Linguistics* (1997).
 [17] Terrell, A. and Bilge, M.: A regression-based approach to modeling addressee backchannels, *Proceedings of the 13th Annual Meeting of SIGDIAL* (2012).