

対話音声聴取時における間投詞の働き

3 J - 2

伊藤敏彦 峯松 信明 中川聖一

豊橋技術科学大学 情報工学系

1 はじめに

本研究では協調的な問題解決の対話音声中に存在する間投詞に着目し「発話中の間投詞は聞き手に対してどのような働きを持つか」、「協調的な応答文生成において間投詞は有効又必要なのか」という観点から、知覚実験による検討を行なった。実験は対話音声より、間投詞部分を1)抜き出して切り詰めた音声試料、2)同一時間長の無音置換を施した音声試料、3)異なる箇所が発声された同一種類の間投詞、4)異なる種類の間投詞と置換した音声試料、5)2)の無音区間の長さを様々に変化させた音声試料、6)間投詞の直前に位置する無音区間を様々に変化させた音声試料、を各々用意し被験者に提示した。1)~4)までの音声試料に対しては自然である(違和感を全く感じない)との反応を示した。5)、6)に対しては「長い無音区間が不自然に感じる」との反応が幾らかあった。以下、本実験の目的・計画・結果・考察について述べる。なお、本稿で言う無音置換とはバックグラウンドノイズとの置換を意味する。

2 研究の目的

現在の音声対話システムの研究対象は自然な発話に関するものが多い。その場合「書き言葉」には見られない、「話し言葉」特有の言語現象(言い直し、間投詞など)を扱う必要がある[1]。最近では、対話音声中の各種の現象を分析した研究が行なわれている[2]が、分析結果を実験者の主観によって分類する作業が一般的であり、実際に音声を聴取している被験者の動向/判断に基づく研究例は少ない。

2.1 間投詞に対する予備的考察

従来の音声認識等では間投詞は冗長語として忌み嫌われ、特別な検出機構を設けることによって[3]最終的には不要語として扱われてきた。ここで、人間が音声を生成・知覚する際の間投詞の役割について考察を行なう。生成の立場から考えると、間投詞の生成理由としては「概念→言語」の符号化に何らかの支障を来した場合、全く発話行為がなければ無音状態に呈することになる。また表層的な表現形式で即座の返答を控える場合もある。この場合も同様に発話行為がなければ、無音状態に陥ることになる。このように「通常の発話行為を継続できない/したくない」心理状況を相手に提示する術として無音を埋めるべく間投詞を使用していると考えられる。もちろん長い無音を呈することでも同様な意図伝達は可能であるが間投詞の方がより効果的に働くと思われる(filled pauseといわれる所以)。

一方、間投詞の存在により、聞き手の理解が助けられる(発話内容を理解するための時間が与えられる)こ

とも考えられる。しかし、聞き手に時間を与えられることが間投詞の主な働きであるとするならば、単なる無音と置き換えても何ら支障を来さないことになるが、果たしてそうであろうか。

さらに、間投詞の部分が無音に置き換えても何ら支障がないならば、どれくらいの長さの無音時間が続く人間は違和感を感じるのか。また違和感を感じる無音時間は絶対的なものなのか、文脈、対話速度、発話権の移動などによって変化する相対的なものなのか。

以上のような考察から、以前行なった講演調音声に対する知覚実験[4]と同様に以下のような項目が実験の検討項目としてあげられる。

間投詞の必要性(応答生成の立場から) 間投詞を抜き出して切り詰めた音声試料は聞き取りやすいのか/難しいのか。

間投詞の種類別の必要性(応答生成の立場から) 間投詞が発話を自然にするなどの理由から文生成において必要であるなら、いろいろな種類の間投詞を使い分ける必要はあるのか。

間投詞と無音の等価性 間投詞は無音と置換しても何ら支障はないのか。その等価性は対話速度などに依存しないのか。等価性があるならば違和感を感じない無音の長さはどの程度なのか。

間投詞の有無などによる無音の許容時間の変化 違和感を感じない無音時間は、無音後の第一発声の間投詞か否かによって変化するのか。またその無音時間は対話速度によっても変化するのか。

3 聴取実験

3.1 実験の目的

本実験では聞き手における「間投詞の必要性」、「間投詞と無音の等価性」、「間投詞間の相互置換性」に着目する。もちろん、間投詞の必要性が低いことが示されても、間投詞が発話文をより自然したり、発話文生成までにかかる時間(一般に無音状態となる)の印象を少しでも良くしてくれるならそれは非常に利用価値があると考えられる。そのため「間投詞による自然性」、「間投詞の有無などによる許される無音時間の変化」にも同様に着目する。

なお先行実験[4]より、無音区間に対する聞き手の反応が発話速度に依存することが示されているのでここでは発話速度も考慮した上で実験を計画した。

3.2 音声試料

協調的対話の音声試料として文部省重点領域研究「音声・言語・概念の統合的処理による対話の理解と生成に関する研究」の対話音声コーパスから「地図課題(主観的に対話理解難易度は高く、対話速度は速い)」と「地理・旅行案内(理解難易度・低、対話速度・遅)」の対話を用いた[5]。「地理・旅行案内」はスキー旅行の予

約を行なうタスクであり、「地図課題」は情報提供者と情報追随者がもつ各々の地図の足りない情報をお互いに交換しながら情報提供者の地図にかかっている順序を情報追随者に教えるタスクである。「地図課題」の音声データは発話スピードが速く、対話内容も少し複雑で理解しにくい。逆に「地理・旅行案内」の方は対話スピードも遅く、その内容も理解しやすい。

実験の結果は音声の性質やタスクの種類に依存した結果であるとみなすべきであるが、間投詞の働きを見るには十分であると思われる。

このそれぞれの音声データにおける間投詞の位置とその前後の無音区間の位置を視察/聴取によって決定し記録した。これらの位置データを利用し以下のような処理を施した。

- (1) 間投詞を切り出して切り詰める。
- (2) 間投詞を同一時間長の無音(バックグラウンドノイズ、以下同様)と置換する。
- (3) 間投詞を異発話の間投詞と時間長が等しい無音と置換する。
- (4) 別の場所の同一種類の間投詞と置換する。
- (5) 異なる種類の間投詞と置換する。
- (6) 間投詞をランダムな無音区間と置換する。
- (7) 発話権の移行に伴う無音区間と語頭の間投詞をランダムな無音長と置換する。
- (8) 発話権の移行に伴う無音区間をランダムな無音長と置換する。

無音に置換する場合は、背景雑音を考慮し実際の対話中の無音部分の波形を用いて置換した。他の間投詞と置換する場合も、ピッチの連続性を考慮して置換対象の間投詞を決定した。異なる間投詞と置換する場合にはその間投詞の意味的素性等は考慮していない。

「間投詞の必要性や無音との等価性」を調査するために、実験に用いる各対話を5等分し各々について(1)・(2)・(4)・(5)・(7)の処理を施したものを用意する。この中から処理が重ならないようにランダムに選択し全対話を構築する。このようにして作られた試料(音声試料1)を4種類作成した(タスク毎に2種類)。「間投詞の自然性や許容される無音時間の変化」を調べるために、上述8種の処理から2つをランダムに選択し、約10ターン程の長さでランダムに抜き出した音声区間に対して適用する(音声試料2,6サブ対話*2種類*2タスク)。

3.3 実験方法

まず、絶対的な評価として前節で示した音声試料1の内2対話(1対話*2タスク)を本大学の学生(3名・本実験+4名・予備実験)にスピーカを通して提示した。被験者には「提示音声には人工的な操作が加えられているので、不自然さを感じる場所があったら知らせて欲しい」という指示を与えた。被験者が知らせてきた不自然さを感じた場所は、実験者が書き起こしデータ上にマークし、音声試料提示終了後にマークされた位置において違和感を感じた理由をインタビューした。その後、簡単なアンケートに記入してもらった。

次に相対的な評価として、音声試料2を1サブ対話(2データ)毎聞かせては、その2種類の音声データを聞いてどちらがより自然かなどのアンケートを記入させた。この場合も先ほどと同様に違和感を感じる場所があればその位置を知らせるように指示した。

3.4 実験の結果

表1に音声試料1中に含まれる間投詞の数を表す。表中の“分子/分母”で示されているのが実験のために用意した2対話のそれぞれの間投詞の数である。

表1: 各音声データ中の間投詞数

	処理(1)	処理(2)	処理(4)	処理(5)	処理(7)
旅行案内	12/12	27/29	33/29	29/33	29/27
地図課題	15/15	13/16	14/21	21/14	16/13

被験者には間投詞部分に関して音声処理を行なったとは明示的には伝えていないため、被験者が違和感を感じた部分の原因は多岐に渡っている(例えば、発話者のアクセント、話し方等)。しかし、音声試料1に関する実験において、処理(1)・(4)・(5)による原因で被験者が違和感を感じることは全くなかった。ただ、処理(2)や処理(7)が原因である不自然に長い無音区間に関しては違和感を感じる事が数回あった。

また音声試料2における比較実験においても、不自然に長い無音区間があるためという原因以外で聞きやすさ、理解しやすさ、自然さなどが大きく変化することはなかった。また印象も被験者によってバラツキが大きく有意な差は見られない。ただし間投詞や無音の数やその長さによって被験者の感じる発話速度を変化させることができることはわかった。さらにどれくらいの無音時間ならば被験者は違和感を感じないかという調査に関しては、状況(音響的・言語的)によって許される無音時間があまりに大きく変化し未だ正確には掴めていない。もう少し細かい調査が必要であると考えられる。

3.5 まとめ

間投詞は、その種類毎に発声者の心理状態を表すとされているが[6]、本聴取実験ではそのような知見は得られず、むしろ無音区間と置換しても違和感を感じないという結果を得た。これは講演調音声の聴取実験の結果[4]と一致する。

参考文献

- [1] 森元他, 日本音響学会誌, vol.50, No.7, pp.558-562(1994)
- [2] 中川他, 日本音響学会誌, vol.51, No.3, pp.202-210(1995)
- [3] A.Kai et al., Proc. ESCA. EUROSPEECH'95, pp2095-2098(1995)
- [4] 峯松他, 音響学会講演集, 3-3-4(1996.3)
- [5] 板橋他, 人工知能学会研究会資料, SIG-SLUD-9701-5(1997.6)
- [6] 田窪, 情報処理学会論文誌, Vol.36, No11, pp.1020-1026(1995)