

共通基盤の構築における名付けの有用性の分析

齋藤結¹ 光田航² 東中竜一郎² 南泰浩¹

¹ 電気通信大学情報理工学系研究科 ² 日本電信電話株式会社

s2231050@edu.cc.uec.ac.jp

{koh.mitsuda.td, ryuichiro.higashinaka.tp}@hco.ntt.co.jp

minami.yasuhiro@is.uec.ac.jp

概要

対話研究において、話者が知識や信念を共有する共通基盤の構築は重要な課題とされてきたが、その構築過程を分析した研究は少ない。我々は先行研究において、二名の作業者が共同で図形を配置する課題を対象に、重要な共通基盤と考えられる図形への「名付け」に着目してアノテーションを実施した。本研究では、アノテーションされた情報に基づき、共通基盤構築における名付けの有用性の検証、課題の成否につながる名付けの調査、発話中の名付けの自動抽出に取り組んだ。その結果、名付けが有意に共通基盤構築に貢献すること、目標に関する名付けが課題の成否に関連すること、単純なモデルを用いて適合率 0.57 で名付けの抽出が可能になったことが明らかになった。

1 はじめに

対話において、話者が相互理解を行うためには話者間で共有される知識や概念といった共通基盤の存在が必要である [1, 2, 3]。しかし、共通基盤の構築過程を分析した研究は少ない [4, 5, 6]。我々は以前の研究で、共通基盤の構築過程を分析するために集められたデータ（共同図形配置コーパス [7]）を調査し、名付けが重要な役割を果たしていると考えた。そして、データから名付けを取り出すための仕様の作成、および、アノテーションを実施した [8]。

図 1 に、2名の作業者が図形を配置する共同図形配置課題における名付けのアノテーションの例を示す。「自動車の車輪」や「トラック」という名付けを通じて作業が進んでおり、これらは共通基盤構築に必須であると考えられる [9, 10]。しかしながら、我々の先行研究では名付けが共通基盤構築に有用であるかや、課題の成否につながる名付けの出現の傾向は調査できていなかった。

ID	話者	発話	作業画面
2	B	よろしくお願いします。円が2つあるので自動車の車輪のように配置するのはいかがでしょうか	
3	A	わかりました。トラックのようなイメージですね	
4	B	いいですね。では四角形をトラックの運転席と荷台の部分として配置しましょうか。	
5	A	直角ではない四角形をトラックの先頭部分にしますか。ひし形は排気ガスというイメージですかね。	
6	B	わかりました。小さい平行四辺形を運転席の窓にするのはいかがでしょうか。	

図 1 名付けを通じて共通基盤構築が進む対話の例、ID は発話 ID を表し、A、B は話者を表し、作業画面はそれぞれの話者が見ている画面を示す。

本稿では、アノテーションされた情報に基づき、共通基盤構築における名付けの有用性の検証、課題の成否につながる名付けの調査、発話中の名付けの自動抽出に取り組んだ結果について報告する。

2 共同図形配置コーパス

本コーパスでは作業員 2 名がペアとなり、ツールを通じて図形の配置と対話を行う [7]。各作業員の画面には、共通の図形が異なる初期配置で表示される。作業員は相手と対話をしながらマウスを使用して図形を動かし、2 名で共通の図形配置を作る。このとき、各作業員は相手の図形配置を見ることはできない。図形の数 5 個または 7 個、種類は重複あり、大きさはランダムである。図形の回転、縮小拡大、削除はできず移動のみが可能である。相手と同じ図形配置を作成すればこの課題は成功となる。

図 2 に共同図形配置課題で利用される図形を示す。1 つ目が上部に示す三角や四角などの単純図形で、2 つ目が下部に示す建物などの前提知識を利用できる建物図形である。それぞれ 10 種類存在する。

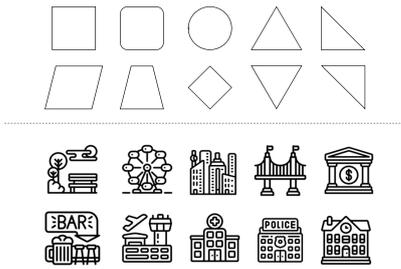


図2 共同図形配置課題で使用された図形

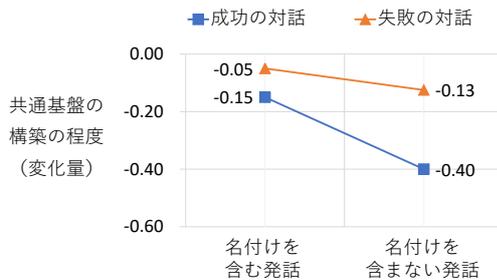


図3 共通基盤の構築の程度に関する二次元配置分散分析(ANOVA)の結果. 縦軸の値が小さいほど共通基盤の構築の程度が大きいことを示す.

3 名付けのアノテーション仕様

名付けのアノテーションデータには2つのラベルが存在する. 1つ目が, 個数・配置に関する分類である. これは, 名付けの対象となる図形の数, および, 対象となる図形が発話時点で配置されているかについての分類であり, 「図形/建物1つ」, 「図形/建物複数・目標」, 「図形/建物複数・その他」の3種類ある. 2つ目が, 名称に関する分類である. これは, その名付けがどのような事物かを表す分類で, 「食べ物」や「登録名」など15種類ある [8].

4 名付けと共通基盤構築の関係

名付けが共通基盤の構築に有用かどうかを調べるために, 光田らが考案した図形配置間距離を指標として利用した [7]. 図形配置間距離は, 図形配置に含まれる任意の2図形間の距離の総和であり, 共通基盤の構築の程度を推定できる. また, 共同図形配置課題では, 図形配置が完全に一致する場合を「成功」, 部分的に一致することを「中間」, それ以外の場合を「失敗」としている. 得られた名付けのアノテーション結果を用いて, 成功対話と失敗対話の共通基盤の構築過程と名付けの関係を調査した.

図3に, 共通基盤の構築の程度について, 二元配置分散分析(ANOVA)を行った結果を示す. 横軸は発話に名付けが含まれているかどうか, 縦軸はその発話における図形配置間距離の変化量(共通基盤

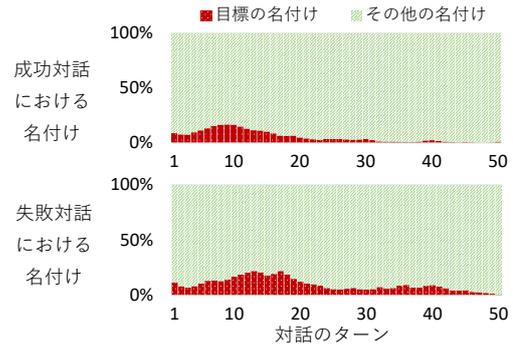


図4 各対話ターンの「目標」と「その他」の名付けのヒストグラム. 上段と下段は, 成功対話と失敗対話のヒストグラムになる.

の構築の程度)を表している. 縦軸の値が小さいほど, 共通基盤の構築が進んでいることを意味する. 2本の線は課題の成否に対応しており, 成功は作業仲間間で図形配置が一致していること, 失敗は一致していないことを示す. 二元配置分散分析の結果, 名付けと共同図形配置課題の結果の主効果は有意であり, それらの交互作用も有意であることがわかった. この結果は, 我々が注目している名付けが共通基盤の構築に有意に影響することを示している. また, 成功対話における名付けが, 失敗対話における名付けよりも共通基盤の構築を有意に進めるという事実は, これらの対話における名付けの質の違いを示唆する.

5 対話の分析

前節の調査から名付けの有用性が確認できたため, 本節では, 課題の成否につながる名付けの出現の傾向を調査する.

5.1 目標とその他の分析

図4は, 対話の各ターンにおける「目標」と「その他」の名付けのヒストグラムを表す. 上段/下段は, 成功対話と失敗対話を示している. 各対話のターン数は, 線形変換を用いて50に正規化した. このグラフから, 成功対話と失敗対話で, 目標に関する名付けの傾向に違いがあることが確認できる. 成功対話では, およそ10ターン目に最も頻繁に出現し, 20ターン目以降には頻度が少なくなっている. 失敗対話では, 目標の名付けは成功対話よりも遅いおよそ15ターン目に最も頻繁に現れ, 20ターン目以降も10%程度現れている. このことから, 初期の段階で目標の名付けについて話し合い, 目標としたものを作るための構成要素の名付けをしながら

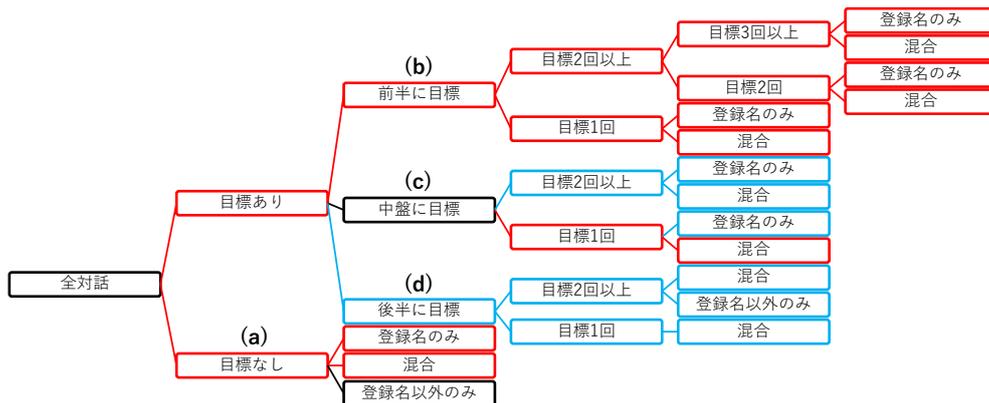


図5 名付けに注目した対話の分類結果. 赤は成功対話のほうが失敗対話より多いことを示し, 青はその逆を示す. 黒は成功対話と失敗対話の数が等しいことを示す.

対話を行うことの重要性が明らかになった. また, 成功対話と失敗対話において, 最初の目標の名付けが現れる対話ターンの平均はそれぞれ 5.7 と 8.1 であり, 両者には有意な差が見られた ($p < 0.05$). これは, 成功対話の初期に目標の名付けが現れることを定量的に示しており, 初期段階での適切な目標の名付けが, 共同図形配置課題を成功へと導くことを示唆している.

5.2 対話の流れの分析

ここまで述べた通り, 対話の初期における適切な目標の名付けが共同図形配置課題を成功へと導くことがわかった. 次に, 目標の出現回数や出現タイミングに注目して, どのような対話の流れが成功につながるのかを調べる. このとき, 目標の名付けに加えて, 最頻の名付けである「登録名」にも着目する. 登録名は共同図形配置課題で使用される図形にあらかじめ付けられている名前であり, 理解が容易なことから, 目標の名付けと合わせて共通基盤の構築に役立つと考えられる.

図5に, 目標と登録名の名付けの有無やその頻度に基づいて対話を著者らが人手で分類した結果を示す. なお, 目標の出現位置によって対話の流れが変化すると考えたため, 対話の前半, 中盤, 後半の3つに分けて分類した. 目標が出現する対話の平均対話長が 22.5 であったため, 前半を 5 発話目まで, 中盤を 6 発話目から 17 発話目まで, 後半を 18 発話目以降とした. 目標の出現回数については, 目標が 2 回しか出現しない対話が全体の約 7 割と多くを占めていたため, 目標の出現回数が 3 回以上のものはまとめた.

名付けに基づく対話の分類から, (a) のように目標

が出現しなかった場合であっても, 多くの対話で課題が成功していることがわかる. このような対話では, 登録名を適切に用いて相手と意思疎通を行っている過程が見られた. (b) に示されるように, 目標が出現する場合は対話の前半に目標を提案することで, 成功につながるということがわかる. 早い段階での方針の決定, つまり, 考えの大枠を対話の冒頭で共有することで, 共通基盤の構築が進むと考えられる. (c) に示されるように, 対話の中盤で目標を提案する場合, 2 回以上提案すると失敗対話のほうが多くなっている. これらについては, 対話がかみ合っていないか, 様々な考えが共有され, どの話をしているのかわからない, といった対話が見られた. (d) のように, 対話の後半に目標を提案すると, 失敗する可能性が上がるということが示唆される. 例えば, 対話の後半で考えがまとまるも時間が足りない場合などが考えられる.

なお, どの対話でも登録名が使われていることがわかった. 作業者の前提知識が利用でき, 図形から容易に想像できる名付けは作業者間で確実に伝わる名付けであることから, 理解が容易な情報は共通基盤の構築に貢献するということが確認できた.

6 名付けの抽出実験

以上の分析から, 名付けが共通基盤の構築に有用であるということが確認できた. 分析を通じて, 我々は今後共通基盤を構築する過程のモデル化に取り組むため, そのためには名付けを特定し, 対応する図形を同定したり, 名付けの意図を理解したりすることが重要となる. そこで本研究では, モデル化の初期段階として, 対話から名付けを抽出する実験に取り組んだ.

表 1 対話を入力とした名付けの抽出と分類の評価結果

	適合率	再現率	F 値
	0.576	0.256	0.354

6.1 データセットとモデル

本研究では最も基本的な名付けとして、図形の種別が一意に決まる名付けを対象に、Bidirectional Encoder Representations from Transformers (BERT) [11] を用いた名付けの抽出実験を行った。入出力としては、対話（開始の発話から何らかの名付けが含まれる発話まで）を入力とし、名付け箇所のスパン、および、図形の種別を出力とする。学習データは 29,055 個、図形の種別は 20 種（図 2）となる。データセットは訓練、検証、評価の割合が 6 : 2 : 2 となるように分けた。モデルには東北大学が公開している BERT の日本語版事前学習済みモデル¹⁾を使用した。入力長は最大で 256 トークンとし、最大エポック数を 5 に設定して fine-tuning を行った。

6.2 評価結果

表 1 に、対話を入力とした名付けの抽出と分類の評価結果（適合率、再現率、F 値）を示す。単純なモデルでもある程度名付けの抽出は可能であり、学習データを増やせば、より高い性能で抽出が可能だと考えている。

図 6 に、モデルが入力対話から名付けを抽出した出力例を示す。登録名の名付けが適切に抽出できていることに加え、「修飾語 + 登録名」のような名付けも一部分ではあるものの抽出ができています。正解の名付けの一部しか抽出できていない場合でも、正解と同じ図形を特定できているため、図形の特定という観点でもある程度の精度が実現できていると考えている。

モデルが抽出した名付けの内、名付けと対応図形が正解と一致した 28 件を以下に示す。

空港、観覧車、橋、銀行、病院、公園、警察、バー、ビル、台形、大きい橋、大きな橋、学校、警察署、大きな観覧車、交番、小さな観覧車、大きい銀行、丸、大きい観覧車、小さな橋、大きい警察、小さい観覧車、小さい橋、大きい公園、小さいバー、小さい銀行、飛行場

抽出された名付けの内、登録名の割合が比較的多い。学習データで使用している名付けの 6 割強が登録名の名付けであることから、登録名が多く抽出されるということを考慮しても、登録名の名付けが多

入力される対話

空港、バーが同じ大きさですのでどうでしょうか？どれかをメインにして町を考える感じです。そうですね。中心は空港かバーがいいと思いますが、どうでしょうか？私も空港を中心にいいかなと思ってました。空港にしましょう。ではどうしましょう。空港の隣に大きな公園があるのはいいなと思ったのですが、どうですか？

正解およびモデルの出力（一部）

正解の名付け	正解の図形	抽出した名付け	推定した図形
バー	bar	—	—
空港	airport	空港	airport
空港の建物	airport	—	—
大きな公園	park	公園	park

図 6 入力される対話と抽出される名付け（一部）の例

い。今回用いた比較的単純な手法では、共通基盤の構築に必要な名付けを全て抽出するのは難しい可能性がある。

抽出した名付けは正解と一致しているが、特定した図形が正解と一致していないものを以下に示す。

ビル、バー、学校、病院、ベンチ、はし、街、公園、空港、銀行、飛行機

街に対応する図形は、作業者によって様々な名付けがされているため、図形の特定が難しい結果となった。

現状の方法では、対話から名付けを抽出できたとしても、登録名の名付けが比較的多く抽出されること、また、文脈によっては登録名でも対応する図形を特定することは困難なことがわかった。今後、共通基盤の構築に必要な名付けの抽出に関する新たな手法の検討が必要である。

7 おわりに

本稿では、共同図形配置コーパス中の名付けのアノテーションを用いて、共通基盤構築における名付けの有用性を明らかにするために、図形配置間距離を用いて二元配置分散分析を行い、有用性を確認した。また、名付けに注目して対話の分析を行い、どのような名付けから共通基盤が構築されるのかを明らかにした。BERT を用いて対話から名付けを抽出するモデルの作成も行い、既存の機械学習モデルを用いた名付けの理解の可能性を調べた。その結果、現状でもある程度名付けの抽出は可能であるが、文脈や作業者の背景知識を必要とする名付けの抽出は難しいことがわかった。今後は、文脈から名付けを適切に抽出する手法や、抽出した名付けを図形と紐付けて理解する手法の確立に取り組んでいく予定である。

1) <https://huggingface.co/cl-tohoku/bert-base-japanese-whole-word-masking>

参考文献

- [1] Herbert H Clark. **Using language**. Cambridge university press, 1996.
- [2] David R Traum. A computational theory of grounding in natural language conversation. Technical report, Rochester Univ NY Dept of Computer Science, 1994.
- [3] Stefan Kopp and Nicole Kramer. Revisiting human-agent communication: The importance of joint co-construction and understanding mental states. **Frontiers in Psychology**, Vol. 12, , 2021.
- [4] Kris Liu, Jean E Fox Tree, and Marilyn Walker. Coordinating communication in the wild: The Artwalk dialogue corpus of pedestrian navigation and mobile referential communication. In **Proc. of LREC**, pp. 3159–3166, 2016.
- [5] Takuma Udagawa and Akiko Aizawa. An annotated corpus of reference resolution for interpreting common grounding. In **Proc. of AACL**, pp. 9081–9089, 2020.
- [6] Cristian-Paul Bara, Sky CH-Wang, and Joyce Chai. Mind-Craft: Theory of mind modeling for situated dialogue in collaborative tasks. In **Proc. of EMNLP**, pp. 1112–1125, 2021.
- [7] Koh Mitsuda, Ryuichiro Higashinaka, Yuhei Oga, and Sen Yoshida. Dialogue collection for recording the process of building common ground in a collaborative task. In **Proc. of LREC**, pp. 5749–5758, 2022.
- [8] 齋藤結, 光田航, 東中竜一郎, 南泰浩. 対話での共通基盤構築過程における名付けの分析. 言語処理学会第 28 回年次大会発表論文集 (NLP2022), pp. 38–42, 2022.
- [9] Herbert H Clark and Deanna Wilkes-Gibbs. Referring as a collaborative process. **Cognition**, Vol. 22, No. 1, pp. 1–39, 1986.
- [10] Anya Ji, Noriyuki Kojima, Noah Rush, Alane Suhr, Wai Keen Vong, Robert D. Hawkins, and Yoav Artzi. Abstract visual reasoning with tangram shapes. In **Proc. of EMNLP**, pp. 582–601, 2022.
- [11] Jacob Devlin, Ming-Wei Chang, Kenton Lee, and Kristina Toutanova. BERT: Pre-training of deep bidirectional transformers for language understanding. **arXiv preprint arXiv:1810.04805**, 2018.