D-44

対話文における双方向文脈補完を用いた言い淀み検出

中島 寛人，嶋田 和孝
九州工業大学大学院 情報創成工学専攻

1 はじめに

話し言葉には「やす，安くて」といったような，単語を途中まで言いかけて中断する言い淀みと呼ばれる現象が多く含まれている．文意に関係しない言い淀みの整形は，話し言葉を処理する際に必要不可欠である．しかし，言い淀みを人手で整形する作業は労力を要する．したがって，言い淀みの自動的な整形や検出は，話し言葉を使用する言語処理タスクの前段処理として重要なタスクとなっている．

複数人の話者が存在し，割り込みや相槌による発話の中断が発生する対話文では，個々の発話の文長が短くなってしまう傾向がある．言い淀み検出では文脈情報を利用した検出が有効だが，発話が短い状態では文脈を十分に得られないため，途切れた文脈を補完する必要がある．しかし，片方向の補完だけでは文長を安定して確保できず，前後に分散している文脈を捉えることも難しい．

そこで，本論文では発話者の前後発話に着目し，深層学習モデルに入力する前に発話の文脈を双方向に補完することで，対話文においても文脈情報を利用できる言い淀み検出を試みる．

2 提案手法

提案手法の概略を図 1 に示す．本論文では深層学習モデルとして汎用言語モデルである BERT[1] を使用し，発話を分解したトークン列に対して言い淀みのチャンク (塊) を発見する系列ラベリングモデルとしてファインチューニングすることで言い淀みの検出を行う．

この際，言い淀み検出のターゲットとなる発話 (以降，これをターゲット発話とする) について，前後双方向に文脈情報の補完を行う．ターゲット発話と同じ発話者がターゲット発話の前後で行った発話を前後文脈として採用し，BERT の特殊トークンである “[SEP]” トークンをデリミタとして結合する．このようにして作成した文を BERT に入力することで，各発話についてより広い範囲での文脈情報を与える．

正解データのラベリングには，従来の研究でも使用されてきたチャンキング形式を使用する．チャンクの始点に B-タグ，始点以外のチャンクの内側に I-タグ，チャンクの外側に O-タグを付与する IOB2 方式に従って，言い淀み部分へのラベリングを行う．なお，本研究では前後発話がもつ表層的な文脈情報のみを利用する．具体的には，結合した前後発話に相当する部分の正解ラベルの有無がターゲット発話中の言い淀み検出の際に有利にならないように，前後発話中の単語はすべてチャンクの外側として扱い，正解ラベルを秘匿化して学習する．

3 実験

対話書き起こし文のデータセットとして，CSJ：『日本語話し言葉コーパス』[2] からインタビューや自由対話など全 58 対話の転記テキストを使用する．これを用いて文脈補完の程度・部分により以下 4 種類のモデルを作成する．

- 文脈補完なし　：ターゲット発話のみ
- 前方向文脈補完：前発話+ターゲット発話

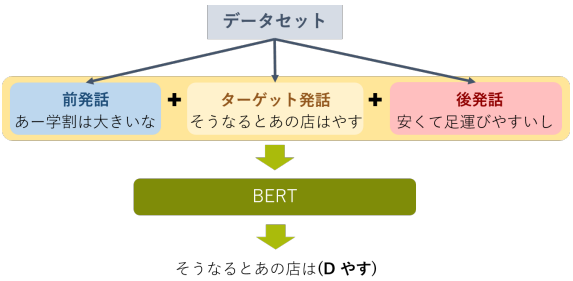

図 1: 文脈補完の流れ．前後発話を結合した後 BERT に入力し，チャンクラベリングを行う．

表 1: CSJ における検出結果．

モデル	Pre.	Rec.	F1
文脈補完なし	0.223	0.564	0.321
前方向文脈補完	0.311	0.527	0.392
後方向文脈補完	0.399	**0.585**	0.474
双方向文脈補完	**0.462**	0.537	**0.496**

- 後方向文脈補完：ターゲット発話+後発話
- 双方向文脈補完：前発話+ターゲット発話+後発話

言い淀み検出の実験は，訓練データ：テストデータが 8：1 になるようにデータセットを対話単位で分割した 9 分割交差検証によって評価する．なお文脈補完を行うモデルでは，予測の際にも前後の文脈を結合しているが，ターゲット発話部分の予測結果のみを評価する．

実験結果を表 1 に示す．双方向文脈補完モデルが Precision, F 値の各スコアにおいて他のモデルより高いスコアを得た．BERT に入力されるトークンの平均長は，文脈補完なしと双方向文脈補完との間で 16 トークンから 45 トークンと約 3 倍に増加した．また，前・後方向文脈補完モデルを含め，入力される平均トークン長が長いほど Precision, F 値が高くなる傾向があった．発話ひとつひとつの長さが短い対話文においても，双方向に文脈情報を補完することで言い淀み検出精度の向上を行うことができたと考える．

4 おわりに

本論文では，対話文の言い淀み検出において，ターゲット発話の前後発話を補完することで精度向上を試みた．今後は発話補完範囲の増加や，生成系モデルの使用など，よりターゲット発話の持つ文脈を補完できる手法に取り組む．

参考文献

[1] Jacob Devlin, Ming-Wei Chang, Kenton Lee, and Kristina Toutanova. BERT: Pre-training of deep bidirectional transformers for language understanding. In *Proceedings of NAACL, Volume 1*, pp. 4171–4186, 2019.

[2] 前川喜久雄. 『日本語話し言葉コーパス』 の概要. 日本語科学, Vol. 15, pp. 111–133, 2004.