

# ことばとは何か： 人の言葉と人工知能のコトバの違いを考える

吉田 江依子

## 1. はじめに

名古屋工業大学では生涯学習活動の一環として、全学的な公開講座を毎年開催している。平成 29 年度は「人口知能の諸相」というタイトルで実施された。下記は本講座全体の概要である。<sup>1</sup>

「人口知能」という言葉はずいぶん昔から使われてきましたが、以前は SF 的な印象が強いものでした。1997 年にチェスの世界チャンピオンがコンピュータに敗れたニュースの衝撃によって、この言葉の現実味が少し感じられるようになったと思います。近年になり、人工知能に関する話題はより多くまた身近なものとなり、それに伴い、「人口知能」は「AI」と略記されるようになりました。ごく最近では、たとえば、自動車の自動運転技術の開発についてのニュースレポートの部分が奪われるかもしれないとも言われています。そして、人工知能を利用した技術は様々な方面で開発・実用化が試みられています。名工大においても、いろいろな観点から人口知能およびその周辺分野の研究が行われてきました。平成 29 年度の全学公開講座は、名工大で行われているそれらの研究に関連するいくつかの講演を、市民の方々に紹介できればと考えています。

(名古屋工業大学ホームページより)

概要に述べられているように、本講座の目的は名古屋工業大学で行われている人工知能およびその周辺分野の研究を紹介することであり、人工知能研究に関わる 5 名が講師として登壇した。筆者は言語学を専門とする立場から、現在の人工知能によって表出されるコトバの仕組みについて概観し、さらにそれを人間の言葉の特性と比較することによって、人工知能のコトバとは何かについて

考察をする。<sup>2</sup> 本稿の内容はその際の発表内容をもとに展開しているが、新たに人工知能のコトバと人間の言語のインターフェイスについて今後の展望についても考察を行っている。

本稿の構成は以下の通りである。2節では、人間の言葉の特性を内的側面および言語習得の観点から論じ、構造階層性および普遍性に特に着目する。3節で人工知能を利用したチャットボットの仕組みとの比較を行い、4節で両者の違いについてまとめの考察を行う。5節では運用面における人間言語の特性を示し、それを応用することでチャットボットとの会話が可能になる可能性について言及する。

## 2. 人間の言葉の特性

### 2.1 内的側面から見る人間の言葉の特性

自明のことのようではあるが、人間の言葉は人間という種のみに与えられた特有のもので他のどの種もそれを共有することはない。それは人間のような发声器官や音声体系を持たないからという意味ではない。人間以外の生物、鳥、昆虫、動物もいろいろな手段を使って情報を伝達することができる。その情報伝達の方法を比喩的に「ことば」という人もいるかもしれない。しかし、それを考慮にいれたとしても人間の言葉には人間の言葉にしかない様々な特性がある。

第一に、(1)に示したように人間の言葉には時間、空間を超えた事象を表現できる超越性という特徴がある。

- (1)
    - a. あそこに食べ物が昨日はあった。 [過去の出来事]
    - b. 明日は、敵は来ないでしょう。 [未来についての予測]
    - c. もうバナナはないよ（本当は残っている） [否定、嘘]
    - d. ボスはお前のことが好きなんだ。 [第三者の感情]
- (中島・外池(1994: 31))

この超越性という特性は人間の言葉にのみ存在するものである。この特性のために、人間は空想や創造の世界も自由に表現することができる。たい

ていの動物は、直接目の前にある事柄についてしか伝達ができない。鳥は危険が迫っている時だけ危険を告げる鳴き声を発する。彼らは時間と場所の離れたところにある危険について情報を送ることはできない。

第二に、人間の言葉には創造性がある。創造性とは新しいことを伝達する特性であり、例えば(2)のように、人はそれまでに経験したことのないようなことでも表現できるし、また聞き手もそれを理解する力を持っている。

- (2) 太郎君は手のひらに載るくらいの小さな籠の中に、お父さんが動物園から  
もらってきた象2頭と太郎君が夢の中で捕まえたカッパ3匹を飼っている。

(ibid.: 32)

人間の言葉には創造性があるので、実生活とかけ離れた突飛な内容を持つ文を作ることが出来、また理解することもできる。動物の伝達方法にはこのような豊かな創造性はなく、大抵の動物は非常に限られた数の伝達内容を送ったり受けたりしているにすぎない。

第三の特性は、恣意性である。恣意性とは両者の間に必然性の見当らない関係のことを言うが、人間のことばとそれが指示するものとの関係は必然的なものではなく、何の根拠もない偶然的なものである。例えばワンワンとなく4本足の全身毛でおおわれている動物を、日本語では「いぬ」という音連続で表すが、同じものを英語では dog、フランス語では chien という音連続で表しているわけで、音とそれが表すものとの結びつきには何の意味もない。この恣意性という特性は一見したところ、人が新しい知識・概念に触れるたびにそれを表す新しい「語」を学習しなければならないという面倒な短所に見えるかもしれないが、一方で、この恣意性は人間の言葉を豊かなものにしてくれている重要な特性となる。恣意性があることによって、人は形のない抽象概念や行動・状態の描写を表すことが可能になる。恣意性があればこそ、人間は言葉を使ってさらに抽象的な概念を構築し、知的活動を発展させることができるるのである。

第四の特性は規則性である。人間の言葉には規則性が存在する。例えば語句のつながりに関して、(3)のようなつながりは許すが、(4)のようなつながりはできない。

- (3) a. 太郎 — が — 私たち — を — 呼んだ  
b. 私たち — が — 太郎 — を — 呼んだ

- (4) \*が — 太郎 — 呼んだ — を — 私たち (ibid: 35)

この文法性の違いは、単語同士の結合の仕方には規則があることを示している。また、こうした規則の存在は文だけでなく単語レベル、音レベルにもある。例えば、「私たち」という句は2つの部分に分けられるがこの場合も、(5a)のような配列は可能だが(5b)の配列是不可能になっている。また、「だ」「ん」「ご」という3つの音を結合させる場合、(6a)のように「だんご」という結びつきは可能だが、「んごだ」は不適格である。これは日本語では、「ん」という音で始まる音の配列の規則がないからである。

(5) 句レベル

- a. 私 — たち
- b. \*たち — 私

(ibid.)

(6) 音レベル

- a. だ — ん — ご
- b. \*ん — ご — だ

第五のそして最大の特性は階層構造性である。言葉は音声（二次的には文字）でもって意味（または概念）を表現することであるが、両者が直接結びついているわけではない。人間言語には音と意味の間に（統語）構造が介在していると言われている。

(7)



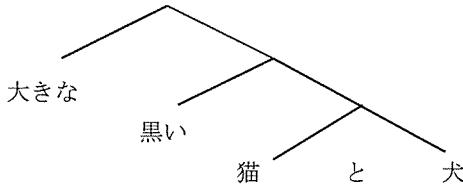
そしてこの構造は線状ではなく階層を持っている。例えば、「大きな黒い猫と犬」という表現を考えてみましょう。これは音のつながりとしては、オオキナのあとにクロイという語が続き、そのあとに、ネコという語、そして最後にイヌという語が続くひとつの音のつながりにすぎない。

- (8) オオキナ + クロイ + ネコ + ト + イヌ

しかし、この音のつながりには3つの解釈が可能である。一つは、大きくて黒い猫と大きくて黒い犬、もうひとつは大きな黒猫と黒くはないが大きな犬という解釈、3つめは黒猫だけが大きくて犬は大きくも黒くもないという解釈である。この意味の違いは、人間の言語は、線状的な横のつながりだけでできているのではなくそれぞれの語句の結びつき方に強弱があり、それぞれがまとまって、さらに別の語句と結びついて構造を作っていることを示唆している。

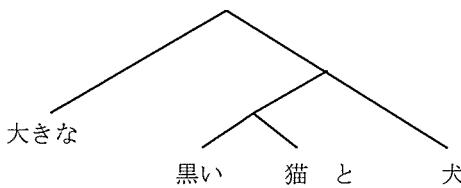
例えば、大きくて黒い猫と大きくて黒い犬という意味は、(9)で示したように、「猫」と「犬」が一つのまとまりをしており、それに「黒い」が結びつき、「黒い猫と犬」となり、それにさらに「大きな」がついていると考えられる。

(9)



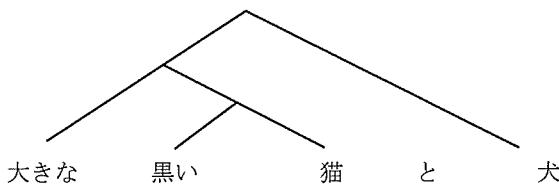
次に、大きな黒猫と黒くはない大きな犬という解釈は、(10)に示したように、「黒い」と「猫」がまず結びつきそれが犬と結びつき、全体が「大きな」と結びつく。

(10)



3つめの 黒猫だけが大きくて犬は大きくも黒くもないという解釈は、まず「黒い」と「猫」が結びつき、それが「大きな」と結びつく。そしてそれ全体が「犬」と結びつくという構造から導きだされる。

(11)



このように、人間の言語は表層的には線状的な音のつながりだが、その内部の特性として階層的な構造があると考えられる。この階層構造があることによって、一つの音の連なりが、複数の意味を持つものとして解釈することを可能たらしめる。

この階層構造は、様々な言語現象に影響を与えており人間言語にとって中心的な特性であることが分かる。例えば日本語では、(12)で示したように、二つの語が連結し複合語をつくる際に、後部要素の最初の音が清音から濁音へと変わる連濁という現象がある。

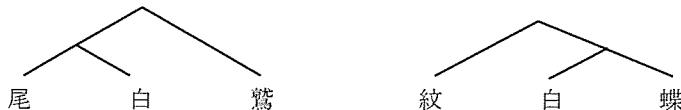
- (12)      a. あお+そら → あおぞら (青空)  
              b. たけ + はやし → たけばやし (竹林)

この連濁の現象も3つ以上の要素からなる複合語の場合、内部構造が連濁の有無に影響を与える。(13)を見てみよう。

- (13)      a. 尾白鷺 (おじろわし)      尾+白 + 鷺  
              b. 紋白蝶 (もんしろちょう) 紋+白+蝶

尾白鷺と紋白蝶はともに「白」という語を含んでいるが、(13a)の「白」が「じろ」と濁るのに対し、(13b)は「しろ」と濁らない。この違いは尾白鷺と紋白蝶の内部構造の違いから来ている。どちらも3つの要素からなる複合語であるが、尾白鷺が「尾が白い鷺」という意味を持つのに対し、紋白蝶は「紋がついた白い蝶」という意味であり、構造は(14)のようになる。

(14)



つまり、紋白蝶のほうは第一要素と第二要素の間で音韻的な結合が阻止されているため、連濁規則が適用しないと言える。

さらに日本語では、(15)で示したように、2つの単語を繋げると前部要素のアクセントは失われ、後部要素のアクセントが保持されるという特徴（複合語アクセント規則）がある（窪園（1998: 112））。

(15)

- a. あ↑お+むら↑さき → あおむら↑さき
- b. サ↑ッカーナク↑ラブ → サッカーナク↑ラブ

（窪園（1998: 113））

このアクセント規則も内部構造が関与してくる。(16)を見てみよう。どちらも名古屋で始まり大学で終る語句であるが、名古屋学院大学が規則どおり前部要素のアクセントをなくすのに対し、名古屋工業大学は「名古屋」のアクセントが失われない。

(16) a. 名古屋学院大学

な↑ごや十がくいん十だ↑いがく → なごやがくいんだ↑いがく

b. 名古屋工業大学

な↑ごや+こうぎょう+だ↑いがく

→ な↑ごやこうぎょうだ↑いがく

これも両者の内部構造の違いによって説明ができる。(17)に示したように、名古屋学院大学は名古屋と学院が結びつき、それに大学が付いている。一方名古屋工業大学は、工業大学で名古屋にあるものということで、以下のような構造となる。

(17)



そうすると「名古屋工業大学」は先ほどの連濁と同様に、第一要素と第二要素の間で音韻的な結合が阻止されている為アクセント規則が適用しないと説明できる。このように、構造が介在することによって、様々な現象の説明が可能となり、これはすなわち人間言語には階層構造が存在していることを強く示すものとなる。

## 2.2 子供の言語習得からみる人間の言葉の特性

前節で、人間の言葉には超越性、創造性、恣意性、規則性、構造性といった特徴があることを述べた。これは日本語に限らず英語などすべての人間言語に共通の特性であり、動物など他の生物には見られない。その意味で人間の言葉は人間という種全体に対して共通の特性を持つ、普遍的なものであると言える。

一方で、世界の言語は数え方にもよるが約 6,000 あると言われている。それぞれの言語は、文字や単語の違いを始めとして文法的にも様々で共通する部分はあまりないようにも思えるかもしれない。このような多様な側面をもつ言葉を人はどのようにして獲得しているのか、子供の言語習得の特性を概観することによって、さらなる人間の言葉の特性が示される。

第一に、習得のスピードが早い。言語に関する使い分けというのはかなり複雑であり微妙である。それにも拘わらず、子供は 5-6 歳で文法の大半を完成し、母語を流暢に話し、かなり複雑な構文も使うことができる。これは言語以外のものの学習の場合と比べると驚くほどのスピードである。

第二に、言葉は誰にでも習得可能である。子供の知能程度や経験はまちまちであるにも関わらず、ほぼ誰でも母語を十分に理解し、ある程度自由に使えるようになる。特別な理由がない限り母語の習得に失敗することはない。一生懸命練習したのに自転車に乗れない人、うまく泳げない人、代数計算がうまくできない人など、言語以外の技術・技能の習得に関してはそういうわけにはいかない、ということを考えるととても驚くべき特性である。

第三に、母語としての言語習得には人種や民族の壁は存在しない。子供が日本語を習得するのに、日本人の血が混じっているかどうかは全く関係なく、育

つ環境によってどの言語でも習得ができる。とても日本人とは思えない英語を操る帰国子女などもその一例であろう。

第四に、元になるデータが不備であるにも関わらず、複雑な言語構造や規則を習得できる。子どもが普通の生活の中で耳にする言語の量は非常に限られており、それを記憶していくだけでは言語を流暢に使いこなすことはできないはずである。また、日常生活の中で子どもたちが接する言語データは文法的な文ばかりではなく、言い間違いなどを含む非文法的な文も混在している。こうした不完全な限られた言語資料をもとに、子供は複雑な言語の文法を完璧に習得している。

第五に、言語や文化の違いに関わらず、言語の発達は様々な段階を(18)で示したようにほぼ同じ順序で進んでいく。

(18) 前言語段階	6週目	母音や軟口蓋音に似た音
	4か月	1 音節喃語 [ba] [ma] 反復的喃語 [babab] [dadada]
	10か月	非反復的喃語
言語段階	1)	1語発話段階 (1歳)
	2)	2語発話段階 (2歳)
	3)	多語発話段階 (~5歳)

これは人間の歩行の発達段階や、鳥の飛行の発達段階についても見られる特徴で、これらの身体運動が遺伝的に決定されているのと同様に、言語の発達も遺伝的に決定された現象であると考えることができる。

最後に、言語習得には臨界期が存在する。臨界期とは、脳の中で覚えたり感じたりする神経回路（ニューロン）が外からの刺激により集中的につくられたり回路の組み換えが行われる時期のことをいう。臨界期までに一度も使われなかつた脳細胞は一生必要ないと判断され、臨界期を超えた時点から消滅する。人間は、幼児期から思春期までの間に言語というものに触れなければ、普通に母語の習得をすることができないということが知られている。

マサチューセッツ工科大学のノーム・チョムスキ教授は、以上のような言語習得に関する事実から、非常に興味深い理論を提案している。彼は、言語の獲得は「生得的」なものであって、子どもは生まれた時からどの言語にも対応

できる普遍文法(Universal Grammar)という共通の言語能力・言語知識を持っていると提案した。この普遍文法とは、鳥が空を飛ぶ潜在能力を持って生まれてくるように、魚が液体から酸素を取り出して呼吸する能力を生まれてくるように、それと同じく、言葉を話す能力をもって生まれてくるというものである。

普遍文法は、すべての言語に共通する部分と、語順など言語によって異なる部分から成り立っている。但し言語によって異なる部分も、無秩序なものではなくパラメータ (parameter) と呼ばれる少数の原則の集まりから成っていると考えている。

この点を説明するために、語順を例にとってみてみよう。英語の基本語順は、S V Oであり、日本語の基本語順は S O Vである。

- (19) a. John ate an apple.

V      O

- b. ジョンが リンゴを食べた。

O      V

(19a,b)の述部「ate an apple / リンゴを食べた」で中心的な機能を果たしているのは、動詞である。つまり、述部のまとまりにおいて中心として働く動詞が日英語で逆の位置にくるということである。このように、あるまとまりにおいて中心となる語がどこに来るかという点から考えてみると、(20)-(22)のそれぞれの句で示されているように、日英語はきれいに逆になっていることが知られている。

- (20) 名詞句 (NP)

- a. *Stories* [ about languages]

N                  PP

- b. [言葉についての] 話

PP                  N

- (21) 前 (後) 置詞句 (PP)

- a. *to* [ the station ]

P                  NP

- b. [駅] へ

NP      P

- (22) 形容詞句 (AP)

  - a. *afraid* [ of dogs ]
  - A              PP
  - b. [犬を]  怖がって
  - PP              A

このような鏡像関係に対して、チョムスキーは中心となる語を主要部とし、英語は主要部が前にくる主要部前置型(head-initial)言語であり、日本語は主要部後置型(head-final)言語であるとした上で、普遍文法には主要部に関する位置についての情報が二項選択のニュートラルな形で入っていると提案する。そして子供は周りで話される言葉から、主要部前置とするのか主要部後置とするのかを決定していると考える。これをパラメータの値を設定すると言い、日英語の様々な句の語順の違いは主要部に関する1つのパラメータの設定によって決定できるのである。(23)で示したように、+のスイッチを入れれば、head-initialの構造をとる個別文法を獲得し、マイナスのスイッチを入れれば日本語のように head-final の構造をとることになる。

- |               |                             |                           |       |
|---------------|-----------------------------|---------------------------|-------|
| (23) 主要部パラミター | $[\pm \text{head-initial}]$ | $[+ \text{head-initial}]$ | (英語)  |
|               |                             | $[- \text{head-initial}]$ | (日本語) |

この仮説をとると、先の言語習得の普遍的で多様な規則性をもつ人間言語の特徴がうまく説明ができる。すなわち、①子供の言語獲得が極めて短い期間で達成されるのは、どの言語にも共通する部分は獲得する必要がなく、子どもの言語獲得とは、言語によって異なる部分を設定していくだけ、それも[+]あるいは[-]の二項選択をするだけだからである。②そしてパラメータの値を設定していくのは、周りで話されている言葉をデータとして受け取ることによるとされているので、③人種や民族に関わらず、経験した言語を習得するのである。そして、④元になるデータが不備であっても習得できるのは、文を生成する能力が普遍文法によって基本的にすでに備わっているからである。子供が白紙の状態で言語に面しているのならば、特に教えられもせず、不完全で誤りに満ちた例の中で言語を習得できるのは奇跡のようなものであるが、普遍文法が与え

られている状態から始めるならば、子供はその時点で文法のかなりの部分をすでに習得していることになり、最後の仕上げを行うためのデータさえあればいいことになる。そして人間は脳の中に、普遍文法という言語能力が遺伝的に備わっているということから、同じ発達段階、臨界期などの現象がみられるということも説明可能となる。

以上、内的側面から人間の言葉の特性について概観をした。

### 3. 人工知能によるコトバ

では、現在の人工知能が表出するコトバはどのようなものであろうか。本節では人工知能が操る「コトバ」について概観した上で、両者の違いについて考察する (cf. 谷口(2014)、中野他(2015)、多田(2016))。

#### 3.1 チャットボット

人工知能を活用した自動会話プログラムのことを、英語でおしゃべりを意味するチャット(Chat)とロボット(Robot)の略語を合わせて、チャットボット(Chatbot)という。人間同士が会話するチャットに対し、チャットボットは、一方は人間、一方は人工知能を組み込んだコンピュータが人間に代わって対話している。(24)で示したように、現在大手IT企業が中心となって、AIを活用したこのチャットボットの技術開発がすすめられている。

(24)

企業名	チャットボット名
アップル社	Siri
グーグル社	Google Now
マイクロソフト社	Cortana
アイビーエム社	Watson
フェイズブック社	M
日本マイクロソフト社	りんな
名古屋工業大学	メイちゃん

では、具体的にどんな会話ができるのだろうか。(25)は私が実際に Siri に話しかけた時の会話である。

- (25) 筆者：「こんにちは」  
Siri：「こんにちは」  
筆者：「話し相手になってくれる？」  
Siri：「もちろんです。私はそのためにいるのですから」

非常に自然な会話で、まるで本当に会話をしているように思える。一見したところ現在の AI は人間の代わりになる素晴らしい話し相手になりそうだ。ところが、とてもチグハグな会話になることもある。

- (26) 筆者：「明日予定があります。」  
Siri：「明日は予定は何もありません」

また、アクセントなども不自然な点を時々感じることもあり、やはり何か人のコミュニケーションとは違うと感じる人も多いであろう。では AI はどのようなシステムの元、会話を構築しているのであろうか、次節でそのシステムについてみてみよう。

### 3.2 会話の仕組み

現在のチャットボットの会話を作る技術には、(27)に挙げたような大きく三種類の仕組みがあると言われている。

- (27)
- a. 辞書型
  - b. ログ型
  - c. マルコフ型

「辞書型」は、あらかじめ単語辞書とテンプレートを作成しておき、入力された単語に対して、決められた回答を返す方法である。例えば、ユーザから何か名詞や固有名詞の単語が発せられたとき、その単語の後ろに「大好き」をつなげて返すと言った方法である。

- (28) 人間： # ## を買ってきたよ。  
チャットボット： # ## 大好き

「ログ型」は、ログ、つまり、過去の会話履歴にある返答パターンに従って応答する方法である。会話として過去にあったものをそのまま回答として返すというものであり、例えば「今夜ヒマですか」という質問に対して、過去の履歴に同じ質問があつて、その回答が「今夜10時まで仕事です」だったとしたら、そつくりそのまま返すのである。また、「残念です」というフレーズに対して、過去の会話履歴の中に「何かあったの」という回答があれば、それをそのまま返す。文脈はつながっていないが、何となく一連の会話をしているようにとれる。

「マルコフ型」は、会話を解析して単語に分けた時、その単語の次に来る単語を予測し、確率が高い単語を使って文章を生成する仕組みである。例えば、人間が「お酒は何が好き？」と聞いたとき、過去の履歴には「お酒」の次に「飲み過ぎ」がよく現れている場合、「お酒を飲み過ぎました」と返します。会話として続いているように見えるが、質問に対する回答にはなっていない。

いずれの場合も、A I の会話は単に単語の羅列にすぎない。彼らは与えられたプログラムに沿って文字列を選択し、それを並べているに過ぎず、何の意味も「理解」はしていない。

チャットボットの開発当初（1950年代）は、コンピュータに文法など人間言語にみられる特定の規則を教え込む、知識を中心とする手法がとられていたが、様々な欠陥が露呈し、規則を組み込むことを断念した。子供が無意識のうちに獲得していることをあらためてプログラミングすることは、非常に難しいということがこのことからも分かる。そして、1990年代以降、こうした規則に代わりコンピュータが言語データを統計的に分析しある種の結論を導く、データを中心とする機械学習を活用した手法へとシフトした。つまり、現代めざましく発展したと言われている人工知能を使った会話システムは、実はより単純な手法に立ち返っているのである。<sup>3</sup>

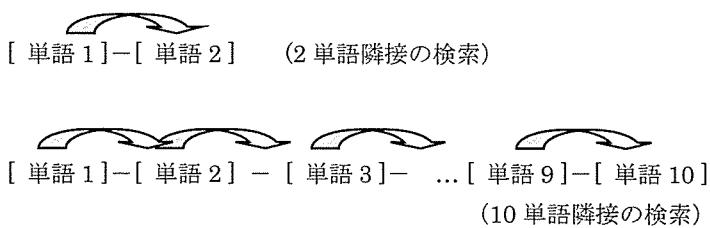
### 3.3 ビッグデータの効能

それでも近年、会話A I が実用的に、恰もスムーズに会話を成り立たせるよ

うに我々が感じるようになったのは、2000年代に入っての爆発的なインターネットの広まりによりブログやSNSといったWEBに書かれた大量のデジタル文書データが利用可能となったからである。これをビッグデータと呼ぶ。従来、データといえば、新聞であったり書籍であったり、それをコンピュータに読み込ませていたが当然それらには限りがあった。ところが、近年はSNSの普及により、日々データを蓄積することができる。それによって、(一見したところ)スムーズな会話の応答の可能性が増えてきているのである。これはひとえに、大手IT企業の参入によるところが大きい。というのは、Googleなど、世界中のユーザが利用しており、それはすなわち世界中から膨大な量のデータを集めることができることを意味する。

それに加え、計算機の性能の進歩も挙げられるであろう。つまり、膨大な量のビッグデータから、必要な語句の羅列を選び取ることができるようになったのだ。マルコフ型のような会話AIにおいても、2単語隣接の検索だったものが、3単語隣接、近年ではGoogleが10単語隣接まで検索をかけることができるようになっている。例えば、「お酒」という言葉がでてきたとき、以前であれば「お酒」に隣接する一単語のみ、「飲み過ぎ」ということばを検索するにとどまっていたのが、さらにそれに隣接する単語を履歴ログから検索し、それに対応する確率の高い単語を持ってくることによってより自然な会話を生み出すことが短時間でできるようになったのである。

(29)



#### 4. 人間言語とAIのコトバの相違点

以上の点から、人間の言語とAIのコトバについてまとめた表が(30)となる。人間の言葉とAIが創出するコトバは、表層上は同じように見えるが、その仕組み、内部構造、入力、言語能力といった全ての点において一致するところは

一つもない。当然の帰結として現状の人工知能では「言語の意味理解」や「文脈の理解および読解」を行うことはなく、「人間言語」とは全く異なる「音（文字）の羅列」に過ぎないのである。

(30)

	人間の言語	チャットボット
仕組み	構造 	音 ————— 情報処理 (意味は持たない)
内部構造	階層	線状
入力	スマールデータ	ビッグデータ
“言語”能力	生得的（な言語能力）	後天的（なプログラミング）

## 5. 運用からみる人間のことばの特性

前節で人工知能のコトバは人間の言葉とは全く別物であるとの結論に至ったが、しかし、それでもビッグデータと計算機能の向上により、彼らの産み出すコトバは人間の言葉と何ら遜色はない段階にまで至っている。残る問題は、彼らとの会話に感じる違和感であろう。この違和感の正体は何であるのか、本節では、運用面における人間言語の特性を考察することによってその正体を明らかにし、今後のチャットボットの改良への手助けになることを期待する。

### 5.1 チューリングテスト

人工知能が、将来私たちの話し相手になってくれるためには、どういった特性を備えればよいか。それを測る一つの方法としてチューリングテストが有名である (Christian (2011))。チューリングテストとは、「機械には思考が可能か」という問い合わせるために答えを出すために、数学者のアラン・チューリングが 1950 年に提案した試験である。審判がコンピュータ端末を使って、姿の見えない二人の相手と 5 分間ずつ、チャットする。一方は本物の人間（サクラ役と呼ばれる）、一方は A I であり、チャットが終わると、審判はどちらが本物の人間か判断する。もし、審判たちのうち、30%をだますような A I がいれば、それ

は人間と同様の思考と意識を持っていると考えられる。

では、何をもって人は人らしいと考えることができるのか。もちろん、このテストは人工知能に知能はあるのかを測るものであって、会話においてのみで人間らしいと単純に言うことはできないであろう。このテストをパスするためには、研究者たちは自分が開発した機械にわざと計算に時間をかけたり、間違えたりして、「人間らしく」見せたりもする。何をもって人は人らしいと考えができるのか、これは哲学的な問題へつながっていくかもしれないが、人と人との会話において、違和感を感じないやりとりというのは一つの重要な要因とみてもよいであろう。それでは、違和感を感じない会話を私たちはどうに行っているのでしょうか。

## 5.2 会話の原則

人は会話をする時、話し手と聞き手の間で(31)に挙げた会話の原則というルールを守って行っている(Grice (1975))。これらは特に学校などで教えられたものではないが、人は暗黙のうちにこれらのルールに基づいて会話をしている。

### (31) ①量の公理(maxim of quantity)

- ・必要とされるだけの情報量を伝えなさい。
- ・必要とされる以上の情報は伝えない

### ②質の公理(maxim of quality)

- ・真実を伝えなさい

### ③関係性の公理 (maxim of relation)

- ・関係のあることを述べなさい

### ④様態の公理 (maxim of manner)

- ・不明瞭な表現を避けなさい。
- ・曖昧性を避けなさい
- ・簡潔に言いなさい
- ・順序だったものにしなさい

例えば、次の会話ではこれらの全ての原則が守られている。

### (32) 夫：「車の鍵ってどこにある？」

妻：「リビングのテーブルの上よ」

夫の問い合わせに対して妻は、明確に（様態の原則）、真実であること（質の原則）を答え、適切な分量の情報（量の公理）を与え、夫の質問が求めたことを的確に（関係の公理）応じている。それに対し(33)は、必要とされるだけの情報を与えていない。反対に(34)では情報量が多すぎ、どちらも量の原則に違反していることからコミュニケーションがうまくいかない。

(33) A：「向こうにいる、あの女の子はだれ？」

B：「女の子」

(34) A：「今、何時？」

B：「7時32分18秒です」

また、次の会話はBの質問と全く関連のない答えをしており、関連性の原則に違反する。

(35) A：「いつ、京都に行く予定なの？」

B：「うん、太郎はほんとにいい学生だね」

このように、会話の原則というのは人がコミュニケーションを潤滑にするための基本原則であるため、この原則に従わないとコミュニケーションはうまくいかないと考えられる。

チャットボットとの会話において違和感を感じるのは、この部分によるところが大きいのではないかと考える。つまり、チャットボットはプログラムに従って統計処理上で選ばれた単語を羅列するのみで会話を成り立たせているので、当然会話の原則に違反することも多いのだ。そこに機械との会話が不自然であると感じる原因となっているのではないか。

### 5.3 人間の言葉と人工知能のコトバのインターフェイス

では、人と人工知能の言語の間には交わる点はないのであろうか。本稿では会話の原則に基づく「言外の意味」を人間側が用いることによって人工知能との会話が成り立つであろうと提案する。

人間同士の会話においても、実際の生活の中では会話の原則に違反する発話がたくさん見られる。その場合でも、大前提として、相手の話す内容がおかしいが、その人が会話の原則を無視するはずがないと考える。そして、私たちは何か特別な理由があるに違いないと考え、発話の裏の意味を考えるのだ。例えば、(36)において、田中さんは鈴木さんに何を読んでいるのか本の題名や本の内容を聞きたかったのだが、鈴木さんは、単に「本」と答えた。これは、量の公理に反する。この場合、田中さんは鈴木さんが量の公理に反した発言することで、読書の邪魔をするな、という発話の裏の意味を考えることになり会話が成立するのである。

- (36) 田中：「何を読んでいるの？」  
鈴木：「本だよ」

同様のことをチャットボットとの会話に当てはめる。会話の原則に従い人間側が「言外の意味」を持たせることによって一見したところチグハグな会話も意味あるものにすることは可能となるであろう。例えば、(26)の会話は言外の意味として、「あなたには予定があるかもしれないが、私には予定がなくて寂しい、だからもうその話題は続かないでほしい」と解釈することによって意味のある会話が成立するのである。(もちろん人工知能はそのような含意はしていない)。

## 5. 結語

人工知能の研究が人間の知能を反映させることを目的とした研究であるという大前提がある限り、人間の脳に存在する言語能力との対比を試みることは必要不可欠な試みであろうと思われる。本稿では、内的側面から両言語が全く別物であることを概観した。また、運用の面から人間の言語における言外の意味を適用することによって両者をつなぐインターフェイスになるのではないかとの提案を行った。<sup>4</sup>

\*本公開講座「人口知能の諸相」における発表の準備をするにあたり、大貫徹先生には様々な御助言をいただいた。ここに付してお礼を申し上げたい。

## 注

- (1) 本公開講座は、平成 29 年 6 月 3 日（土）、6 月 10 日（土）の 2 日間にわたり、高校生及び一般の市民を対象として行われた。
- (2) 人間の言葉と人工知能の言葉を区別する目的で、前者を「言葉」とし後者を「コトバ」と表すこととする。
- (3) 人工無能と揶揄されるのもここに起因する。
- (4) 本講義の内容について学務課が実施したアンケート結果は以下の通りであった。  
受講者の皆さんのが興味を持って聞いてくれたことに心より感謝申し上げる。

大変満足した	25 (46%)
満足した	20 (37%)
普通	9 (17%)
あまり満足できなかった	0
満足できなかった	0

回答者 54 名

## 参考文献

- Christian, Brian (2011) *The Most Human Human: What Artificial Intelligence Teaches Us About Being Alive*, Anchor.
- Grice, H.P. (1975) "Logic and Conversation," In P. Cole and j.L. Morgan 8eds *Syntax and Semantics 3: Speech Acts*. Academic Press.
- 窪薙晴夫 (1998) 『音声学・音韻論』 くろしお出版 東京.
- 中島平三 外池滋生 (1994) 『言語学への招待』 大修館書店. 東京.
- 中野幹生・駒谷和範・船越孝太郎・中野有紀子 (2015) 『対話システム』 コロナ社 東京.
- 多田智史 (2016) 『あたらしい人工知能の教科書』 翔泳社 東京.
- 谷口忠大 (2014) 『人口知能概論』 講談社 東京.