

国立国語研究所学術情報リポジトリ

On the Characteristics of Introducing New Referring Objects in Japanese Map Task Dialogue

メタデータ	言語: jpn 出版者: 公開日: 2020-02-06 キーワード (Ja): キーワード (En): 作成者: 川端, 良子 メールアドレス: 所属:
URL	https://doi.org/10.15084/00002562

地図課題対話における参照導入方法の特徴

川端 良子 (国立国語研究所音声言語研究領域) *

On the Characteristics of Introducing New Referring Objects in Japanese Map Task Dialogue

Yoshiko Kawabata (National Institute for Japanese Language and Linguistics)

要旨

対話において、相手が知っているかどうか不確かな対象に言及する際、話し手はどのようにその対象を対話に導入するのだろうか。本研究では『日本語地図課題対話コーパス』を用いて、特定の対象が最初に対話に導入される際の言語活動の分析を行った。本稿は、(1) 発話機能、(2) 相互行為、(3) 言語形式の3つの観点からその言語活動の特徴を報告する。

1. はじめに

対話の中で特定の対象が参照されるとき、話し手と聞き手が同一の対象を想定していることは、コミュニケーション成立のための不可欠な要素の一つである。そのため、話し手は、自分が意図した対象と同じ対象を聞き手が特定できるように、対象に関する聞き手との共有信念を想定して発話を計画する。しかし、話し手がいつも確信をもって聞き手との共有信念を想定できるとは限らない。聞き手がその対象を知っているか否か不確かな場合もある。このような場合に話し手はどのように対象を対話に導入するのかということが本研究の関心である。

本研究を推進するにあたり、分析対象として『日本語地図課題対話コーパス』(以降では、マップタスクと呼ぶ)(堀内ほか 1999)を選ぶことにした。その理由は、マップタスクではコーパスの設計上(詳しくは後述)、共有信念が不確かな地図上のランドマークに、対話参加者が何度も言及するため、参照対象の導入方法について定量的な調査が可能だからである。具体例を(1)に示す。これは、「出発地点」と「石の砂漠」の2つのランドマークが実験の1セッション内で初めて言及されたときの音声言語列の一部を転記したものである。

(1) えと出発地点の下に石の砂漠がありますよね。

(マップタスク:j1n2)

本報告では3つの観点から対象の対話への導入方法の特徴について述べる。1つめの観点は、発話の機能に着目したものである。(1)は一つの発話単位(Jap 2017)となっている。末尾の句点「。」は発話単位の末尾であることを示している(以降の事例でも同様)。発話単位は、コミュニケーション上のなんらかの行為の遂行と捉えることができる(居關ほか 2018)。この発話単位は「石の砂漠」が相手の地図上に存在することを聞き手に確認している。一方、「出発地点」は聞き手の地図に存在することを前提にしており、相手からの情報を要求していない。

* kawabata@ninjal.ac.jp

このように、発話単位の中で参照対象がどのように提示されるかについて、マップタスクでは典型的な方法がいくつかある。この方法を分類し、ランドマークがどのように対話に導入されるのか、その傾向を報告する。英語版地図課題対話コーパス (Anderson et al. 1991) で行われた同類の分析 (Anderson and Boyle 1994) では、導入の形式を”Question” と”Non-question” の2種類に分けている。本稿では、”Non-question” をさらに細かく分類する。

2つ目の観点は、発話の途中で行われる対話参加者間のやりとりに着目するものである。(2) は「銀鉱」(実際は「銀の鉱山」という表現が用いられている) が対話に導入されたの言語活動を記したものである。一番左の列は当該発話内での連番、2列目は発話者の識別子 (G/F)、3列目は発話の開始時間と終了時間 (単位:秒)、4列目が転記テキストである。転記における {} 内の数字は発話の途中の無音区間で示している (単位:秒)、「?」記号は直前が上昇調で発話されていることを示す (以降の転記でも同様の表記を用いる)。ここでも「銀鉱」は(1)と同様に確認によって対話に導入されているが、対象が互いの地図上にあることを共有するタイミングが(1)とは異なっている。2行目のFの「ある」という発話は、Gの発話の途中の休止の間行われているため、1行目のGiverの発話が完了する前に、Followerの地図上にも銀鉱が存在することが両者にとって既知となっている。こうしたやりとりについての研究 (Sacks and Schegloff 1979, Clark and Wilkers-Gibbs 1986, 串田秀也 2008) はあるものの、実際にどのような状況でどの程度使用されているのかということについて十分な調査は行われていない。

(2)	No	話者	開始 - 終了	転記テキスト
	1	G	23.536 - 27.900	でさ {0.7}銀の鉱山がすぐ下に {0.8} あるでしょう?。
	2	F	26.848 - 27.008	ある。
	3	F	27.856 - 28.000	うん。 (マップタスク:j2n4)

3つ目の観点は、言語表現に注目するものである。日本語では、対象に対する相手との知識に差がある場合、「って」「という」などの引用の形式が用いられることがあることが知られている (田窪 1989)。上述した発話機能ごとに引用形式がどの程度使用されているか報告する。

2. 地図課題対話コーパス (マップタスク)

『日本語地図課題対話コーパス』は、2名の被検者が地図課題と呼ばれる共同的課題の遂行中に行われた言語活動が収録されている。地図課題では参加者の一方に、経路が描かれた地図が渡され、他方には同じ地域で経路が描かれていない地図が渡される (図 1)。経路が描かれた地図をもった参加者は Giver、経路が描かれていない地図を持つ参加者は Follower と呼ばれる。両者の課題は、Giverの地図に描かれた経路の情報を、音声言語を用いて Follower に伝達し、Followerの地図上に Giverの地図と同じ経路を再現することである。実験を行う際、両者はそれぞれ別の部屋に入り、お互いの作業の様子が見えない状態で課題を遂行する。参加者は4人で一組になり、対話相手と役割を交換しながら、一人4回 (Giver2回, Follower2回) 課題を行う。このような4人組が16組あり、地図課題全体では128対話が収録されている。

地図の経路は、図1のように複数のランドマークの間を通過して描かれている。そのため、対話参加者は経路を説明するために複数回ランドマークに言及する必要がある。ランドマーク

は、Giver の地図と Follower の地図の両方に存在する場合もあれば、片方の地図にのみ存在する場合がある。ただし、出発地点は両者の地図上にあり、目標地点は Giver の地図のみにある。しかし、そのことは両者に明示的には知らされていないため、出発地点と目標地点を含め、対話に導入するまでは相手が特定のランドマークを知っているかどうか不確定な状況になっている。

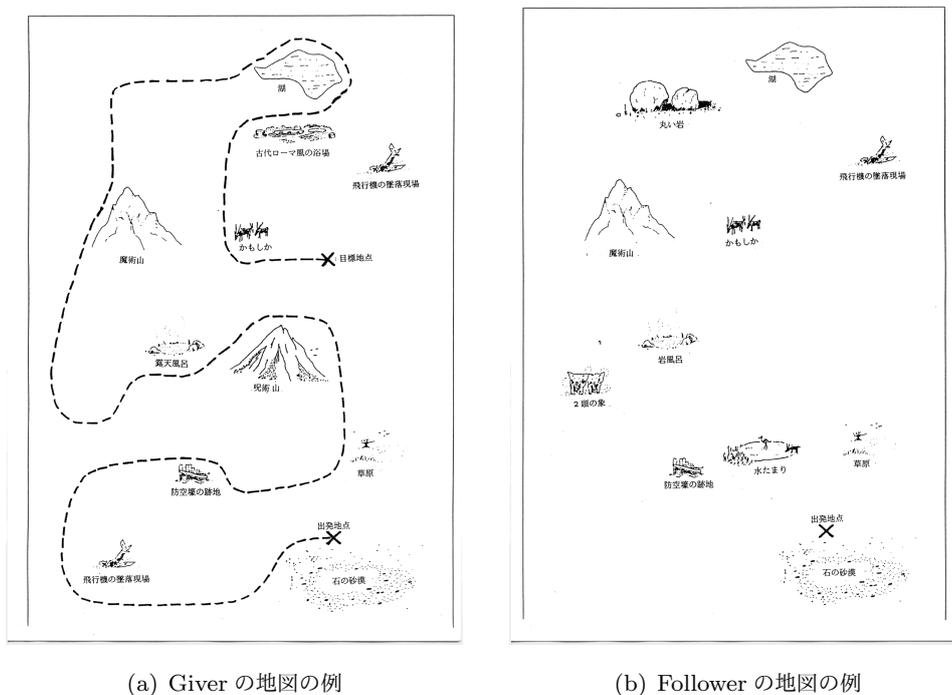


図1 地図課題でそれぞれの参加者に渡される地図の例

3. アノテーション

分析の手順として、まずランドマークが最初に対話に導入された箇所を抽出する。抽出する単位として、「発話単位」(Jap 2017)を用いた。発話単位とは、話し手と聞き手が行為や情報を交換する際の基本単位として提案されたもので、統語的、談話的、相互行為的な一まとまりに対応する単位とされる。最初に導入されたランドマークとは、ランドマークを含む発話単位の開始時点が一番早いものとした。ただし、その発話が聞き手に聞こえなかったと判断できる場合や、話し手が発話を言いかけて途中でやめた場合は、その後の発話の中で最初にランドマークを含む発話単位を分析対象とする⁽¹⁾。次に、ランドマークごとに以下のアノテーションを施す。

⁽¹⁾ 地図課題のランドマークのうち、2to1条件のものは2つ存在するランドマークのどちらを参照しているか区別することが困難だったため区別せず、最初に参照されたランドマークのみを分析に用いた。また、名称変更条件で名称が異なっていたとしても、同じ対象を参照しているものとして、最初に言及されランドマークだけを分析に用いた。

3.1 発話機能による分類

各ランドマークの対話への導入方法を，ランドマークが含まれている発話単位の機能から分類を行う。分類は「存在確認」「存在提示」「存在前提」「非明示」「介入停止」「予告」「独り言」「その他」の8種類がある。ランドマークのそれぞれについて，どの方法によって対話に導入されているかを判定する。1つの発話単位内に初めて言及されるランドマークが複数ある場合は，同じ発話単位に対してランドマークの数だけ分類を行う。

■**存在確認** 発話単位が，当該のランドマークの有無について聞き手に情報を要求する場合，その導入方法を「存在確認」とした。(1)の「石の砂漠」の導入方法がこの分類となる。「出発地点」については後述の「存在前提」となる。

■**存在提示** 発話単位が，自分または相手の地図に当該のランドマークがあることや，ランドマークの特徴を相手に伝える場合，その導入方法を「存在提示」とした。(3)の1行目では「ゴルフ場」が自分の地図上にあることをFollowerがGiverに伝えている。このような発話が典型的である。また，(4)のような質問に対する答えも本分類とした。

- (3) 1 F 373.586 - 375.413 えとここにゴルフ場てのがあんですよ。
 2 F 375.413 - 376.571 そちらに無いようです。
 3 G 376.166 - 376.512 無いよ。

(マップタスク:j1n1)

- (4) 1 G 434.394 - 435.857 えと魔術山の上に何かありますか。
 2 F 436.256 - 438.400 えと上に丸い山 丸い岩。

(マップタスク:j1n2)

■**存在前提** 上記の2つのタイプはランドマークの存在について情報の授受を行うことが発話の機能であったが，ランドマークが存在していることを前提にした発話を行う場合も少なくなかった。これを「存在前提」という分類にした。(1)の「出発地点」や(5)の「牧場」のようにランドマークが修飾部に表われる場合がその典型である。

- (5) 1 G 479.216 - 480.800 で今度牧場の上を通る。
 2 F 480.992 - 482.352 その牧場てゆうのは無いから。

(マップタスク:j1n3)

■**非明示** ランドマーク導入の方法には，(4)のような質問に対する答えでないにもかかわらず，文末の要素がなく「存在確認」なのか「存在提示」なのか判断できない場合がある。これを「非明示」とした。(6)に非明示の具体例を示す。この場面でFollowerは「洞窟」とだけ発話しており，質問なのか情報提供なのか判断できないために本タグを付与する。

- (6) 1 F 441.299 - 441.829 洞窟。
 2 G 442.523 - 444.305 えと 洞窟はここに無いんですけど。

(マップタスク:j1n4)

■介入停止 上記の「非明示」は話し手が自発的に発話を止めているのに対して、聞き手が発話を開始したために、途中で発話を停止する場面があった。これを「介入停止」とした。(7)に介入停止の具体例を示す。この場面では、Followerが「踏切」に言及して発話を行う途中で、Giverが踏切は地図上にないことを伝えている。そしてその後、Followerは発話を打ち切っている。このような場合、「踏切」を含むFollowerの発話の機能が判断できないため、「介入停止」というタグを付与する。

- (7) 1 F 558.304 - 559.664 と踏切てゆうのを。
 2 G 559.680 - 560.208 無いです。
 3 F 560.288 - 560.784 無いですか。

(マップタスク:j4n6)

■予告 ランドマークの有無や存在する場所に関する情報をその発話以降で説明するために、ランドマークの名称のみを対話に導入する場合を「予告」とした。具体例を(8)に示す。この場面のよう「なんですけど」という表現や「ですが」という表現が典型的に用いられる。

- (8) 1 G 331.488 - 333.232 で最後の目標地点なんですけど。
 2 F 333.120 - 333.296 はい。
 3 G 333.904 - 340.704 えと {0.1} 紙の {1.1} 左端から {0.2} あえと一番下の端から {0.5} 三センチ右で。

(マップタスク:j1n5)

■独り言 聞き手に向かって発話しているというよりは、自問しており、それにもかかわらず聞き手がそれに応答してしまう場合を「独り言」とした。(9)に具体例を示す。この場面で導入されようとしているランドマークは「クレバス」であるが、Followerは「クレバス」が何であるか認識できておらず、2行目で自問している。それに対してGiverが反応し、対話に導入される結果となっている。

- (9) 1 F 435.460 - 437.912 じゃねそこにね私の絵が描いてあんだよ。
 2 F 437.912 - 440.192 何かククレバスてなんだ。
 3 G 439.919 - 440.619 クレバス?。
 4 F 440.493 - 441.953 クレバスてゆう絵があんだよ。

(マップタスク:j4n7)

■その他 上記に該当しない発話を「その他」とした。

3.2 発話の分割提示

1節の(2)のように、話し手は発話途中に休止を入れ、発話を分割して提示することで、発話の途中で聞き手の応答を獲得することができる。このような発話を、Clark and Wilkers-Gibbs (1986)は'Installment noun phrase'と呼んでいる。本稿では、このように休止によって発話を分割して提示することを「分割提示」と呼び、ランドマークを最初に対話に導入する際、分割提示が行われているかどうか判定を行う。ランドマークに関する知識について相手からの応

答を獲得するために分割提示を行うかを判定するため、ランドマークの直後に休止がある場合に分割提示ありとする。つまり、(2)で「銀の鉱山」の直前の0.7秒の休止はここでの分割提示の対象とはせず、直後の0.8秒の休止を対象とする。休止の長さはおよそ0.3秒以上の場合とするが、話し手の話速からあきらかに休止を入れているととらえられる場合は、それ以下の長さの休止でも分割提示ありとする。

3.3 引用表現の使用

ランドマークの直後に副助詞「って」や格助詞「と」を伴う場合を「引用あり」とし、ない場合を「引用なし」とする。

4. 結果

4.1 発話単位の機能からみた導入方法の使用傾向

地図課題対話全128中の64対話においてランドマークが最初に対話に導入される方法を「出発地点」「目標地点」「ランドマーク一般(一般)」別に集計した結果を表1に示す。出発地点と目標地点を他のランドマークと分けて集計する理由は、両者は他のランドマークとは条件が異なるため、会話への導入方法にも違いが生じる可能性があるためである。

表1 ランドマーク別導入方法の使用頻度と比率

	出発地点		目標地点		一般	
	頻度	(比率%)	頻度	(比率%)	頻度	(比率%)
存在確認	29	(46.0)	13	(21.7)	485	(60.2)
存在提示	12	(19.0)	32	(53.3)	135	(16.7)
存在前提	16	(25.4)	5	(8.3)	109	(13.5)
非明示	3	(4.8)	0	(0)	48	(6.0)
介入停止	0	(0)	0	(0)	18	(2.2)
予告	1	(1.6)	8	(13.3)	0	(0)
独り言	2	(3.2)	2	(3.3)	5	(0.6)
その他	0	(0)	0	(0)	6	(0.7)
合計	63	(100)	60	(100)	806	(100)

「出発地点」と「一般」はどちらも「存在確認」の使用が最も多く、「目標地点」では「存在提示」が多かった。「出発地点」は2番目に多い方法が「存在前提」で、「一般」では「存在提示」が多かった。この違いは、各ランドマークに対する対話参加者の認識の違いを反映したものと考えられる。出発地点と目標地点は、すべての地図課題で存在し、しかも全て同じ条件(出発地点は互いの地図に存在、目標地点はGiverの地図のみ存在)で存在する。したがって、参加者は課題を繰替えし行うことで、出発地点は互いの地図にあること、目標地点はGiverの地図にのみ存在するという予測が立つ。このことが、出発地点はランドマーク一般と比べて「存在前提」の使用が多く、目標地点では「存在確認」が少なく「存在提示」が多くなる理由と考えられる。本研究の関心は、相手が知っているか否か不確かな対象物の導入方法であるため、条

件を揃えるため、以降は「ランドマーク一般」のみを分析対象とする。

表2はランドマーク一般の対話への導入方法を話し手の役割別に集計し、役割によって導入方法の使用傾向が異なるかを分析した結果を示している。各導入方法の使用/不使用を従属変数（二項分布）、役割を固定効果、個人を変量効果として、一般化線形混合モデルを適用して分析を行った。分析にはRのlme4パッケージ(Bates et al. 2015)を用いた。導入方法と役割に有意な関係がみられた種類のp値を太字で示している(以降の分析も同様)。表2が示すように、GiverはFollowerに比べて「存在確認」を用いる傾向があり、FollowerはGiverに比べて「存在提示」を用いる傾向がみられた。また、「介入停止」についてはFollowerに多い傾向がみられた。

表2 ランドマーク一般の役割別対象導入方法の使用頻度と比率

	Giver		Follower		p 値
	頻度	(比率 %)	頻度	(比率 %)	
存在確認	414	(71.9)	71	(30.9)	< .001
存在提示	42	(7.3)	93	(40.4)	< .001
存在前提	78	(13.5)	31	(13.5)	.917
非明示	29	(5.0)	19	(8.3)	.114
介入停止	9	(1.6)	9	(3.9)	.054
予告	0	(0)	0	(0)	-
独り言	3	(0.5)	2	(0.9)	.573
その他	1	(0.2)	5	(2.2)	-
合計	576	(100)	230	(100)	

4.2 分割提示の使用傾向

発話を分割して提示する頻度と発話機能の関係について集計し、分析した結果を表3に示す。分割の有無を従属変数（二項分布）、各導入方法の使用の有無を固定効果、個人を変量効果として、一般化線形混合モデルを適用して分析を行った。

表3に示されているように、「存在前提」の発話で発話を分割して提示する傾向がみられた。一方、「存在確認」と「非明示」では分割する頻度が少ない傾向がみられた。

4.3 引用表現の使用傾向

引用形式としては、表4の表現が使用されていた。引用形式を使用する頻度と発話機能の関係について集計し、分析した結果を表5に示す。引用形式の有無を従属変数（二項分布）、各導入方法の使用の有無を固定効果、個人を変量効果として、一般化線形混合モデルを適用して分析を行った。

表5に示されているように、「存在確認」の発話の際、引用形式を用いる傾向がみられた。一方、「存在提示」、「存在前提」、「非明示」では引用形式が使用されない傾向がみられた。

表3 ランドマーク一般の導入方法と分割提示の頻度

	分割なし		分割あり		p 値
	頻度	(比率 %)	頻度	(比率 %)	
存在確認	424	(61.5)	61	(52.1)	.033
存在提示	119	(17.3)	16	(13.7)	.347
存在前提	76	(11.0)	33	(28.2)	< .001
非明示	46	(6.7)	2	(1.7)	.045
介入停止	15	(2.2)	3	(2.5)	.739
予告	0	(0)	0	(0)	-
独り言	4	(0.6)	1	(0.8)	.601
その他	5	(0.7)	1	(0.8)	-
合計	689	(100)	117	(100)	

表4 使用されていた引用形式

て	と
ての	
て + 名詞句	
ていうの	とていうの
ていう + 名詞句	とていう + 名詞句
て書いてある + 名詞句	とかいう

表5 ランドマーク一般の導入方法と引用形式の使用頻度

	引用なし		引用あり		p 値
	頻度	(比率 %)	頻度	(比率 %)	
存在確認	229	(46.7)	256	(81.0)	< .001
存在提示	96	(19.6)	39	(12.3)	.003
存在前提	101	(20.6)	8	(2.5)	< .001
非明示	45	(9.2)	3	(0.9)	< .001
介入停止	10	(2.0)	8	(2.5)	.244
予告	0	(0)	0	(0)	-
独り言	4	(0.8)	1	(0.3)	.175
その他	5	(1.0)	1	(0.3)	-
合計	490	(100)	316	(100)	

5. 考察

Giver と Follower で対象の対話への導入方法が異なること (表 2) については、次の理由が考えられる。マップタスクにおいて課題を成功させるためには、Follower に正しい経路を描写させる必要がある。それには、Giver が Follower に理解可能なように経路説明を行う必要がある。そのためには、Follower の地図上にどのようなランドマークがあるかを知っていることが得策である。よって、Giver は Follower からランドマークに関する情報を要求するために「存在確認」を多く用いたと考えられる。一方、Follower は、自分の地図上のランドマークの有無を Giver に伝えることが課題の遂行に有効だが、相手の地図上にランドマークがあるかどうかは必ずしも課題の遂行に関わらない。この課題と役割の違いが言語活動に差を生じさせたのではないかと考えられる。

ランドマークの「分割提示」が「存在前提」の場合だけ有意に多い (表 3) 理由については、次の理由が考えられる。分割提示は使用することで発話時間が長くなり、途中で聞き手に割り込まれる可能性が高くなることから、特に相手の反応を求める場合以外は効率的な方法ではないと考えられる。「存在前提」は、相手が対象を知っていることを前提とするために、知らなかった場合にコミュニケーション上のトラブルが生じる可能性がある。そのため、「存在前提」によって対象を導入する場合は、相手の反応を伺いながら発話を行なうことでトラブルを回避しようとしていたのではないだろうか。

引用形式の使用の差異 (表 5) について、田窪 (1989) によれば、話し手・聞き手の非共有知識を指すときに引用形式が用いられる (p.226)⁽²⁾。「存在確認」は、相手と対象を共有していない可能性を想定して行うものであり、「存在前提」は、相手と対象を共有していると想定して行うものである。この理由から、「存在確認」では引用形式が多く、他の導入方法では引用形式が少なかったと考えられる。

本研究の結果 (表 1, 表 2, 表 3, 表 5) は上述したように我々の直感に沿うものである。言いかえれば、今回提案した分類方法は、地図課題対話において対象を導入する際に行われていることを特徴づけることに成功しているといえる。そして、本研究によって、地図課題対話における対象の導入方法の概要が捉えられたと考えられる。

6. まとめと今後の課題

今回の研究では、対象を対話に導入する際の言語活動の特徴付けする方法を提案し、提案した方法を用いて、全体的な傾向をとらえることができた。今後は、本結果を基礎にして、さらに詳細な分析を行う予定である。具体的には、親近性の有無や課題の熟練度による参照対象の導入方法の違いや、参照対象の導入方法の違いによってその後の会話の流れがどのように変化するかなどについて分析する予定である。また、マップタスク以外のコーパスでも分析を行い、汎用的な分析方法の確立と理論の構築を行うことも課題である。

謝 辞

⁽²⁾ 正確には、引用形式ではなく、「メタ用法」という表現が用いられている。

本研究は JSPS 科研費 JP19K13196 の助成を受けたものです。

文 献

- 堀内靖雄・中野有紀子・小磯花絵・石崎雅人・鈴木浩之・岡田美智男・仲真紀子・土屋俊・市川熹 (1999). 「日本語地図課題対話コーパスの設計と特徴」 人工知能学会誌, 14:2, pp. 261–272.
- Japanese Discourse Research Initiative (2017). 『『発話単位ラベリングマニュアル』 version 2.1』. <http://www.jdri.org/resources/manuals/uu-doc-2.1.pdf>
- 居關友里子・門田圭祐・伝康晴 (2018). 「日本語日常会話コーパス」への談話行為アノテーションの試み：タグ選択が困難な事例に焦点を当てて」 言語資源活用ワークショップ発表論文集 3 巻, pp. 47–56. 国立国語研究所.
- Anne H. Anderson, Miles Bader, Ellen G. Bard, Elizabeth Boyle, Gwyneth Doherty, Simon Garrod, Stephen Isard, Jacqueline Kowtko, Jan Mcallister, Jim Miller, Catherine Sotillo, Henry Thompson, and Regina Weinert (1991). “The HCRC Map Task Corpus.” Technical report.
- Anne H. Anderson, and Elizabeth A. Boyle (1994). “Forms of introduction in dialogues: Their discourse contexts and communicative consequences.” *Language and Cognitive Processes*, 9:1, pp. 101–122.
- Harvey Sacks, and Emanuel A. Schegloff (1979). “Two preferences in the organization of reference to persons in conversation and their interaction.” *Everyday language: Studies in ethnomethodology*, pp. 15–21.
- Herbert H. Clark, and Deanna Wilkers-Gibbs (1986). “Referring as a collaborative process.” *Cognition*, 22:22, pp. 1–39.
- 串田秀也 (2008). 「指示者が開始する認識探索：認識と進行性のやりくり (<特集> 相互行為における言語使用：会話データを用いた研究)」 社会言語科学, 10:2, pp. 96–108.
- 田窪行則 (1989). 「名詞句のモダリティ」 仁田義雄・益岡隆志 (編) 『日本語のモダリティ』 くろしお出版 pp. 211–233.
- Douglas Bates, Martin Mächler, Ben Bolker, and Steve Walker (2015). “Fitting Linear Mixed-Effects Models Using lme4.” *Journal of Statistical Software*, 67:1, pp. 1–48.

関連 URL

『日本語地図課題対話コーパス』 <http://research.nii.ac.jp/src/MapTask.html>