

演劇界の常識「距離感」の実証への取り組み

Efforts to substantiate the common knowledge “the sense of distance” in the theater field

村上 遥^{1,2}

Haruka Murakami^{1,2}

¹ 国立情報学研究所, ² 東京大学附属病院 22 世紀医療センター

¹National Institute of Informatics, ² 22nd Century Medical and Research Center, The University of Tokyo Hospital

Abstract: English Abstract. 近年、AI 合成音声は自然に聴こえるようになってきたが、人間のような表現豊かな発話は出来ず、現在は AI 音声に感情を付加する研究が活発になされている。しかし、演劇界では感情表現を行う際「人が対話相手との物理的・心理的距離感によって無意識的に発話方法を変えている」という「距離感」と呼ばれる現象の意識が重要視されており、合成音声でも検討する必要があると考える。「距離感」は科学的検証例がないため本研究では実験を行い、実証を試みた。

1. はじめに

1.1 AI 合成音声の現状

本文昨今、Alexa(Amazon)、Google アシスタント(Google)、Siri(Apple)などのスマートスピーカーの普及により、我々が日常的に AI 対話システムと接する機会は増大した。また、音声合成システムの進化によりその読み上げ音声も Deep Learning の登場により以前と比較して自然となり、NHK では”ニュースのヨミ子”と呼ばれる AI アナウンサーがニュースを読む場面があるほど実用性のあるものとなった。現在の技術はナレーションのようなフラットに情報を伝える音声合成では実用にたえるものの、それ以上の話声による情報を伝えるまでには至っていない。そのため、より人間に親和性のある表現をするためにこうした音声合成の研究では近年感情を付加するという試みが多くなされている[1][2][3]。しかし、この方法では人間が定義した感情のみを極地的に表現することしかできない。特に日本人は日頃から明確な喜怒哀楽を表現している時間は多くないため、この方法では発話時間の大半が何も表現をしない時間になってしまう。

1.2 人間による表現の重要要素「距離感」

では、人間に何かしらの意図を持たせて読み上げや発声をさせる時には何が起きているのか。それを

考えることで感情以外の付加すべき情報が分かるはずである。仕事とする声優業に関して筆者が声優及びその講師、演出家の下で学んだところ、「演技、表現には距離感が最も大事で、距離感のない表現は棒読みになる」という話をたびたび聞き、「距離感」という概念が演劇界において常識となっていることを知った。「距離感」とは演者とその対話相手、もしくは視聴者との①物理的距離感覚及び②心理的距離感覚を指す。

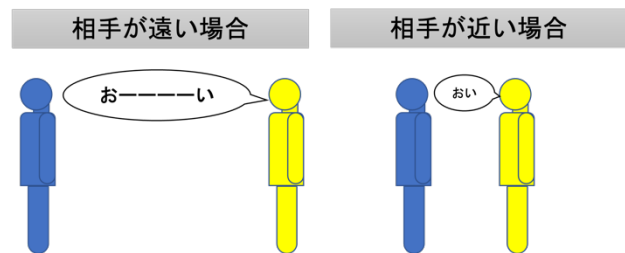


図 1. 物理的距離感の例

対話相手や呼びかける相手が遠い時、声の大きさだけでなく、話すスピードが遅くなる。

①物理的距離感とは発話者とその伝える対象が物理的に離れている距離によって表現を変える感覚のことを指す。これは無意識的に行われており、最も顕著に我々が実感しやすい例としてやまびこに向かって「やっほー」と叫ぶ例が挙げられる。その際、声は通常より大きく、発話は間延びしたり速度が遅くなる。一方、一部の若者で友人に「やっほー」とい

う挨拶をする人がいるが、教室から出た瞬間に会うなど接近している際には「やほ」と言うように発話は短くなる。この発話をやまびこと友人で逆にした場合、「空気が読めない」と感じられる。これが人が無意識で行なっている「物理的距離感」による発話の調整であり、これに失敗すると不自然さを感じたり伝わらない表現となる。

②心理的距離感とは発話者と対話相手の心の距離によって発話者が無意識的および半意識的に発話を調整する感覚のことであり、親しさという心の距離だけでなく、上司と部下、年功序列などの上下関係、親子、師弟、敵対関係、恋人など関係性の種類も絡んだ単純なスカラー距離ではないものを表す。こちらはそうは思っていないのに、妙に馴れ馴れしい人のことを「あの人は距離感がバグっている（おかしい）」と言うことがあるように、我々は心理的距離感に沿った行動をすることが当然だという共通認識があり、一般的に単に距離感と表現するときにはこの心理的距離感を指す。明確な態度として意識的に距離感を表す他に、発話において人によって半意識的に目上の人への声を抑えたり、友人になったばかりの人には「おはよう」と言う挨拶が慣れ親しんで来ると無意識的に「おはよ」など同テキストでもやや雑に聞こえる言い方に変わるなどがこれに当たる。

距離感にはこのように二種類の距離感が存在し、人が表現をして相手に伝える時にはこれら距離感が“合っている”ことが何よりも前提なのだ。また、距離感が互いに合致していることは互いを理解しているというメッセージの一つになると考えられるため、距離感を考慮しない感情表現は親和性が低くなる可能性がある。

1.3 本研究の目的

この「距離感」が表現に与える効果が普遍的なものであるならば、現在の音声合成で感情研究のみを行なっても、受取手の人間にとっては独りよがりな表現となり、伝わらない表現となる。よって、筆者は感情を付加する以前に、距離感というコンテキストを音声合成に付加する必要があると考える。

しかしながら、この共通概念は科学的に検証されたことがなく、書籍等で体系化されたり分析された例もない。そのため、i)聴き手側が距離感を聞き分けられる、ii)発話者が距離感によって発話を変えているという現象についてまずは証明する必要がある。本稿ではこの現象が実際に観察されるものかどうかを検証する第一歩として、物理的距離が離れた際にどの程度角度の違いに対して聴き手が聞き分けられるかを調査することとした。人は音声定位が可能であることが分かっているため、より明確な実験結果

を得やすいと判断し、角度の検証から取り組むこととした。自分に向けて話しかけられた場合に、角度がどの程度であれば認識できるかを調べるのが本実験の目的である。

2. 検証方法

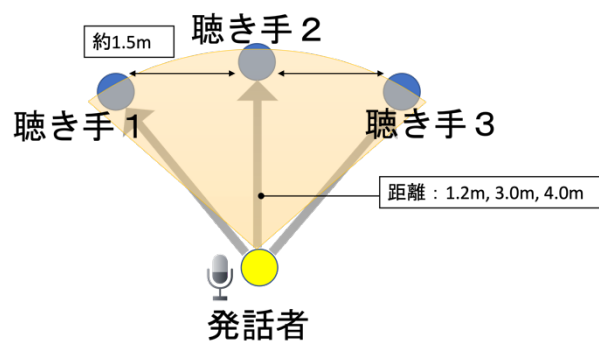


図 2. 被験者の配置

どの程度の距離から話しかけると聞き分けられるのかについて調べるため、被験者を発話者 1 人、聴き手 3 人の 4 名で図 2 のように聞き手同士がやや隣り合った状態で配置した。被験者は 20 代の日本語を母語とする成人男性 3 名と日本語を話せるが母語ではない成人女性 1 名である。聴き手は発話者に背を向け、発話者は実験実施者の指での 1、2、3 の指示に従い、指示された聴き手番号の人に向けて「すみません」と声をかける。聴き手は自分が話しかけられたと思った場合は振り返り、他の人が話しかけられたと思った場合にはその人を指差し、分からない場合には頭の上でバツを作るというジェスチャーをすることで、話しかけられたかどうかの判断表示をもらった。

角度は現実世界で聴き手側が調整することは考えづらいため、発話者が相手から距離がどれくらい離れると音声定位が難しくなるかを調べることとし、発話者が聴き手から離れる形で距離 $l = 1.2, 3.0, 4.0$ [m] と変化させて実験を行った。これは、Sommer によって提唱され[4]、Hall によって分類されたパーソナル・スペース[5]の概念を参考とし、固体距離 (0.45m~1.2m)、社会距離 (1.2m~3.5m)、公共距離 (≤ 3.5 m) を参考にし、公共距離に近づくほど識別精度が下がると仮説を立てて設定した。密接距離 (0m~0.45m) は角度が大きすぎるにより明確に聞き分けられることから割愛した。

3. 結果および考察

結果を表1に示す。

表1. 聴き手によるターゲット当ての結果

距離(m)	各試行回数	1,2,3,0=分らない			
		聴き手1,2,3 ターゲット	聴き手左	聴き手中央	聴き手右
4m	1	2	3	2	0
	2	3	3	2	0
	3	1	0	2	1
	4	2	3	2	0
	5	1	1	2	3
	6	3	0	2	2
	7	1	1	2	3
	8	2	0	2	3
	9	3	2	2	2
3m	1	1	2	2	2
	2	2	0	2	2
	3	3	3	3	3
	4	1	1	1	1
	5	3	3	3	3
	6	2	2	2	2
	7	1	0	1	1
	8	3	3	3	3
	9	2	2	2	2
1.2m	1	1	1	1	1
	2	2	2	3	3
	3	3	3	3	3
	4	1	1	1	1
	5	3	3	2	3
	6	2	2	2	2
	7	1	1	1	1
	8	3	3	3	3
	9	2	2	2	2

結果として、発話者が聴き手に近づいた方が精度は上がった。社会距離の範囲内である4.0mにした途端に聞き分け精度が大きく下がった。この結果は、角度差が大きい方が聞き分けやすいというだけではなく、社会距離から急激に精度が下がったことを考えると、人間の耳が自分にとって普段重要なスペース（距離内）にいる人の声に対する判断力がそうでないスペースの人よりも感度が高い可能性も考えられる。

また、被験者から聞き取った感想として、「左から右もしくは右から左など大きく距離が変わった場合には誰に向けて話しているかわかりやすいが、端から真ん中、真ん中から端など少し変わっただけの時には、真ん中か端のどちらの人が呼びかけられているかが分かりづらい」という声があった。「誰に向けて呼びかけたかの正解を毎回フィードバックした時の方が分かりやすい」という感想もあり、絶対的な角度が呼びかけ対象の判断に重要というよりも、振り幅などの相対値やその環境での聞こえに慣れたかどうかということが、呼びかけられたという認識に影響を与える可能性が示唆された。

4. 今後の取り組み

本稿では「距離感」という現象の実証実験の第一

手として発話者が距離をやや開けて互いに隣り合う複数の聴き手から離れた場合に、聴き手側がどの程度の精度で自分が話しかけられていると認識できるかの検証を行った。今後は物理的距離感の実証実験を被験者実験やその実験での録音音声の分析を通して行なっていく予定である。

謝辞

本研究の一部はJSPS 若手研究 JP21414497 の助成による。

参考文献

- [1] 森山 剛, 森 真也, 小沢 慎治: 韻律の部分空間を用いた感情音声合成, 情報処理学会論文誌, Vol. 50, No. 3, pp. 1181-1191, (2009)
- [2] 加瀬 嵩人, 能勢 隆, 千葉 祐弥, 伊藤 彰則: 発話状態推定に基づく協調的感情音声合成による音声対話システムの評価, 電子情報通信学会論文誌 A, Vol. J99-A, No. 1, pp. 25-35, (2016)
- [3] 松本 剣斗, 原 直, 阿部 匡伸: WaveNetによる言語情報を含まない感情音声合成方式の検討, 研究報告音声言語情報処理 (SLP), Vol. 2019-SLP-127, No. 61, pp. 1-6, (2019)
- [4] Robert Sommer: Studies in Personal Space, Sociometry, American Sociological Association, Vol. 22, No. 3(Sep.), pp.247-260, (1959)
- [5] Edward T. Hall: The hidden dimension, Doubleday, Garden City, N.Y., (1966)