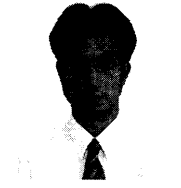


特集 「編集委員今年の抱負 2013」

人間のようなターンテイキング

駒谷 和範 名古屋大学大学院工学研究科



人間のように対話を行うシステムを考えるとき、①何を、②いつ、③どのように、システムに発話させるのかを考える必要があります。このうち、①は対話システム一般の、古くからの問題であり、③の解決には表現力豊かな音声合成が鍵となります。ここでは、②の「いつ話し始めるか」について考えてみたいと思います。

「システムがいつ話し始めるか」は、ターンテイキング (turn taking) に関する問題として知られています。ターンテイキングとは話者交替のことで、「二者が交互に話す」ことを指します。人間同士の対話の多くでは、ターンテイキングは、無意識のうちに、円滑に行われています。しかしそれが破綻したとき、その存在に気がきます。例えば、両者が同時に話し始めたり、まだ話し続けたいのに相手に割り込まれたり、といった場合です。またテキストによるコミュニケーションでも、これに類する現象はあります。例えば「メールの返信がすれ違う」のは、ターンテイキングの不具合に相当します。この場合、両者は対話の内容が伝わっているかどうか (つまり共有信念に齟齬が生じていないかどうか) を不安に思い、送信時間を確認したうえで、「メールがすれ違ってすみません」と問題を共有して、修復を図ります。

音声対話システムにおいて、人間のようなターンテイキングを行うといっても、それは一様ではありません。まずよく知られていることとして、文化による違いがあり、この典型例として、あいづちの頻度があげられます。私は2008年に一年間米国に滞在する機会を幸運にも得たのですが、その当初に戸惑ったことの 하나가これでした。相手があいづちも打たずに (単語が出てこずにこちらが困っているのに)、黙ってこちらの目をじっと見ている、という状況です。自分の英語の拙さから、そもそも話が全く通じていないのだろうかとかよく心配になりました。しかしこれは、「日本人はアメリカ人の3倍のあいづちを打つ」という研究結果があるように、文化の差によるようです。またあいづちを打つ箇所も、日本語だとあらゆる文節境界であいづちを打てるが、英語だと基本的に文末のポーズでしかあいづちを打たない。つまり、「相手が話している途中に口を挟むのは失礼」という文化のようです。アメリカでのほかの例として、アメリカ人の知人が、電話で病院の予約をするのを聞いていると、電話が相手に通じた後、すべての必要情報を、一気にほぼ話者交替することなく伝えていました。日本だと、要件を区切り、話者交替をしながら少しずつ伝えるでしょう。これは、「自分と相手の立場は違うので、まず自分の要求

をきちんと伝えないと始まらない」という文化と、「自分が相手と同じ背景知識をもっていると仮定でき、共同で目的を達成する」という文化の違いなのかもしれません。

次に、最近、大学で日本人の学生と話しているときにも、ターンテイキングに違和感を覚えることがしばしばあります。こちらが話しているにもかかわらず、彼らが話し始めるという現象です。しかも、その内容は、そこでどうしても伝えたいことがあったというよりは、単なる復唱 (感想?) だったり、特に重要でない内容だったりします。このような発話行動をとるのは、均質化された範囲内 (友人や家族) でのコミュニケーションに慣れ、世代や立場が異なる人と話す経験が少ないことに起因するのでしょうか。つまり、どういうタイミングで話し始めても、聞き入れて理解してもらえぬ環境に長くいたのかなど考えたりします。そんな彼らが社会人との「未知との遭遇」を果たすシューカツ (就職活動) では、もし誰にも指摘されなければ、「〇〇させていただく」を乱用し、許容され難いターンテイキングをしながら、面接に臨むのかもしれませんが。

言語表現としての敬語は世に広く認識され、しばしばその乱れが話題になります。この敬語と同様に、ターンテイキングにおいても、親しい人と話す場合と、オフィシャルな会話を行う場合で、本来使い分けがあるはずですが、友人と話するときには、相手の発話末に重ねるように同意したり、相手の途中までの発話につなげて発話を完結させる「共話」を行ったりすることで、より会話を楽しめるでしょう。一方で、面接でそのようなターンテイキングを行うと、あまり考えずに話す人だ、という印象を与えかねません。

音声対話システムでも、基本的なターンテイキングが実現できた後には、タスクや状況 (話し相手を含む) に応じて、ターンテイキングの戦略を使い分ける必要があるでしょう。つまり、雑談的な対話をする場合と、タスクを確実に遂行すべき場合では、我々が敬語を使い分けするように、ターンテイキングの戦略も異なるはずです。

しかしそもそも、システムが自分の発話の途中で口を挟んでくるのを、私は許容できるのでしょうか。よほど気の利いたことを言うのなら、それにより楽しい対話が行えそうですが、音声認識の問題が解決されたとしても、しばらくは、邪魔だ、なれなれしい、と感じてしまいそうです。発話内容とともに、ターンテイキング戦略を自在に使い分けられる音声対話ロボットの実現、これを今年の抱負とするには時期尚早のようです。