

ヒューマンロボットインタラクションコーパスへの 焦点アノテーションの基準と予備的分析

Criteria and Preliminary Analysis of Annotation of Focus for a Human-robot Interaction Corpus

春日 悠生¹ 井上 昂治^{2*} 山本 賢太² 高梨 克也² 河原 達也²
Yuki Kasuga¹, Koji Inoue², Kenta Yamamoto²,
Katsuya Takanashi², Tatsuya Kawahara²

¹ 京都大学 大学院人間・環境学研究科

¹ Graduate School of Human and Environmental Studies, Kyoto University

² 京都大学 大学院情報学研究科

² Graduate School of Informatics, Kyoto University

Abstract: Focus in dialogue is defined as an element that adds new information to what is already shared among the speaker and hearer (old information), and it tends to become the next dialogue topic. It is important for dialogue systems to recognize focus in a user utterance in order to generate the proper response. In this study, we present an annotation scheme of focus by considering the type of dialogue act (DA) segmented by long utterance units (LUU). Using our annotation scheme, we annotated focus for a human-robot interaction corpus collected by our research group. We report the results of the preliminary analyses of the annotation.

1 はじめに

現在の音声対話システムは特定の対話課題について実用化されているが、将来はより多くの社会的な場面において深く長い対話を扱えるようになることが期待される。我々の研究グループでは、自律型アンドロイド ERICA のための音声対話システムの研究・開発を進めているが [1, 2], 想定する対話タスクとして、傾聴、就職面接、研究室紹介、お見合いを選び、取り組んできた。これらの社会的な対話タスクにおいてより自然で深く長い対話を実現するためには、ユーザの発話を表層的に理解するのではなく、可能な限り深いレベルにおいて理解する必要がある。

こうした深い理解を実現するための要素技術の一つとして焦点解析が挙げられる。対話における焦点 (focus) とは、話し手と聞き手の間で既に共有されている情報 (旧情報 [3]) に対して新たに付加される要素であり、その後の主題の元になりやすいものである。直観的に言えば、ユーザ発話の焦点が認識できることは、その発話の「要点」を理解できているということであり、深い理解の実現にとっては応用可能性が高い技術である

といえる。また、音声対話システムにとって、ユーザ発話中の焦点を正しく認識することは、ユーザ発話を理解するだけでなく、適切な応答を生成する際にも非常に重要になる。例えば、情報案内システムにおいては、対話状態 (dialog state) を表す要素として、ユーザ発話内の焦点の有無が考慮されている [4]。アンドロイド ERICA の傾聴対話システムでも、ユーザ発話内の焦点となる単語の繰り返しや、疑問詞と焦点を組合せた掘り下げ質問を生成している [5]。ユーザ発話内の焦点を自動的に認識するためには学習データが必要となるが、焦点は対話の状況や文脈に依存する。したがって、明確なアノテーション基準とこれに従って人手で付与された高精度のアノテーションが求められる。そこで本研究では、自然な対話の実現を指向した焦点のアノテーション基準を提案する。

2 焦点に関する理論的背景

焦点を定義するにあたり、言語学における語用論や談話分析において、これまでに蓄積されてきた談話における焦点に関する知見を援用する。焦点の定義は研究によって多少異なるが、およそ「伝達される文内に

*連絡先：京都大学 大学院情報学研究科 知能情報学専攻
京都市左京区吉田本町
E-mail: k.inoue@i.kyoto-u.ac.jp

おける、話し手と聞き手によって共有されている情報¹に対して新たに付加される要素」を指す。例えば(1)において、(2a)への返答としては the pictures が、(2b)への返答としては Mary が焦点であるとされる。この場合、(1)は先行する質問によって異なる焦点を持ち、英語では音調の差で区別されるが、日本語のように語順が比較的柔軟な言語では、焦点となる要素は文中のなるべく後方に置かれやすい。

- (1) John showed Mary the pictures.
(2) a. What did John show Mary?
b. Who did John show the pictures?

(文献 [8] 244 頁)

文の形式が同じであっても、音調による焦点の違いが命題に影響を与える場合もある。(3a)と(3b)は文は同じであるが、焦点の違いにより only によって修飾される語が異なっている。焦点の箇所は「《・》」で表されており、(3a)では the pictures, (3b)では Mary がそれぞれ焦点である。

- (3) a. John only showed Mary 《the pictures》.
b. John only showed 《Mary》 the pictures.

(文献 [8] 244 頁)

コーパスに対して焦点アノテーションを行っている先行研究では、ドイツ語学習者のための設問とそれに対する学習者の返答を収録した CREG Corpus を用いて、返答部分に対する焦点アノテーションの基準が定められている [9]。ただし、この研究では、質問に対する返答部分へのアノテーションが主な対象となっている。これに対して本研究では、返答部分だけでなく、対話に現れるその他の発話に対しても焦点付与の判断が可能となるような基準の作成を試みる。

3 焦点アノテーションの基準

本研究では、長い発話単位 (Long Utterance Unit: LUU) [10] で分割される各発話に対して、対話行為 (Dialogue Act: DA) [11] 毎に異なる基準を用いて焦点を付与する。焦点は理論的には文を単位として付与されるものであるが、対話における文の判定は困難であるため、これに相当する単位として LUU を用いる。また、従来の焦点アノテーションが質問への返答に限られていたことから、対話における質問・言明・応答などの行為の種類に応じて焦点の性質が異なること

¹この情報を指す用語は研究者によって異なるが、旧情報 (old information) [3], 主題 (topic) [6], 前提 (presupposition) [7] などと呼ばれることが多い。

が予想される。一方、本研究で用いる対話コーパスには、LUU ごとに対話行為のラベルが付与されているため、これを参照しながら焦点の情報を付与した。以下では、焦点アノテーションの基準について述べる。はじめに、対話行為の種類に依らない一般的な原則について述べ、その後、対話行為ごとの基準を述べる。また、例文などにおいて焦点の箇所は「《・》」で表す。

3.1 原則

焦点の定義は、「話し手と聞き手の間で既に共有されている情報 (旧情報) に対して新たに付加される要素」とする。ただし、こうした要素が複数ある場合は、韻律的に強勢が置かれているものを選択する。各 LUU に対して、1 個もしくは 0 個の焦点を付与する。ただし、対比、並列 (等位接続)、列挙、選択質問などの場合は、例外的に 1 つの LUU に対して、複数個の焦点を付与してもよい。焦点の最小単位は形態素とし、1 つの焦点の範囲はできるだけ短いものにする。また、情報の授受に関する行為でない行為には、焦点を付与しない。

3.1.1 焦点の種類

焦点は原則として、(i) 名詞焦点、(ii) 述語焦点のいずれかとする。名詞焦点の際は、とくに強勢が置かれな限りは助詞は焦点に入れない。述語焦点の際は、言い切り可能な語幹 (+否定) を焦点とする。モダリティはそこが強勢されていない限りは焦点としない。(例「《大きい》です」、「《食べた》かもしれない」、「《走ってない》」、「《綺麗》だった」) 特別な場合として、そこだけが新情報として韻律的に卓立されている場合は、(iii) 修飾語焦点、(iv) 助詞のみの焦点も認める。

焦点は「短く付与する」ことを原則とするが、名詞と述語、修飾語と被修飾語などがともに新情報であり、かつ片方だけを焦点とすると発話の意図が通らなくなる場合は、(v) 名詞+述語焦点、(vi) 修飾語+被修飾語焦点のようなより大きな単位を焦点とすることも認める。このような例は、取り立て助詞による限定的な否定や、逸脱的な修飾語を用いる際によく見られる (例「《プロになれるほどは強くない》です」、「昨日、《虹色の蛙》を見ました」)。

3.2 対話行為ごとの基準

文献 [11] で定義される対話行為には、言明 (statement)、質問 (question)、応答 (response)、行動 (action)、その他 (other) といった上位分類があり、さらにそれぞれの下位分類がある。以下ではこれらの上位分類ごとに焦点の付与基準を述べる。さらに下位分類において基準が異なる場合は、下位の基準を設ける。

3.2.1 言明 (statement)

言明は、「問い返し可能性」を原則的な基準として焦点を付与する。例えば、「昨日、フェリーで別府に行ったんです」という発話は、特別な文脈を設定しなければ「別府ですか」という問い返しが自然であるといえるため「別府」が焦点となる。しかし、いつも飛行機で別府に行っている人との会話であれば「フェリーですか」という問い返しが自然であるため、この場合は「フェリー」が焦点となる。このように、言明の焦点は発話の文脈に依存する。以下では順に、(i) 単文、(ii) 等位的な複文、(iii) 従属的な複文、(iv) 補文節を項としてとる述語の場合について述べる。

単文 新情報が1つしかない場合はそれを焦点とする。新情報が複数ある場合には、韻律的な卓立や問い返し可能性の観点から際立っているものを焦点とする。3.1.1節で見たように、より大きな単位を焦点にすることも認める。

等位的な複文 前件と後件のそれぞれに対して、単文の基準に準じて焦点を付与する（例「週末は、《買い物》に行ったり、《料理》をしたりします」）。

従属的な複文 従属節が旧情報の場合は、主節のみに焦点を付与する（例「初めての海外旅行だったので、《ソウル》に行ったんです」）。ただし従属節も新情報である場合は、従属節と主節のそれぞれに単文の基準に準じて焦点を付与する。

補文節を項としてとる述語 「と思う」「と言う」のような軽い主動詞の場合は、補文節内で単文の基準に準じて焦点を付与する。一方、「が大変だ」などの比較的重い主動詞の場合は、主動詞を述語焦点として認定する。

3.2.2 質問 (question)

質問は、大きく命題質問 (propositional question)、選択質問 (choice question)、集合質問 (set question)、確認質問 (check question) の4つに下位分類される。以下ではそれぞれの焦点付与基準を述べる。

命題質問 相手に真偽を確認する質問であり、真偽を問うている部分、すなわち否定されうる部分を焦点とする（例「ERICAさんは、今《学生》ですか」、「《イクラを食べたことがあります》か」）。確認質問もこの基準に準ずる。

選択質問 相手に複数の選択肢から正しい選択肢を選ばせる質問であり、例外的に複数の選択肢をそれぞれ焦点とする（例「太郎さんは《学生》ですか、それとも《社会人》ですか」）。

集合質問 相手に不完全な命題の空所を埋めさせる質問であり、答えとなるべき疑問詞を焦点とする（例「今は《どこ》で働いていますか」）。ただし、「どこか」「何か」などの「疑問詞+カ」は単なる副詞とみなし、疑問詞として扱わない。

疑問詞が変項名詞句 [12] を修飾している場合は、その名詞句を含めた範囲を焦点とする（例「《どういった本》が好きですか」「その映画は《どういった風》でしたか」）。また、変項名詞句のみによる集合質問の場合は、当該名詞句を焦点とする（例「《行きたいところ》ってありますか」「何か《スポーツ》はされてましたか」）。

3.2.3 応答 (response)

応答は、返答 (answer)、同意 (agreement)、不同意 (disagreement)、訂正 (correction)、受諾 (accept)、拒否 (decline) の6つに下位分類される。このうち、受諾と拒否は情報の授受に関する行為ではないため、焦点は付与しない。

命題質問・確認質問への返答 命題の真偽が問われているため、真偽を確定させる述語部分を焦点とする。原則として内容語を焦点として認定するが、内容語が省略された場合は、それに準ずるものに焦点を付与する（例「はい、《食べました》」、「いや、《食べてない》です」、「はい、《そう》です」、「《はい》」）。ただし、条件付きの返答の場合、条件節は新情報となるが、例外的に「接続詞+真偽を確定する述語部分」のみを焦点とする（例 質問「リンゴは好きですか」返答「オープンで焼いたもの《なら好き》です」）。

選択質問・集合質問への返答 質問文中の命題内の空所が問われているため、選択質問の場合は選択された部分が焦点となる。集合質問の場合は、疑問詞に対応する部分が焦点となる（例 質問「どこで働いていますか」返答「《京都》で働いています」）。

同意・不同意 命題への同意・不同意であるため、その対応する述語部分に焦点を付与する（例 質問「昨日地震があったじゃないですか」返答「《ありました》ね」、言明「寒いですね」返答「《そう》ですね」、言明「お若いですね」返答「《そんなことない》ですよ」）。

訂正 訂正した部分を焦点とする。形態素より小さな単位や音調などが訂正の対象となる場合もあるが、今回は形態素を最小単位として焦点を付与する。

3.2.4 行動 (action)

行動は、申し出 (offer)、約束 (promise)、提案 (suggestion)、依頼 (request) の4つに下位分類されている。しかしこれらは情報の授受に関する行為ではない

ため、原則として焦点は存在しないと考え、付与しない。ただし例外的に、取り立て助詞などを用いて新情報が付加されている場合が存在するため、その場合は名詞焦点として認定する(例「《本》をお持ちでしょうか」「じゃあぜひ《東野圭吾》は読んでみてください」)。このような例はあくまでも名詞焦点に限り、述語焦点は原理的に存在しない。

3.2.5 その他 (other)

その他に分類される行為には、相槌、挨拶、感謝、謝罪、第三位置的応答(「そうですか」など)、評価発話などが挙げられる。これらは情報の授受に関する行為ではないため、原則的には焦点は付与しないが、評価発話だけは情報提供の側面があるため、焦点付与を行う。以下では、評価発話とその返答に対する焦点付与の基準を述べる。

評価発話 形容詞部分に焦点を付与する(例「それは《すごい》ですね」)。

評価発話への返答 まず、肯定的返答に関しては、内容語を伴った第二評価が現れた場合はその内容語に焦点を付与する(例「《すごい》ですよ」)。ただし、それ以外の「はい」「そうです」などの返答には焦点を付与しない。次に、否定的返答に関しては、否定部分に焦点を付与する(例 評価発話「すごいですね」返答「いや、《そんなことない》ですよ」)。ただし、「いえいえ」などの儀礼的のみなされる返答には焦点を付与しない。

3.3 その他の細則

以下では、上記で扱うことのできなかった例に関する細則を述べる。

同格 焦点にあたる同格の名詞が並列して現れる場合は、より具体的なほうに焦点を付与する(例「あのアニメの4話目の、《主人公が走ってるやつ》です」)。

畳語 焦点にあたる語が連続して現れる場合、(i) 強調のための畳語なら前者に焦点を付与する(例「《そう》そうそうそう」、「これが《重たくて》重たくてね」)。一方、(ii) 言い換えによる畳語なら、後者に焦点を付与する(例「出身は宮崎、九州の《宮崎県》です」、「食べた、《食べてきた》んです」)。

要約 今まで自分が話したことを要約するような発話は、内容としては旧情報であるが、談話的に半活性的情報を活性化させる機能 [13] を持っているため、まとめ表現に焦点を付与する(例「まあ、《そういうこと》がありました」)。

分割された焦点 名詞+述語焦点をとる際などに、焦点でない語によって焦点が分割される場合がある。このような場合は例外的に、間の部分を含めて全体を焦点とする。

含意 言いさしなどで、焦点部分が省略される場合がある。陽に表れていないものに焦点を付与することはできないので、この場合は焦点を付与しない(例 質問「そのゲームはよくやっていますか」返答「あんまりおもしろくないので(焦点なし)」)。

独言 独言には焦点を付与しない。

4 ヒューマンロボットインタラクションコーパスへのアノテーション

第3節の基準に従い、我々が収録を進めているヒューマンロボットインタラクションコーパスに対して焦点のアノテーションを実施した。予備的な分析結果を以下に述べる。

4.1 コーパスの概要

本対話コーパスには、被験者とアンドロイド ERICA の一対一の対話が収録されている。ERICA は別室にいるオペレータによって操作されており、オペレータの話した音声を ERICA 本体に配置したマイクからそのまま再生した。したがって、音声に関しては人間どうしの自然な対話に近いものといえる。また、ERICA の視線、うなずき、ジェスチャなどの非言語動作はオペレータの手元のコントローラで操作してもらった。

対話のタスクは傾聴、お見合い、就職面接の3種類である [14]。傾聴では、ERICA は聞き手として、語り手である被験者の話を促すような聞き手応答を適宜発話した。お見合いでは、ERICA は初対面会話の練習相手としてふるまった。就職面接では、ERICA は面接官役として応募者役の被験者に対して様々な質問を行った。これらの対話タスクを選択した理由は、ERICA と被験者との間での対話の主導権のバランスが幅広くカバーされるようにするためである。今回焦点アノテーションを行ったセッション数は、傾聴が9、お見合いが18、就職面接が14である。1セッションあたりの平均時間は、傾聴が11分17秒、お見合いが10分55秒、就職面接が8分48秒であった。

4.2 アノテーション結果の傾向

今回対象としたセッションには既に長い発話単位(LUU)と対話行為(下位分類を含む) [11] のアノテーションがなされており、これらの情報を参照しながら焦点アノテーションを実施した。はじめに、LUU と焦点の頻度を対話行為の大分類毎に調べた。表1に結果を示す。質問、言明では焦点の数はLUUと同程度、応答では約8割、行動では約半数となった。「その他」で焦点なしとなったものの多くは相槌や「そうですか」などの第三位置的応答である。

次に、第3.1.1節で述べた焦点の種類について分析した。今回のアノテーションでは、焦点の種類自体は記述していないため、形態素解析によりおよその分類を試みた。ここでは、「名詞のみ（名詞焦点）、動詞のみ（述語焦点）、名詞+動詞（名詞+述語焦点）、形容詞のみ（述語焦点または修飾語焦点）、形容詞+名詞（修飾語+被修飾語焦点）、助詞のみ（助詞のみ焦点）、それ以外」とした。ただし、助詞や助動詞が、名詞や動詞に付随する場合には、名詞や動詞にまとめて扱った。形態素解析には MeCab²を用いた。辞書には mecab-ipadic-NEologd³を用いた。結果を表2に示す。全体の4割程度が名詞のみの焦点となり、続いて、「名詞+動詞」である名詞+述語焦点が多かった。これらだけでなく、動詞のみ、形容詞のみも十分に多い。また、その他には副詞が含まれる場合が多くみられた。

以上の結果より、各対話行為において、十分な比率でかつできるだけ短い範囲の焦点を特定できていることがうかがえる。焦点ができるだけ短い範囲でかつ確かなものであることは、対話システムにおける応答生成という応用的観点からも重要であるといえる。例えば、第1節で述べたERICAの傾聴対話システム[5]では、ユーザ発話の多くは言明で、その焦点は「名詞」または「形容詞+名詞」と仮定している。今回のアノテーションにより、このような焦点を自動で認識できるようになれば、これを用いて的確な繰り返し応答や掘り下げ質問が生成できるようになる。

4.3 アノテーションの例

各対話タスクの一部の対話を焦点アノテーションの例とともに以下に示す。ただし、ERICAと被験者の発話をそれぞれEとSで表す。LUUの境界は「/」で表す。各LUUの末尾にそのLUUの対話行為の大分類を示す。言明は[s]、質問は[q]、応答は[r]、行動は[a]、その他は[o]でそれぞれ表す。なお、一部の相槌、フィラー、言い淀みは対話内容の理解し易さのため省略してある。また、読点は読み易さのために便宜的に付与したものである。

4.3.1 傾聴

傾聴では、被験者（語り手）の言明とそれに対するERICA（聞き手）の反応（評価発話など）が多くなる傾向にある。語り手（ユーザ）の発話の焦点を的確に捉えることで、聞き手（システム）は、効果的な繰り返し応答、掘り下げ質問、評価応答（「いいですね」など）が生成できるようになると期待される。

S1 今日は《京都検定》のお話をしようかと思っておりますが [s]

²<https://taku910.github.io/mecab/>

³<https://github.com/neologd/mecab-ipadic-neologd>

表 1: 対話行為の大分類毎の LUU と焦点の頻度

対話行為	#LUU	#焦点
言明 (statement)	2,177	1,986
質問 (question)	949	916
応答 (response)	1,556	1,268
行動 (action)	141	78
その他 (other)	8,606	274
計	13,429	4,522

表 2: 形態素による焦点の種類分布

焦点の形態素	頻度
名詞	1,814
動詞	235
形容詞	272
助詞	16
名詞+動詞	672
形容詞+名詞	158
その他	1,355
計	4,522

- E1** 《京都検定》ですか [q]
S2 はい [r] / 《京都の文化とか歴史》に関する検定試験、ございまして [s]
E2 あ [o] / ああー [o]
S3 それを《一昨日》受験しましたので [s]
E3 あ [o] / そうなんですね [o]
S4 はあい [o] / これは、あの一、《なぜ》受験したかと申しますと [s] / あの一、アルバイトの、するのにですね《観光ガイド》をするのに、必要な、あの、資格の一つなものですから [s] / 《それで受験をしました》 [s]
E4 はあー [o] / なるほど [o] / 《観光ガイドをされている》んですね [s]

4.3.2 お見合い

お見合いでは両者の質問、応答、言明がバランスよく出現する傾向にある。ユーザの質問中の焦点を的確に捉えることで、より正確でかつ具体的な応答が生成できるようになると期待される。

- S1** えっと、休日は《何》をされていますか [q]
E1 お休みの日はだいたい《お掃除》をしたり、まあ、《洗濯》をしたりしています [r]
S2 へえー [o] / 結構《インドア》なほうですか [q]
E2 そうですね [r] / あまり《活発に外に出ることはない》ですね [r]

- S3** そうなんです [o] / 僕も休みの日は《全然》外に出ないです [s]
- E3** そうなんです [o] / お家の中で《何》をさけますか [q]
- S4** ああ [o] / でも、暇な時は《映画》とか見たりしてますね [r] / 基本的には [r]
- E4** そうなんです [o] / 《私も》映画好きです [s]
- S5** そうですか [o] / 《最近見た映画》とかありますか [q]
- E5** 最近、あの、DVDというか、あの、動画配信サイトで見たんですけども [r] / あの、《神様メール》というベルギーの映画だったかなを見ました [r]
- S6** 《どういうあらすじ》とか覚えてらっしゃいますか [q]

4.3.3 就職面接

就職面接ではERICA（面接官）の質問と被験者（応募者）の応答や言明が中心となる傾向にある。応募者（ユーザ）の応答や言明中の焦点を的確に捉えることで、面接官（システム）はより効果的な掘り下げ質問が生成できるようになると期待される。

- E1** それではあなたの《志望動機》を教えてください [a]
- S1** はい [r] / 御社は《コンサル》と《プログラミング》、この二つを二大業務として挙げてらっしゃいますね [s] / 私は《心理学》と《プログラミング》を大学で学んできました [s] / 心理学では相手の《本音を聞き出す手法》を学び [s] / そしてプログラミングではより《効率の良いそして役立つソフトウェア》を作るすべを学んできました [s] / この二つの心理学とプログラミングはコンサルとおたくのやっている、あ、御社のやっているプログラミングにすごく《貢献できる》のではないかと思ひこちらを志望しました [s]
- E2** はい [o] / ありがとうございます [o] / 《なぜ》そのお仕事をしようと思われたのですか [q]
- S2** 私は昔からコンピューターとか、えーとー、ファミコンですね、《ゲーム機械》が大好きでした [r] / そこから《どうしてこれらの機械が動いているか、動いているんだろう》と思って調べていくと、プログラミングという技術によって動いているということを知ったのでプログラムを組んだりする仕事に《惹かれて》って [r] / 最終的に《クライアントさんと相談をしながらいいものを作り上げていく》という仕事に興味を惹かれました [r]
- E3** 分かりました [o] / それでは今おっしゃったプログラミングという言葉について《もう少し詳しく》説明してください [a]

5 おわりに

本稿では、対話における焦点アノテーションの基準について述べた。長い発話単位（Long Utterance Unit）で分割される各発話に対して、対話行為（Dialog Act）毎に異なる基準を作成した。我々が収録を進めているヒューマンロボットインタラクションコーパスに対して、上記の基準で焦点アノテーションを実施した。分析の結果、LUUの数に対して、質問、言明といった対話行為では焦点の数はLUUと同程度、応答では焦点の数はLUUの数の約8割となった。焦点の形態素を調べたところ、4割ほどが名詞のみ、次いで1.5割ほどが名詞＋助詞＋動詞という形態素となった。今後は、アノテーションデータを用いて焦点の自動認識に取り組む予定である。

謝辞

本研究は、JST ERATO 石黒共生ヒューマンロボットインタラクションプロジェクト JPMJER1401 の支援を受けて実施した。

参考文献

- [1] Koji Inoue, et al. Talking with ERICA, an autonomous android. In *SIGDIAL*, pp. 212–215, 2016.
- [2] Tatsuya Kawahara. Spoken dialogue system for a human-like conversational robot ERICA. In *IWSDS*, 2018.
- [3] 久野暉. 談話の文法. 大修館書店, 東京, 1978.
- [4] Koichiro Yoshino and Tatsuya Kawahara. Conversational system for information navigation based on POMDP with user focus tracking. *Computer Speech & Language*, Vol. 34, No. 1, pp. 275–291, 2015.
- [5] Divesh Lala, et al. Attentive listening system with backchanneling, response generation and flexible turn-taking. In *SIGDIAL*, pp. 127–136, 2017.
- [6] Knut Lambrecht. *Information structure and sentence form: Topic, focus, and the mental representations of discourse referents*. Cambridge University Press, 1994.
- [7] Noam Chomsky. *Deep structure, surface structure and semantic interpretation*. Indiana University Linguistics Club, 1969.
- [8] Manfred Krifka. Basic notions of information structure. *Acta Linguistica Hungarica*, Vol. 55, No. 3–4, pp. 243–276, 2008.
- [9] Ramon Ziai and Detmar Meurers. Focus annotation in reading comprehension data. In *The 8-th Linguistic Annotation Workshop (LAW VIII)*, Vol. 8, pp. 159–168, 2014.
- [10] Yasuharu Den, et al. Two-level annotation of utterance-units in japanese dialogs: An empirically emerged scheme. In *LREC*, 2010.
- [11] Harry Bunt, et al. Towards a standard for dialogue act annotation. In *LREC*, pp. 2548–2555, 2010.
- [12] 西山佑司. 日本語名詞句の意味論と語用論: 指示的名詞句と非指示的名詞句. ひつじ書房, 東京, 2003.
- [13] Wallace Chafe. *Discourse, consciousness, and time: The flow and displacement of conscious experience in speaking and writing*. University of Chicago Press, 1994.
- [14] 河原達也. アンドロイドERICAによる人間レベルの音声対話. 人工知能学会研究会資料, SLUD-B802, 2018.