

# コンテキストに適した英語表現学習支援のための 映像作品の字幕データのセンチメント分析と可視化

## Visualizing Sentiment Analysis on Subtitles in Dramas for Learning English Conversation Complying with Context

北畑 由莉香<sup>1</sup> 山西 良典<sup>1\*</sup>  
Yurika Kitahata<sup>1</sup> Ryosuke Yamanishi<sup>1</sup>

<sup>1</sup> 関西大学総合情報学部

<sup>1</sup> Faculty of Informatics, Kansai University

**Abstract:** This paper describes the sentiment analysis for English subtitles of dramas, aimed at enhancing the learning of conversational expressions within their contextual use. Traditional educational materials may not always encompass expressions and vocabulary encountered in daily interactions among friends, including informal and colloquial language ( occasionally, even profanities) which can be essential in making friendships. Focusing specifically on drama subtitles, we conducted a detailed sentiment analysis to visualize the context and construct of sentences where specific words are used. This paper discusses the efficacy of such visualizations in facilitating the acquisition of conversational language skills.

### 1 はじめに

急速に進む国際化やインターネットの普及に影響され、言語・文化が異なる人とのコミュニケーションの機会は増加傾向にある。特に、国際コミュニケーションの手段として世界的に使用され、影響力を拡げている英語 [Quirk 85] を用いたコミュニケーション能力の獲得は、非ネイティブ話者にとって共通の課題である。生涯にわたって様々な場面で必要とされるその能力を身に着けるため、英語能力の獲得への需要はますます高まるものと考えられる。

Computer Assisted Language Learning (CALL) は、コンピュータによる言語習得の支援を目指した研究分野 [Graham 01] であり、様々な技術や事例が報告されている [Hughes 23, Hamad 17]。英語能力の習得を支援するために開発された CALL システムはこれまでも数多く存在するが、それらの研究は単語学習 [内田 15] や構文学習 [阿部 21]、発音学習 [葉田 01]、リスニング学習 [Winke 10, 九鬼 19] を支援し、語彙や文法規則、リスニング能力などの基礎能力の習得と定着を支援するものが多い。しかし、英語でのコミュニケーションの実現場には、これらの基礎的な知識や能力だけでなく、「状況に応じて適切な表現を選択する能力」が求められる。この能力は、主に「聞く」「読む」「話す」「書

く」などの実践的コミュニケーションを繰り返し経験することによって身につけられる [山西 23] が、実践的コミュニケーションを経験するための課題も多い。非ネイティブ話者が経験していく中では、自身がコミュニケーションの中で用いた表現 (単語、熟語、フレーズ) が、その「場面」「状況」「背景」「対象」などと照らし合わせて適切か否かを学習者自身からは判断不可能である。例えば、コミュニケーションを取る相手が家族の場合、友人の場合、目上の人物の場合では、同内容であっても異なる表現を用いて会話することが容易に想定される。また、同一の相手であっても、空間 (例えば、フォーマル/インフォーマルの違い) の違いによる言葉の使い分けや文脈次第で大きく意味が変化する言葉を学ぶ必要がある。既存の学習教材の多くは、文法的に正しいとされる表現や広く一般的に使用可能な表現を学ぶうえでは有用である。一方で、口語で頻出するコミュニケーションの場面や状況、背景、対象に影響されて意味や印象が変化する表現は少なく、様々な文脈に合わせた表現を学ぶ手段が少ない現状にある。

多種多様なパーソナリティの話者が背景設定・関係性・場面状況・文脈の中でコミュニケーションを行う様子を観測可能なエンタテインメントとして映像作品 (映画・TV番組) の存在 [Birulés-Muntané 16, 植松 04] がある。映像作品で出現するセリフは、文脈に適應した言語表現の学習用教材としての活用 [Shan 21, 九鬼 19, 吉野 03, 湯舟 15] が考えられる。映像作品では字幕セ

\*連絡先： 関西大学総合情報学部  
〒 569-1095 大阪府高槻市霊仙寺町 2-1-1  
E-mail: {k940098,ryama}@kansai-u.ac.jp

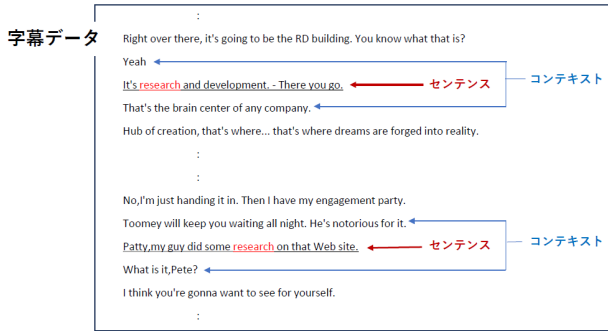


図 1: 分析センテンスと分析コンテキストの例

表 1: センチメント分析モデルの感情評価の精度確認の結果

		Model の出力ラベル		
		positive	neutral	negative
主観ラベル	positive	77	45	3
	neutral	30	287	46
	negative	1	41	109

リフをテキスト化した字幕が用意されていることが多い。この字幕を分析することで、任意の単語が「どのような文脈のどのような発話の中で使用されているのか」を可視化可能であると考えた。

本稿では、「字幕に登場する単語」と「その単語を含む発話」、「発話が行われた文脈」と「文脈と発話から感じ取れる感情」の4点に着目する。映像作品の英語字幕に対してセンチメント分析を適用して可視化可能なシステムを構築した。各感情の文脈での単語出現頻度の多寡を視認可能にすることで、文脈を考慮した英単語の学習支援の可能性を検討する。

## 2 分析手法と可視化方法

任意の単語が「どのような文脈のどのような発話の中で使用されているのか」を分析するために、まず文脈と発話を定義する。任意の単語が含まれる一文を発話文 (i.e., 分析センテンス) とし、センテンスの前後それぞれ1文を発話が行われた文脈 (i.e., 分析コンテキスト) とする。例えば、図1に示すような字幕データ (発表年:2007, imdbId:1003353, 字幕ファイル ID:3140984, 行:272-278, 506-510) で ‘research’ について分析する場合を考える。文中に ‘research’ が含まれている、3行目 “It’s research and development. - There you go.”, 8行目 “Patty, my guy did some research on that Web site.” の2文が分析センテンスとなる。このとき、2行目 “Yeah” と4行目 “That’s the brain center of any company.” は3行目に対応した分析コンテキスト、7行目 “Toomey will keep you waiting all night. He’s

表 2: 感情ラベルごとのセンチメント分析モデルの評価指標値 (小数点第二位以下四捨五入)

	再現率 (%)	適合率 (%)	F 値 (%)
positive	61.6	71.3	66.1
neutral	79.1	76.9	78.0
negative	72.2	69.0	70.6

notorious for it.” と9行目 “What is it, Pete?” は8行目の分析コンテキストとなる。

### 2.1 センチメント分析

分析センテンスと分析コンテキストそれぞれにセンチメント分析を行って、それぞれの極性を得る。得られる分析センテンスと分析コンテキストそれぞれの極性を組み合わせによって、単語の使用傾向を表現可能にする。

Loureiro らによって作成された、約12,400万の tweet によって英語のセンチメント分析用に調整された RoBERTa モデル [Loureiro 22] (以降センチメント分析モデルと呼称)<sup>1</sup> を使用する。センチメント分析モデルによって、入力されたテキストに対して positive, neutral, negative の3種類の感情それぞれの確信度が出力される。本稿では、確信度が最も高く出力された感情ラベルを入力テキストの感情として扱う。

### 2.2 センチメント分析モデルの精度確認

tweet によって学習・構築されたセンチメント分析モデルによって字幕のセンチメント分析が可能であるかについて事前検証を行った。既存研究 [徐 22] を参考にし、2022年8月時点の Netflix で「オリジナル言語」<sup>2</sup> が「英語」である映像作品の字幕から1,000会話文を用意した。この会話文集合から、ランダムに639文を抜き出して評価用のデータセットを作成した。評価用データセットに対して、第1著者の主観評価で各文が3種の感情のうちどの感情を示すかラベル付けを行った。一方で、これらの文に対してセンチメント分析モデルを用いた極性評価も行った。

表1に、主観とセンチメント分析モデルの評価結果の比較を示す。表2に、感情ラベルごとのセンチメント分析モデルの評価指標値 (再現率, 適合率, F 値) を示す。文献 [Loureiro 22] では、TweetEval ベンチマーク [Barbieri 20] を用いたモデルの精度評価が行われており、センチメント分析タスクに対し、マクロ平均再現率が73.7%であった。ここで、クラス数と*i*番目の

<sup>1</sup><https://huggingface.co/cardiffnlp/twitter-roberta-base-sentiment-latest> (2023年12月1日確認)

<sup>2</sup>その映像作品が制作された言語



図 2: 単語入力画面



図 3: グラフ・用例提示画面

クラスの再現率をそれぞれを  $n$  と  $(x_{Recall})_i$  とすると  
 マクロ平均再現率は下式で求められる。

$$x_i = \frac{1}{n} \sum_{i=1}^n (x_{Recall})_i. \quad (1)$$

表 2 と数式 1 から、本稿でのマクロ平均再現率は  
 70.9%となり、文献 [Loureiro 22] と同程度の値が得ら  
 れた。このことから、本稿で用いるセンチメント分析  
 モデルは、字幕の会話文に対しても妥当な感情推定結  
 果を出力可能であることが示唆された。

## 2.3 可視化方法

提案システムは、図 2 に示す「単語入力画面」と図  
 3 に示す「グラフ・用例提示画面」の 2 画面から構成さ  
 れる。ユーザの入力や操作に基づいて 2 画面を切り替  
 え、単語入力画面でユーザが入力した単語のセンチメ  
 ント分析の結果を示すグラフと用例を提示する。

単語入力画面(図 2)では、画面上部中央の入力フォー  
 ムに英単語を入力し、入力フォーム右下の「check」ボ  
 タンを押下する。フォームへ入力された英単語(品詞  
 別)のセンチメント分析結果のデータがシステム内に  
 存在するか走査する。走査の結果、結果データが存在  
 した品詞に対応するボタンのみが有効化される。ユー  
 ザが、有効化された任意の品詞ボタンを押下するこ  
 とで、グラフ・用例提示画面に遷移する。

単語入力画面で有効化された任意の品詞ボタンを押  
 下することで、グラフ・用例提示画面(図 3)に遷移す  
 る。グラフ・用例提示画面ではまず、押下された品詞  
 の単語が登場するセンテンスとそのコンテキストに対  
 するセンチメント分析結果を可視化したグラフ(図 3 中  
 左)とグラフ下部に 9 つのボタンが表示される。ここ  
 で、9 つのボタンはコンテキスト(C)とセンテンス(S)  
 の極性(positive:p, neutral:-, negative:n) 組み合わせに  
 対応する。ユーザがボタンを押下すると、ボタンが示  
 す極性の組み合わせに対応した用例がランダムに 3 件  
 選出され、表示される。

## 3 分析対象とする字幕データ

本稿では、Lison らが構築した、OpenSubtitle コー  
 パス [Lison 16, Tiedemann 16] を対象として分析する。  
 OpenSubtitles<sup>3</sup>は、ユーザーによって作成・投稿された  
 映画や TV 番組の字幕データを集積し、60 以上の言語、  
 300 万以上の字幕データから成る大規模な字幕データ  
 ベースを提供する Web サイトである。OpenSubtitles  
 コーパスは、OpenSubtitles で提供されている字幕デー  
 タに対して、字幕の誤りの訂正、文章の整理などが処  
 理されている。さらに、識別子などのメタデータを付  
 与した 62 言語の字幕データを、任意の 2 言語の文単位  
 の対訳データの形式で得ることができるパラレルコー  
 パスとなっている。

OpenSubtitles コーパスはデータを xces 形式や tmx  
 形式等で提供しており、本稿では HuggingFace 上<sup>4</sup>の  
 英語字幕と日本語字幕の対訳データ 810,000 件が含ま  
 れるファイルを対象とする。表 3 に、データセットの  
 一部を例として示す。以降に示す字幕データは、全て  
 このファイルから引用する。

表 3 に示すように、データセットは ID、メタ情報、  
 英語字幕文と日本語字幕文の対訳の 3 つの情報で構成  
 される。このうち、英語字幕文と日本語字幕文の対訳  
 から英語字幕文を抽出した。次に、英語字幕文に含  
 まれるすべての英単語の種類・数の把握するため、英  
 語字幕文の分かち書きと各単語の品詞の取得を行った。  
 同一単語であっても、異なる品詞タグが付与されたも  
 のは別単語として扱い、単語と品詞のペアが重複を許  
 さないようにしたところ 80,333 種類の英単語が得られ  
 た。このうち、アルファベット以外の文字列(数値、記  
 号、英語以外の言語の文字)が含まれておらず、出現  
 回数が上位 10% より多い単語を抽出した。さらに、こ  
 れらの単語に表 4 に示す 11 種の品詞で絞り込みを行っ  
 て 8,713 単語を分析対象および可視化用の単語として  
 準備した。

<sup>3</sup><https://www.opensubtitles.org/>(2023 年 12 月 1 日確認)

<sup>4</sup>[https://huggingface.co/datasets/open\\_subtitles](https://huggingface.co/datasets/open_subtitles)(2023  
 年 12 月 1 日確認)

表 3: OpenSubtitles データセットの一部. ID は字幕固有の番号, メタ情報には公開年 (year), IMDb 識別子 (imdbId), 字幕ファイルの ID(subtitleId, 英語と日本語それぞれ), 訳文がそれぞれ字幕ファイルの何行目か (sentenceIds, 英語と日本語それぞれ) の情報を含む. 対訳は英語と日本語文の対訳.

ID	メタ情報	対訳
0	"{'year': 1920, 'imdbId': 10323, 'subtitleId': {'en': 5777654, 'ja': 5996100}, 'sentenceIds': {'en': array([1], dtype=uint32), 'ja': array([1], dtype=uint32)}}"	"{'en': '""THE CABINET OF DR. CALIGARI""', 'ja': 'カリガリ博士の小屋'"}
:	:	:
102	"{'year': 1924, 'imdbId': 15324, 'subtitleId': {'en': 5530518, 'ja': 5620547}, 'sentenceIds': {'en': array([2, 3], dtype=uint32), 'ja': array([2], dtype=uint32)}}"	"{'en': '""Don't try to do two things at once and expect to do justice to both. This is the story of a boy who tried it.""', 'ja': '""二兎を追う者は一兎をも得ず""'"}
103	"{'year': 1924, 'imdbId': 15324, 'subtitleId': {'en': 5530518, 'ja': 5620547}, 'sentenceIds': {'en': array([4], dtype=uint32), 'ja': array([3, 4, 5], dtype=uint32)}}"	"{'en': 'While employed as a moving picture operator is a small town... theater he was also studying to be a detective.', 'ja': 'しかし この物語の主人公はあえてそれに挑戦した 彼は町の映画館に勤める 映写技師であり また同時に 探偵修業中の 身でもあった'"}

表 4: 使用する品詞タグ (11 種) と対応する品詞名の一覧

品詞タグ	対応する品詞名
NN	名詞
NNS	名詞の複数形
JJ	形容詞
JJR	形容詞の比較級
JJS	形容詞の最上級
VB	動詞の原形
VBD	動詞の過去形
VBG	動名詞または現在分詞
VCN	動詞の過去分詞形
VBZ	動詞の三人称単数以外の現在形
VBZ	動詞の三人称単数の現在形

## 4 字幕データの分析と可視化

3 節で用意した単語集合に対して, 単語のセンテンスとそのコンテキストに対しセンチメント分析を行った. 紙面の都合上, 興味深い結果が得られた damn(動詞の原形 (VB)) と incredible(形容詞 (JJ)) の可視化結果を示して考察する. 図 4 と図 5 に, それぞれ damn(動詞の原形 (VB)) と incredible(形容詞 (JJ)) が含まれるセンテンスとコンテキストのセンチメント分析の可視化結果を示す. 以下, コンテキストとセンテンスの組み合わせを示すために, 提案システムでのボタンの表示に合わせた (C:p, S:p) のような表記も用いる.

### 4.1 damn(動詞の原形 (VB)) のセンチメント分析結果

'damn' は, 一般的には「永遠に罰する, 地獄に落とす, 酷評する, けなす, ののしる, のろう, ちくしょう!」といった意味をもち<sup>5</sup>, negative な極性の単語と考えられる. 同図から, 'damn' は negative な会話の流れ

<sup>5</sup><https://ejje.weblio.jp/content/damn>(2023 年 11 月 24 日閲覧)

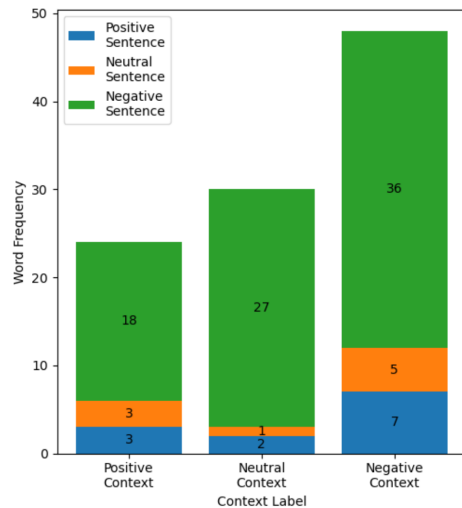


図 4: damn の用法の可視化結果

で登場することが多く, negative な発話で用いられる. 一方で, 'damn' は negative なコンテキストやセンテンスのみで用いられるわけではなく, 'damn' という単語そのものが持つ, negative なイメージと離れた positive なセンテンスやコンテキストでも用いられる.

表 5 にそれぞれ, damn が含まれるセンテンスで, コンテキストがそれぞれ positive, neutral, negative であるものの例を示す. 表中, センテンスの極性が negative((C:p, S:n), (C:-, S:n), (C:n, S:n)), もしくは極性の組み合わせが (C:n, S:-) となっている例では, 'damn' は主に非難や罵りの意味で用いられている. 一方で, コンテキストが positive, 極性の組み合わせが (C:-, S:p), (C:-, S:-), (C:n, S:p) となっている例では, 'damn' は後続する言葉を強調する目的で用いられており, 強い肯定や, 賛辞の強調を示している.

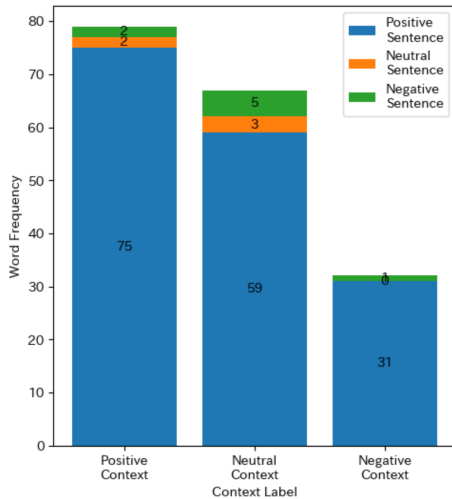


図 5: incredible の単語特徴可視化結果

negative なセンテンスやコンテキストに登場した場合、‘damn’ は非難や罵りなどの negative な意味で用いられており、‘damn’ という単語が一般的にもつ negative なイメージと一致する。一方で、positive あるいは neutral なセンテンスやコンテキストで用いられる場合、‘damn’ は強い肯定や賛辞の強調として用いられており、一般的なイメージとは乖離した用法であると考えられる。辞書の意味とは異なる用法をシステムの出力結果から一瞥して理解可能であると考ええる。

## 4.2 incredible(形容詞 (JJ)) のセンチメント分析結果

‘incredible’ は、一般的には「信じられない、信用できない」と「驚くべき、非常な、途方もない」といった意味をもつ単語<sup>6</sup>である。図 5 より、‘incredible’ は positive なコンテキスト、もしくはコンテキストを問わず positive なセンテンスで登場することが多く、neutral, negative なセンテンスへの登場は極端に少ない。

表 6 に、incredible が含まれるセンテンスで、コンテキストの極性がそれぞれ positive, neutral, negative であるものの一例を示す。同表から、‘incredible’ という単語が示す 2 種類の意味（「信じられない、信用できない」と「驚くべき、非常な、途方もない」）は、コンテキストやセンテンスの極性によらず、どちらの極性でも用いられることがわかる。例えば表 6 中、(C:p, S:p) のセンテンス “you did it darling, that was incredible.” と、(C:n, S:n) のセンテンスである “incredible. so sorry.” で、incredible は「信じられない」を示す言葉として用

いられている。これらの 2 例文で “incredible” は異なるニュアンスを示している。(C:p, S:p) の例では、「(予想していたより良い結果を得られ)信じられない」という positive な反応である一方、(C:n, S:n) の例では「(期待外れの出来事が起き)信じられない」という negative な反応として ‘incredible’ が使われている。同一の単語であるにもかかわらず、異なる意味・感情として捉えられる龍としてセンテンス中の他の表現による影響が考えられる。(C:p, S:p) の例文中には ‘darling’ という親しい人物への呼びかけの言葉があり、コンテキストには “nice job sweetheart.” という恋人を褒める positive な文が含まれている。一方で、(C:n, S:n) の例文中には、“so sorry” という謝罪の言葉が含まれている。これらのことから、‘incredible’ は使用されるコンテキスト・センテンス中に登場する表現の極性に影響を受ける単語であることが理解できる。

以上の考察をまとめると、‘incredible’ は登場するコンテキスト・センテンス中に登場する表現の極性に影響を受け、示す意味の極性を変化させる単語である。また、‘incredible’ は、positive なコンテキストやコンテキストを問わず positive なセンテンスで用いられることが多く、一般的に positive な意味合いで用いられる単語であることがわかる。

## 5 おわりに

本研究の最終目標は、映像作品の英語字幕を英語学習に活用することで、英語表現学習における既存の課題を解決することにある。本稿では、「字幕に登場する単語」と「その単語を含む発話」、「発話が行われた文脈」と「文脈と発話から感じ取れる感情」の 4 点に着目し、映像作品の字幕に登場する任意の単語が「どのような文脈のどのような発話の中で使用されているのか」を可視化した。出現傾向を示す積み立てグラフと具体的な用例を合わせて提示することで、任意の単語の主な使われ方、登場頻度は少ないが特徴的な使われ方の双方を理解することができる可能性が示唆された。今後は、日本語の対訳データを利用したコンテキストに応じた単語の使い分けを分析するほか、提案システムによる学習効果の検証も行う。

## 謝辞

本稿では、分析および表 3、表 5、表 6 での例示に OpenSubtitle コーパス [Lison 16, Tiedemann 16] のデータ ([https://huggingface.co/datasets/open\\_subtitles](https://huggingface.co/datasets/open_subtitles)) を引用させて頂いた。本稿の執筆にあたって、立命館大学情報理工学部の Jeremy White 教授に助言を頂いた。記して謝意を表す。

<sup>6</sup><https://ejje.weblio.jp/content/incredible>(2023 年 11 月 24 日閲覧)

## 参考文献

- [Barbieri 20] Barbieri, F., et al.: TweetEval: Unified Benchmark and Comparative Evaluation for Tweet Classification, in *Findings of the Association for Computational Linguistics: EMNLP 2020*, pp. 1644 – 1650 (2020)
- [Birulés-Muntané 16] Birulés-Muntané, J. and Soto-Faraco, S.: Watching Subtitled Films Can Help Learning Foreign Languages, *PLOS ONE*, Vol. 11, No. 6, pp. 1 – 10 (2016)
- [Graham 01] Graham, D.: CALL (computer assisted languagelearning), <https://web-archive.southampton.ac.uk/www.llas.ac.uk/resources/gpg/61.html> (2001)
- [Hamad 17] Hamad, M.: Using WhatsApp to Enhance Students’ Learning of English Language “Experience to Share”, *Higher Education Studies*, Vol. 7, p. 74 (2017)
- [Hughes 23] Hughes, L.: Effects of Dynamically Computer-Mediated Communication and ‘Pseudo’ Communication on L2 Learning, *CALL-EJ*, Vol. 24, pp. 106–137 (2023)
- [Lison 16] Lison, P. and Tiedemann, J.: OpenSubtitles2016: Extracting Large Parallel Corpora from Movie and TV Subtitles, in *Proc. the Tenth Intl’ Conf. on Language Resources and Evaluation*, pp. 923 – 929 (2016)
- [Loureiro 22] Loureiro, D., et al.: TimeLMs: Diachronic Language Models from Twitter, *CoRR*, Vol. abs/2202.03829, (2022)
- [Quirk 85] Quirk, R. and H.G.Widdowson, : *English in the World: Teaching and Learning the Language and Literatures*, Cambridge University Press in association with The British Council (1985)
- [Shan 21] Shan, J., et al.: アニメの部分会話の難易度の推定による日本語のリスニング練習支援, *芸術科学会論文誌*, Vol. 20, No. 2, pp. 108–119 (2021)
- [Tiedemann 16] Tiedemann, J.: Finding Alternative Translations in a Large Corpus of Movie Subtitle, in *Proc. the Tenth Intl’ Conf. on Language Resources and Evaluation*, pp. 3518 – 3522 (2016)
- [Winke 10] Winke, P., et al.: The Effects of Captioning Videos Used for Foreign Language Listening Activities, *Language Learning and Technology*, Vol. 14, (2010)
- [阿部 21] 阿部 真由美, 向後 千春: 英語多聴学習での拡散的好奇心による学習行動と学習意欲への影響, *日本教育工学会論文誌*, Vol. 45, No. 1, pp. 127 – 134 (2021)
- [植松 04] 植松 茂男: DVD 映画教材利用時の英語字幕が英語学習に与える影響について, *メディア教育研究*, Vol. 1, No. 1, pp. 107 – 114 (2004)
- [内田 15] 内田 千尋, 橋田 光代: TypeSinger: 歌うように“弾く”歌詞タイピングゲーム, *エンタテインメントコンピューティングシンポジウム 2015 論文集*, Vol. 2015, pp. 247–250 (2015)
- [九鬼 19] 九鬼 慧太 他: 映画を用いた英単語学習における話者追従字幕がコンテンツの記憶保持に与える影響, *電子情報通信学会技術研究報告*, Vol. 119, No. 190, pp. 79 – 84 (2019)
- [徐 22] 徐 敏徹: ;研究ノート;OTT サービスを利用したパラレルコーパスの構築方法, *京都大学言語学研究*, Vol. 41, (2022)
- [葉田 01] 葉田 善章 他: Viclle: 会話映像の添削による語学学習支援システム, *情報処理学会論文誌*, Vol. 42, No. 6, pp. 1412 – 1423 (2001)
- [山西 23] 山西 良典 他: 英会話はハートがすべてなんで!? キャンユースピークイングリッシュ?, *エンタテインメントコンピューティングシンポジウム 2023 論文集*, pp. 201 – 207 (2023)
- [湯舟 15] 湯舟 英一, 田淵 龍二: 映画場面検索サイト Seleaf を利用した授業の学習効果, *Language education & technology*, No. 52, pp. 389 – 410 (2015)
- [吉野 03] 吉野 志保: 英語学習に効果的な字幕提示タイミングの検討 (;特集; 第二言語学習とその支援に関する教育工学研究), *日本教育工学雑誌*, Vol. 27, No. 3, pp. 237 – 246 (2003)

表 5: damn の含まれるセンテンスの一例

(C: コンテキストの極性, S: センテンスの極性)	センテンス (括弧内はデータセットから抽出した, センテンスに対応している日本語訳)	センテンスと対応するコンテキスト (上: センテンスの前の文, 下: センテンスの後の文)
(C:p, S:p)	feels so warm, i can't help it. i want you. <b>damn</b> . (温かい体だ 君が欲しくてたまらない)	don't you know i couldn't lose you? (君が大切だから...)
		i love the way you smell. i love you, and i always have. i want to tell you... (ずっと好きだったと 打ち明けたいよ)
(C:p, S:-)	<b>damn</b> the meaning, i'll memorize it. (意味はいいんだ 丸暗記するから...)	our director has unceasingly served... 'served' means... (学長は ここで奇跡を成し遂げた この言葉の意味は...)
		chatur was the introductory speaker at the teachers' day function. (チャートルは 先生の日の スピーチに選ばれた)
(C:p, S:n)	stay in there, <b>damn it!</b> (そこから出てくるな!)	don't you dare follow us! (2度と後を追ってくるな!)
		thanks again for the money, clothes and everything. (お金から服から 何から何までありがとう)
(C:-, S:p)	you're <b>damn right</b> you're glad. (あなたの心遣いが嬉しいわ)	well, i'm glad i'm not working for you. (嬉しいよ 君が僕のために働いてなくて)
		so, do i have to smack you in the mouth, lieutenant? (それで 僕は口移しで 薬を飲ませればいいのか?)
(C:-, S:-)	doc, what's going on here, <b>damn it!</b> everybody. y'all (博士, 何が起こってる?)	30 minutes to decontamination. (汚染浄化まで 30分)
		get your stuff and let's go! (みんな急ぐんだ さあ!)
(C:-, S:n)	see how polacks have worked on me? <b>damn the basterds!</b> (俺は地主どもにやられて こんな有様だからな)	i've sent her for some home brew. (酒を買いに出したんだ)
		you'd better tell me. (聞きたい事があるんだけど)
(C:n, S:p)	it'd <b>damn well better</b> . (一見させるさ)	can it copy dvds? (見れるか?)
		that's not how it opens! (ここはだめだ ティビ)
(C:n, S:-)	i'm a cop, <b>damn it</b> . i'll show you my identification. (私は警察だ ID も持っている)	why should i trust you? you could be the maniac. (信用できないわ あなたが異常者かもしれない)
		here. look, can you see it? (ほら 見えるか?)
(C:n, S:n)	<b>damn</b> the day i met you, you <b>basterd!</b> (あんたと結婚したのが 間違いよ)	go! and may you rot there! (いいわゾーンに行行って 死になさい)
		god himself has cursed you by giving you this child! (だから呪われた子供が 生まれてー)

表 6: incredible の含まれるセンテンスの一例

(C: コンテキストの極性, S: センテンスの極性)	センテンス (括弧内はデータセットから抽出した, センテンスに対応している日本語訳)	センテンスと対応するコンテキスト (上: センテンスの前の文, 下: センテンスの後の文)
(C:p, S:p)	you did it darling. that was <b>incredible</b> . (良くやったわ 信じられない)	katniss everdeen, the girl on fire! (カットニス・エヴァディーン ”炎の少女”でした！) nice job sweetheart. (良くやったな)
(C:p, S:-)	<b>incredible?</b> what is? (”ものすごいこと”って どのようなこと?)	what would you do, if i became incredible? (そしたらどうする? あたしが ものすごいことになっちゃったら) you know, everything! (だから”何もかも”だよ)
(C:p, S:n)	i don't like surprises. i have an <b>incredible</b> ... (いやよ 驚くなんて)	what? no. (何?) trust me, you're gonna love this. (信じて きっと気に入るよ)
(C:-, S:p)	being with him was an <b>incredible</b> adventure. (彼という事は 素晴らしい冒険だった)	he saw things between lines of text that nobody else could. (書物の行間を 読む人よ 誰も出来ないわ) so you did love him? (彼を愛していたと?)
(C:-, S:-)	temperatures are rising with <b>incredible</b> velocity in hot zones across the globe. (全地上に渡って 地温が 急上昇しています)	it's not just yellowstone, sir. (イエローストーンだけではないのです) are you absolutely sure about this, adrian? (確信を持っているのか? エイドリアン)
(C:-, S:n)	what's <b>incredible</b> , though, is how spectacularly ignorant he is (「驚くべき事に特定の分野に ついては全く無知なのだ」)	! sherlock sees through everything and everyone in seconds. (「シャーロックは一瞬で どこの誰だかを見抜くが」) about some things. hang on, i didn't mean that... (待ってくれ 変な意味じゃ...)
(C:n, S:p)	what is <b>incredible</b> is i just finished the mazethis morning, and i'm not gonnateach her how to run it until hour from now. (何がすごいかって、今朝この迷路を 完成したばかり なんだよ 数時間後に道を教える 予定だったんだ)	i'm sorry. how is a rat running through a bloody maze so incredible? (迷路をねずみが抜けれるのが なぜそんなにすごいって言うんだ?) s so so you you sent herto the future. (君がこいつを未来へ 送ったって言うのか?)
(C:n, S:-)	なし	なし
(C:n, S:n)	<b>incredible</b> . so sorry. (信じられん ー 申し訳ございません)	this is not a flee market. (ここは蚤の市じゃない) i have to make a phone call. excuse me. (ちょっと電話してくる、失礼)