

言語習得はどこまで社会的か？

— 社会学習と CDS 研究のメタ分析から迫る言語習得の姿 —

吉川 正人 (慶應義塾大学)

1. はじめに

言語習得は社会的な学習であり、言語は他社との相互行為という社会的な営みを通して習得されるものだというの
は一つの通説となっていると言える (e.g., Tomasello, 2003). しかし、様々な研究を概観してみると、言語習得のどの
部分が、どの程度「社会的」であるか、という点は、見解の一致をみていないように思える。また、いくつかの研究
成果は、言語習得の社会性に関して互いに矛盾する結果を提示しており、その意味するところを考えると、通説とし
て議論されているものとは異なる「社会性」が見えてくるように思える。

本発表では、「社会学習」として言語習得をとらえる研究 (e.g., Kuhl, Tsao, & Liu, 2003; Kuhl, 2007) と、それらの研
究とは矛盾するように思える研究成果 (e.g., Dayanim & Namy, 2015), およびその関連で、言語習得にとって極めて
重要な役割を果たすとされる Child-directed Speech (CDS) に関する研究成果 (e.g., Cooper & Aslin, 1990; Martin et al.,
2015) を総括することで、言語習得の社会性について考察する。

結論は以下のとおりである:

- 音素の区別を習得するには社会的学習が必要とされている (Kuhl et al., 2003) が、それを断定できるほどの証
拠が得られているとは言えない
- 仮に音素の習得が社会的であるとすれば、それは Kuhl et al. (2003) が指摘するような統計学習にとって有益と
いうことではなく、音素そのものが「社会的」実態だからである (Cf. Port, 2007).
- 言語習得の社会性を示す根拠として、動画を見るだけでは習得が不可能であるという議論があるが、それと矛
盾する研究成果もあり (e.g., Dayanim & Namy, 2015), 社会的要因が不可欠とは断定できない
- 社会的要因が言語習得に有効であるのは、周囲の大人が幼児に話しかける際の特有の話し方 (CDS) が幼児の
注意を引くからであり、社会的要因そのものが重要なわけではない

2. 背景と問題

言語習得の社会性は以前から指摘され (Cf. Bruner, 1983; Vygotsky, 1962), 理論的にも実証的にも研究の蓄積が存
在する。しかしながら、上に述べたように、言語習得にとって何がどの程度社会的なのか、という点に関しては疑問
の余地がある。そこで、本節では、言語習得の社会性を示す研究、特に、音素の習得にかかわる社会的な刺激の不
可欠性を示す研究 (Kuhl et al., 2003; Kuhl, 2007) を概観したのち、それらと矛盾するように思える研究成果を紹介し
(e.g., Dayanim & Namy, 2015; Saffran, Newport, & Aslin, 1996), 論点を整理する。

2.1 言語習得の社会性 (音素の習得を中心に)

Tomasello (2003) は、言語習得を「社会認知能力 (social-cognitive skills)」の発露としてモデル化し、特に「共同注
意 (joint attention)」を中心とした「他者の意図を理解する」能力を言語習得の基盤とする議論を展開している。他者
と同一物に対する注意を共有する、共同注意への参与は、生後 9 か月ごろから可能となり、語彙の習得に極めて重
要な役割を担うとされている (e.g., Tomasello, 2003: Chapter 3). 同様の研究は多数に存在するが (e.g., Bruner, 1983;
Baldwin, 1995), 語彙習得に関するものが中心であり、最初期の、音声言語にかかわる習得に関する研究はそこまで
盛んではない。音声言語、特に音素の習得に関しては、Patricia Kuhl らによる研究が代表的である。

音素の体系は言語によって異なり、その習得は特定の言語を母語として習得するのに不可欠であると言える。しか
しながら、よく知られているように、生後 1 年に満たない乳児は、母語として習得中の言語には存在しない音素の差
異を認識することができる。例えば、日本語の音韻体系には *l* と *r* の音を区別が存在せず、双方共に *l* 行の子音 (*l*/*r*)
として認識・産出されるが、生後半年頃であれば、日本語を母語とする環境にいても、*l* 音と *r* 音の区別が、英語を
母語とする環境にいる乳児と同程度に可能である (Kuhl, 2007: 110). このような音素の識別能力は 1 歳の誕生日を迎
える頃には消失し、母語には存在しない音素の差異は認識できなくなってしまう。従って、音素の習得とは、このよ

うにあらゆる言語の音素を認識できていた状態から、自身の母語となる言語特有の音の区別のみが可能な状態に適応していく過程であると言える。

Kuhl et al. (2003) は、母語となる言語以外の言語における音素の差異を認識し続けるためには、どのような刺激がどの程度必要であるかを実験的に検証している。実験の対象となったのは英語を母語として習得する環境にいるアメリカ人の乳児 32 名(実験スタート時の月齢は生後 40±3 週 = 9-10 か月児, 平均 9.3 か月)で、その半数には母語話者による標準中国語 (Mandarin Chinese) のインプットを与え、残りの半数は、統制群として同様の環境で英語母語話者が英語によるインプットを与えた。実験は、25 分のセッションを 12 回、4 週間に渡って行われた。各セッションは、それぞれの言語による絵本の読み聞かせ (10 分間) と、玩具による遊びの時間 (15 分間) で構成されている。12 セッション全て完了したのは、標準中国語群で 10 名、統制群で 11 名の、計 21 名であった (Kuhl et al., 2003: 9097)。また、同一条件で別の乳児 32 名 (同様に平均月齢は 9.3 か月) に対し、半数に標準中国語母語話者とのセッションの録音音声のみを聞かせ (A 群)、残りの半数には映像と音声を記録した DVD を見せる (AV 群) セッションを行った。全て完了したのは、A 群で 13 名、B 群で 15 名の計 28 名であった (Kuhl et al., 2003: 9099)。

このようなインプットを得るセッションを終えたのち、各グループの乳児に対し、標準中国語には存在するものの、英語には存在しない音素の対立である、/tɕ/ と /ç/ の識別が可能かどうかの検証が行われた。検証には、選好振り向き法 (head-turn conditioning procedure) が用いられ、スピーカーから継続的に発されている一方の音素の発声が他方の音素の発声に切り替わったタイミングで適切に注意を向けるかどうかのカウントされた (Kuhl et al., 2003: 9097-9098)。

検証の結果、直接母語話者と触れ合う形で標準中国語のインプットを得た乳児のみが、母語話者並みの当該音素対の識別能力 (平均 65.7%) を維持し、全く標準中国語のインプットが無かった統制群乳児はもちろん、音声のみのインプットを得ていた A 群や映像と音声のインプットを得ていた AV 群も、統制群と同程度の識別能力 (平均 56.7%) に低下してしまっていた (Kuhl et al., 2003: 9098-9099)。このことから、音素の識別能力を維持するには、単に定期的なインプットが与えられるだけでなく、他者とのインタラクションが必要であり、その意味で、社会的な手がかりが必要不可欠であるという見解が得られている (Kuhl et al., 2003: 9100)。

2.2 社会的要因不在の言語習得

このような、言語習得の社会性を示す研究成果に対して、生身の人間とのインタラクションを介さないという意味で、社会的要因不在の状態における言語習得の可能性を示す研究成果も存在する。Dayanim & Namy (2015) は、言語未習得児とのコミュニケーション手段として用いられる、身振り言語である、所謂「ベビーサイン」に注目し、ベビーサイン習得を目的とした市販の乳幼児向け動画教材が、効果的に学習を促進することを実験的に示している。

実験では、アメリカ在住の 92 名の 15 か月児を対象に、動画教材に含まれている 18 のベビーサインの習得可否を検証した。検証にあたって、実験対象者は以下の 4 群に分けられた: 1) 動画を見るだけで学習するグループ (20 名)、2) 動画を親と一緒に見るグループ (27 名)、3) 動画ではなく絵本を用いて親がサインを教えるグループ (21 名)、4) 何も学習させないグループ (統制群, 24 名)。統制群以外の 3 群に対しては、自宅で一回 15-20 分のサイン学習時間を与え、それを週 4 日、3 週間行い、学習時間以外はサインに触れることがないように、両親に指示が与えられた (Dayanim & Namy, 2015: 802)。学習の可否は、親による報告と、実験期間における週 1 度の理解テスト、および実験期間終了 1 週間後に行われた産出テストによって検証された (Dayanim & Namy, 2015: 803)。結果、統制群以外のすべての群で有意な学習が確認され、親による報告、および理解テストに関しては、群間に有意な学習成績の差は認められなかった。従って、動画のみの学習でも、親の教示や共同学習と変わらない学習効果が得られたことになる。

この研究の優位性は、動画による学習効果を確実に検証できる形にしている点にある。というのも、例えば動画による学習効果の検証に実在する語を利用してしまうと、動画以外からのインプットがコントロールできないからである。実際、「動画からの習得は不可能」という結論に至っている研究の中には、高頻度の実在する語 (e.g., table, clock) を学習のターゲットとしてしまっているために、統制群にも学習効果が生じてしまい、結果の解釈に難があるものも存在する (e.g., DeLoache et al., 2010)。この点を配慮し、低頻度語 (e.g., medallion) をターゲットに含めている研究も存在するが (e.g., Krcmar, 2011)、ビデオ以外の学習効果は完全にはコントロールできない上に、単語そのものの学習は生じにくいとしても、単語の一部部分や音韻構造など、利用可能な情報は無いわけではない (ただし、Krcmar (2011) の研究に関しては後に再検討する)。その点、Dayanim & Namy (2015) はホームサインを学習ターゲットとしており、このような問題を排除できていると言える。

同様の環境設定は、人工的に作り出した無意味語や、無意味語の配列を生成する文法を学習させる、所謂人工言語習得 (artificial language learning) の研究にも見て取れる。人工言語習得の研究では、そもそも学習のターゲットが実在しない「言語」であるため、実験環境以外での学習の可能性は完全に排除できる。また、人工言語習得研究の多く

が、スピーカーから流れる音声やモニターに映る映像のみから学習させるようなデザインになっており、1歳に満たない乳児でもそのような環境下で学習が可能であることを示している。例えば Saffran et al. (1996) は、9か月児を対象に、babupu, bupada のような3音節からなる無意味語を連続で聞かせ、繰り返される音のパターンから単語境界を統計的に学習させる実験を行っているが、学習に際してはスピーカーから流れる音のみを用いており、社会的な要因は介在していない。このような単語音節など音声的なパターン発見のみならず、語順や予測関係で表される依存構造、および句構造規則で表現される統語構造などを学習可能であることも示されているが (e.g., Gomez & Gerken, 1999; Saffran, 2002), どれも手法は基本的に同質であり、言語の形式的なパターンの習得には特に社会的要因は不可欠ではないことが示唆される。

また、語学習における共同注意の役割についても疑問が提示されている。Akhtar & Gernsbacher (2007) は、共同注意なしでも語学習は達成されること、逆に、共同注意が可能であってもその度合いと語学習の程度が相関しないダウン症児の症例を提示し、語学習にとって共同注意が必要条件にも十分条件にもならないことを示している。

3. 注意と学習

言語習得の社会性を示す研究成果と、社会的要因不在の言語習得の可能性を示す研究成果とが存在するのは、いったいなぜであろうか。本稿では、それぞれの示す実験的研究の成果そのものは誤ったものではないが、その解釈に問題があると考えられる。本節では、それぞれの研究成果を矛盾なく説明できる、言語習得の社会性に関する新たな解釈を提示する。結論から言えば、Kuhl et al. (2003) などの示す実験結果は、習得における社会的要因の必要性を示しているわけではなく、あくまで社会的要因は習得の「促進要因」であると考えられる。では、何が言語習得にとって主要な要因となるのであろうか。本稿では、それは「注意」であると考えられる。

Kuhl et al. (2003) は、セッション中における学習素材 (標準中国語話者、絵本、玩具およびモニター) への注視の度合いを評定しており、その度合いは、映像や音声のみのインプット条件と、生身の人間からのインプット条件との間に有意な差が存在することを報告している (Kuhl et al., 2003: 9099)。そもそも、社会的要因が音素の習得に不可欠である理由の一つとして、社会的なインタラクションが注意を喚起するものであるという点を挙げており (Kuhl et al., 2003: 9100; Kuhl, 2007: 114), 社会的要因に先立って、「注意」という要因が言語習得の促進要因としてより基本的なものであることが伺える。

これに関連して、Kuhl et al. (2003) でも重要視されているが、大人が乳幼児に対して話しかける際に現れる特有の話し方である、Child-directed Speech (CDS, あるいは Infant-directed Speech, IDS) の持つ役割について考えることが有益である。CDS は、反復の多用や発音の明瞭さ、高いピッチ、大きい抑揚などで特徴づけられるが、このような話し方が、大人同士の発話に比べて、有意に子供の注意を引き付けることが実験的に検証されており (e.g., Cooper & Aslin, 1990), 言語習得を促進させる効果があるとされている。

音素の習得に関しては、CDS において、ミニマルペアとなる母音の音素対 (e.g., 日本語における母音の長短: /E/ と /ee/) の発音上の周波数分布 (frequency distribution) が、母語特有の対立に対応するように有意に差別化されていることが示されているが (e.g., Werker et al., 2007), 対照的に、CDS よりも大人同士の発話の方が、わずかではあるが類似音の差別化が有意に大きくなされている、とする研究成果も存在する (Martin et al., 2015)。前者は実験ベースの研究であるが、後者はコーパスベースであり、規模も大きく、またかなり精緻な音声コーディングを行ったもので、方法論的な優位性がある。

このような研究成果を鑑みると、CDS が言語習得に有益に働くのは、その音声的特徴が故というよりは、理由はともあれ、それが子供の注意を引き付ける性質を持っているからであると考えれば、全ての研究成果を矛盾なく包括的に説明できるように思える。つまり、生身の人間とのインタラクションが音素の識別能力を維持させたのは、CDS が注意を引き付けたからであって、インタラクションという社会的な要因が故ではなく、共同注意と語彙習得の関連を指摘する研究成果は、共同「注意」が一つの注意の形であることに由来し、逆に言えば、「共同」でなくとも注意を引き付ける要因があれば学習は可能であるということになる。

このような解釈に問題があるとするれば、先に挙げた Krcmar (2011) による実験的研究において、動画教材に対する乳幼児の注視の度合いと、学習効果の関係性が検証されており、両者には相関がないことが示されている点である (Krcmar, 2011: 790)。この点に関しては、Krcmar (2011: 790) も述べているが、動画への注意とは必ずしも学習対象となる語やサインへの注意とは限らず、単に映像中の事物やキャラクター、色、光、音声などに向けられているだけである可能性も高いため、単純な検証は難しい。ただ、動画による言語習得の可能性を示した Dayanim & Namy (2015) の研究で学習対象となっていたのが音声言語ではなく視覚言語であるベビーサインであったということは、こ

の問題に一つの答えを与えることにつながるかもしれない。というのも、動画教材を使用した音声言語習得の場合、知覚モダリティが映像における視覚刺激と音声という聴覚刺激に分離しており、それらを統合する必要が生じるが、視覚言語の習得であれば、複数の情報を統合する必要は依然として残るものの、知覚モダリティは視覚のみで処理可能であるため、注意のリソースが限定的でも学習が可能であるとも考えられるからである。

4. 結語: 言語の社会性

本稿では、言語習得の社会性を示す研究と、それとは矛盾するように思える研究成果を紹介し、両者を矛盾なく説明できる解釈の可能性を探ることで、言語習得の社会的性質に関する考察を行った。結果、現在得られている研究成果に関しては、社会的要因の不可欠性は見取ることができず、あくまでそれは、乳幼児が学習対象に「注意」を向けるための習得の促進要因であると解釈することができる、という知見を得た。

最後に、このような考察から浮かび上がってくる言語の社会的性質に関して付記しておく。Kuhl et al. (2003) が研究対象とした音素の区別は、各言語コミュニティにおける、どのような音の区別が有意義であるかの取り決めであり、語の意味などと同じように、それ自体が社会的性質を持ったものと言える (Cf. Port, 2007)。一方、本稿で概観した社会的要因不在の言語習得の可能性を示す研究では、その多くが、人工的な言語 (ベビーサイン、人工言語) を学習対象としており、純粋に社会的な対象とは言い難いということも指摘できる。ここから、言語習得の社会性を示す研究成果は、実は言語習得というプロセスの社会性を示しているのではなく、その習得対象である言語の社会性を示している、という可能性が浮かび上がってこよう。

参考文献

- Akhtar, N., & Gernsbacher, M. A. (2007). Joint attention and vocabulary development: A critical look. *Language and linguistics compass*, 1(3), 195–207.
- Baldwin, D. A. (1995). Understanding the link between joint attention and language. In C. Moore & P. J. Dunham (Eds.), *Joint attention: Its origins and role in development* (pp. 131–158). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Bruner, J. S. (1983). *Child's talk: Learning to use language*. New York: W.W. Norton.
- Cooper, R. P., & Aslin, R. N. (1990). Preference for infant-directed speech in the first month after birth. *Child development*, 61(5), 1584–1595.
- Dayanim, S., & Namy, L. L. (2015). Infants learn baby signs from video. *Child development*, 86(3), 800–811.
- DeLoache, J. S., Chiong, C., Sherman, K., Islam, N., Vanderborght, M., Troseth, G. L., Strouse, G. A., & O'Doherty, K. (2010). Do babies learn from baby media? *Psychological Science*, 21(11), 1570–1574.
- Gomez, R. L., & Gerken, L. (1999). Artificial grammar learning by 1-year-olds leads to specific and abstract knowledge. *Cognition*, 70(2), 109–135.
- Krcmar, M. (2011). Word learning in very young children from infant-directed dvds. *Journal of Communication*, 61(4), 780–794.
- Kuhl, P. K. (2007). Is speech learning 'gated' by the social brain? *Developmental science*, 10(1), 110–120.
- Kuhl, P. K., Tsao, F.-M., & Liu, H.-M. (2003). Foreign-language experience in infancy: Effects of short-term exposure and social interaction on phonetic learning. *Proceedings of the National Academy of Sciences*, 100(15), 9096–9101.
- Martin, A., Schatz, T., Versteegh, M., Miyazawa, K., Mazuka, R., Dupoux, E., & Cristia, A. (2015). Mothers speak less clearly to infants than to adults: A comprehensive test of the hyperarticulation hypothesis. *Psychological science*, 26(3), 341–347.
- Port, R. F. (2007). How are words stored in memory?: Beyond phones and phonemes. *New Ideas in Psychology*, 25(2), 143–170.
- Saffran, J. R. (2002). Constraints on statistical language learning. *Journal of Memory and Language*, 47(1), 172–196.
- Saffran, J. R., Newport, E. L., & Aslin, R. N. (1996). Word segmentation: The role of distributional cues. *Journal of memory and language*, 35(4), 606–621.
- Tomasello, M. (2003). *Constructing a language: A usage-based theory of language acquisition*. Cambridge, MA.: Harvard University Press.
- Vygotsky, L. S. (1962). *Thought and language*. Cambridge, MA.: MIT Press.
- Werker, J. F., Pons, F., Dietrich, C., Kajikawa, S., Fais, L., & Amano, S. (2007). Infant-directed speech supports phonetic category learning in english and japanese. *Cognition*, 103(1), 147–162.