HOKUGA 北海学園学術情報リポジトリ
学校法人北海学園
北海学園大学
北海商科大学

タイトル	TV ドラマ・コーパス情報に基づく英語基礎口語定型表現リストと教材の作成
著者	田中，洋也；TANAKA，Hiroya
引用	北海学園大学人文論集(72)：41-69
発行日	2022-03-31

Hokkai-Gakuen Organization of Knowledge Ubiquitous through Gaining Archives.

TV ドラマ・コーパス情報に基づく英語基礎口語定型表現リストと教材の作成

田　中　洋　也

要　　旨

本稿では，アメリカでテレビ放送されたシチュエーション・コメディ（situation comedy，以下，sitcom）の脚本コーパス情報に基づいた英語基礎口語定型表現リスト（Basic Spoken Formula List）およびリストに基づいた教材の作成について報告する。リストの作成にあたり，2005 年から 2016 年までに放送された 18 作品のドラマ脚本から構築した約 525 万語のコーパス（American Sitcom Corpus, ASC）を用いた。コーパスから抽出した 2 語から 5 語の語彙連結（lexical bundle），15,415 項目を対象に，本研究の目的に合わせて設定した抽出規準に基づき，最終的には 2,585 項目の基礎口語定型表現（Basic Spoken Formula）による英語基礎口語定型表現リストを作成した。さらに，その項目について英語母語話者である外国語教育研究者の協力を得て，American Sitcom Corpus を活用した英語例文による教材を作成した。本稿では，これまでの定型表現リスト研究との関連から，本基礎口語定型表現リストおよび例文による教材の作成過程の各段階について報告し，今後のリストと例文教材の活用に向けた課題を整理する。

1．研究目的

本研究は，英語学習者が初級段階から用いることができる基礎口語定型表現リストと教材の作成，作成したリストと教材の有用性の検証を目的と

している。本稿では，英語定型表現研究の内，定型表現リストを作成した研究を概観し，本研究における基礎口語定型表現リストおよび教材の作成過程について報告する。

外国語学習において語彙知識は，その中心を成すものであり，これまでも様々な単語項目によるリストが作成されている(例, Browne, 2014; West, 1953)。日本人英語学習者の学習ニーズを踏まえた語彙リストを見ても，学術的，教育的，様々な目的で作成された前例が数多くある（例，大学英語教育学会基本語改訂特別委員会，2016；東京外国語大学投野由紀夫研究室, 2020)。また, 語彙リストとは別に, Longman Dictionary of Contemporary English（2009）の Communication 3000 のようにコーパスの頻度情報に基づいた辞書内での学習重要度の指標を明示したものがある。さらに，根拠は明示されないものの，ジーニアス英和辞典第5版（南出，2014)，ウィズダム英和辞典第4版（井上・赤野，2018）のように学習段階（中学，高校，大学生・社会人）の指標が用いられている例もあり，その指標は学習語彙リストと同様の機能をすると考えられる。

一方，2語以上から成る定型表現に目を向けると，リストの作成や利用の目的が多岐に渡ることも影響しているのか, その例は必ずしも多くない。学習用の教材では，日本では伝統的に「熟語」として多くの教材，書籍が存在している。「熟語」に含まれる言語要素には，動詞句（例，take the place of —), 名詞句(例, academic achievement, the wide spread use of —)，副詞句（例，as soon as possible)，句動詞（例，work on)，構文（例，what if —)，決まり文句（例，that's a shame)，コロケーション（例，a piece of cake）など様々な項目が含まれる。こうした市販教材には，いくつかの課題が考えられる。一つは，前述した「熟語」に含まれる項目が多岐に渡る点である。学習者は，示された項目が，どのような場面や状況で必要となるのかを明確に区別するのが難しい状況で，入学試験，特定の英語技能試験の対策などを目的として，これらの複数語句を学習することになる。また，見出し項目の選出根拠が明確に提示されていない教材が多いことも，学習の利便性に関わる課題であると指摘できる。一方，コーパスを根拠と

して，学術的に作成された定型表現リストにも様々なものがある（例，Ackermann & Chen, 2013, Martinez & Schmitt, 2012; Shin & Nation, 2007, Simpson-Vlach & Ellis, 2010）。これらについては後節で本研究との位置付けから概観することとする。

2．定型言語と定型表現リスト

定型言語（formulaic language）の重要性については，これまで著者も報告しているが（田中，2015，2020，2021），本研究との関係から，下記に改めて整理する。

2.1　定型表現研究と本研究における定型表現

文字で表記した際にスペースで区切られる単語が複数のまとまりとなって，一定の意味や機能を持つ場合，もっとも包括的に用いられる用語としては「定型言語」（formulaic language）が挙げられる。定型言語の下位分類となる用語としては，相互に重なりがあるものの，主に下記の用語が用いられる。

定型表現（formulaic sequence / formulaic expression）
複数語句単位（multi-word unit）
語彙連結（lexical bundle）
イディオム（idiom）
句動詞（phrasal verb）
コロケーション（collocation）
慣用表現（conventional expression）
決まり文句（formula）
語用論的定型表現（pragmatic routine）

このうち，定型表現（formulaic sequence）は，Wray（2002）により，「連

続的であれ非連続的であれ，あらかじめ組み立てられている，または組み立てられているように思われる，語や他の意味を持つ要素によるつながりであり，文法によって生成，分析されるものというよりも，ひとつのまとまりとして記憶に貯蔵され，言語使用の際に想起されるもの」(p. 8）と定義されている。また，その観察可能な特徴として，(1）頻出で親密度が高いこと（例，have a nice day)，(2）意味的に不透明で，形式的に不規則であること（例，bullet point)，(3）産出や理解が容易であること（例，happy birthday)，(4）１語より長いまとまりであること（例，Before I answer that, can I just —)，(5）意味的または語用論的な付与があること（例，break a leg)，(6）話者の集団的アイデンティティーを伝えること（例，And I'm so like.)，(7) 形として定着していること (例，the text of the Quran, see Saleem, 2015）などが挙げられている（Wray, 2017)。定型表現は，これらの特徴をすべて満たさなければならないものではなく，いくつかの特徴を持ち合わせたものとして解釈される。その他の用語も，個別にその様態や目的を表し，上記７つの特徴のいずれかを満たすものと言える。例えば，イディオムは，慣用的に用いられ意味的に不透明な複数語句（例，kick the bucket）であるため，少なくとも上記の（2)，(4)，(5）を満たす。

また，口語に特徴的な語用論的定型表現は，Bardovi-Harlig（2012）によって次の特徴を持つとされる。

(a）少なくとも２つの形態素からなる，
(b）音韻的にまとまっている，すなわち流暢に調音され躊躇がない，
(c）繰り返しいつも同じ形式で用いられる，
(d）状況依存的である，
(e）言語共同体で広く用いられている。(p. 208)

語用論的定型表現は，Wray（2017）の（2）以外の全ての特徴を満たすもので，より言語機能との結びつきがあるものを指す。

本研究では，外国語学習環境で英語を学ぶ学習者を対象として，頻繁に用いられる基礎口語定型表現のリストの作成を目的としているため，コーパスとして用いる，アメリカ英語を基準としながらも，Wray（2017）の（1），（3），（4），（5），（7）の条件を満たす複数語句単位の抽出を目指す。そのため，結果としてイディオムのように意味的に不透明なものが含まれることを妨げないが，その特徴を満たすものを抽出すること自体を目的とはしない。定型表現を表す英語の用語としては，formula(s)を用いることとした。さらに，後述する学術目的で作成された定型表現リスト，Academic Formulas List（Simpson-Vlach & Ellis, 2010）と使用の目的を区別するために，リストは，「基礎口語定型表現リスト」（Basic Spoken Formula List）と名付けることにした。

2.2　言語学習者にとっての定型表現の重要性

定型表現知識とその運用能力が学習者にとって重要である理由は，これまでも指摘している（田中，2015，2020，2021）。要点をまとめると，言語使用に定型表現が占める割合が高いこと（例，Biber et al., 1999; Erman & Warren, 2000; Foster, 2001），定型表現を用いる学習者が言語知識測定などで高く評価されること（例，Hsu & Chiu, 2008; Keshavarz & Salimi, 2007），その使用により，多く慣用的発話行為が実現されること（Schmitt, 2010）が挙げられる。

本研究の目的である基礎口語定型表現リストおよび教材に関しては，下記の Schmitt and Schmitt（2020）による定型表現の有用性のまとめが最も関連性が高いと考えられる。ここでは，具体例の記述例を絞って，下記に示す。

（1）考えや意味を表現する（例，Too many cooks spoil the soup. により一般的な信念や助言を表す）。
（2）言語機能を実現する（例，I'm really sorry about X. で謝罪を行う，I'm sorry to bother you, but X —. でポライト

ネスを表現する)。

(3) 会話を維持する（例，How are you? See you later. などで会話やりとりの段階を知らせる)。

(4) 社会的相互作用を促進する社交的表現を提供する（例，Nice weather today. などの意味のある情報ではなく，会話を円滑に進めるために相手への関心を示す)。

(5) 談話構成を標示する（例，口語では，on the other hand, as I was saying など)。

(6) 専門的定型表現を提供する（例，管制官の「着陸許可」を明示するための固定表現としての，Cleared to land.)。

(7) 言語使用を展開するための構成要素（building blocks）を提供する（例，That's an interesting point. は相手の意見への同意としても対照的意見を述べる前の合図としても用いられる)。

(8) 語用論的価値を表現する（例，I don't know if X は命題について不確かであることを表す)。

(9) 流暢性を促進する（例，より速く読む，より流暢に話すなどの流暢性)。

(10) 意味を形成，定義，増大する（例，border は単独では「あるものの縁や端」を意味するが，border on では，His passion for self-improvement bordered on the pathological. のように，その使用の75%が否定的な比喩表現として用いられる)。

(Schmitt & Schmitt, 2020, pp. 86-88)

定型表現自体は，話し言葉にも書き言葉にも大きな割合で存在していると指摘されているが（Biber et al., 1999; Erman & Warren, 2000; Foster, 2001)，上記10の定型表現の有用性は，口頭での対人コミュニケーションおよび語用論的言語機能に特に関わりが強い。本研究で目指す，定型表現

リストと教材も同様の観点に立っている。特に，外国語学習環境にあり，教科書以外での話し言葉インプットの量や質の課題を抱える学習者に対して，受容的，産出的言語活動における流暢さを高めるために有効な定型表現の抽出を目指す。

2.3　定型表現研究・定型表現リスト研究

八木・井上（2013）によれば，教育学的立場による定型表現の研究は，神田・南田（1909）による「英和雙解熟語大辞典」（A Dictionary of English Phrases）にまで遡ることができる。その後，辞書の中で定型表現を記述する試みが国内外で続くことになる（例，Benson et al., 1986, 1997, 2010; 勝俣, 1939, 1958, 1995; Mackin & Cowie, 1983）。また近年は，辞書ではないが，大規模なコーパスを基にして，語彙文法（lexico-grammar）の観点から，定型表現の一形態である語彙連結（lexical-bundle）を含めて，文法を記述しようとする試みもある（Biber et al, 1999）。

本節では，本研究と関連のある 4 つの定型表現リスト研究について概観する。定型表現リストは，そのリストを作成する目的によって項目が大きく異なっている。Simpson-Vlach and Ellis（2010）では，学術場面における話し言葉，書き言葉の定型表現リストである，「学術定型表現リスト」（Academic Formulas List）を作成した。定型表現の抽出には，話し言葉，書き言葉それぞれ約 210 万語のコーパスが用いられた。コーパスから，100 万語あたり 10 回以上の頻度で用いれた，3 語から 5 語の単位，約 14,000 項目を抽出した。その後，学術場面，非学術場面コーパスを比較した対数尤度，サブコーパスとして用いられた学術分野における分布範囲（range）による取捨選択を経て，話し言葉 979 項目，書き言葉 712 項目の定型表現を抽出した。最終的には，話し言葉，書き言葉コーパスともに 100 万語あたり 10 回以上の頻度で，全 9 種類のサブコーパスのうち，6 以上に分布する主要学術定型表現 207 項目（Core AFL）を同定した。教育上の価値を基準とした FWT（Formula Worth Teaching）による上位 10 項目は，1. in terms of，2. at the same time，3. from the point of view，4. in

order to，5. as well as，6. part of the，7. the fact that，8. in other words，9. the point of view of，10. there is a のようになっている。さらに定型表現の各項目は，その機能により，指示表現（referential expression），態度表現（stance expression），談話構成表現（discourse organizing expression）に分類された。例えば，in terms of は，指示表現のうち，特性の詳述（specification of attribute）であり，実体のない構成属性を表す語（intangible framing attribute）に分類されている。

同じく，学術場面に焦点を当てて連語リスト（collocation list）を作成した研究に，Ackermann and Chen（2013）がある。彼らは，2,500万語からなる Pearson International Corpus of Academic English（PICAE）の学術的書き言葉コーパスを用いて，学術連語リスト（Academic Collocation List）を作成した。項目の選定は，４つの段階を経て行われた。最初は，コンピューターによる連語候補の抽出で，これは，内容語±３の範囲での連語の抽出，一般的な内容語の除外，機能語の除外の過程を経て，約130,000項目の候補が残された。第二段階では，学際的に用いられる項目を選別するため，頻度などの統計情報，品詞タグ付与と品詞の組み合わせ，研究者２名による評価により，4,558項目が残った。第三段階は，言語学研究者，文学研究者，辞書編集者など専門家６名による，教育的観点からの４段階の評価であった。その結果，全体の８％にあたる385項目が除外された。第四段階は，項目の体系化で，原形を用いたリスト化，アメリカ英語，イギリス英語の区別による整理，定冠詞・不定冠詞の付与，共起する前置詞情報の付与などが行われた。結果として，2,468項目によるリストが完成した。リストは，品詞ごとの組み合わせにより，（1）名詞結合（例，anecdotal evidence，assessment process），（2）動詞＋名詞・形容詞による結合（例，undertake research，consider appropriate），（3）動詞＋副詞による結合（例，explicitly state，vary considerably），（4）副詞＋形容詞による結合（例，highly controversial，(be) markedly different）の４種類に分類された。

教育者と学習者が受容語彙として教育，学習に用いること，受容語彙知

識の測定に用いること，語彙習得過程を測定する補助手段として用いることを主な目的として作成された定型表現リストに，Martinez and Schmitt（2012）による複数語句表現リスト（PHRASE List）がある。リストは，5,000 項目を上限と設定し，項目は，高頻度で，意味があり，構成性が低い（個別の単語の意味からは理解が難しい）という 3 つの規準をもとに選定された。対象となる複数語句は，下記のように定義されている。

> A phrasal expression is hence defined as a fixed or semi-fixed sequence of two or more co-occurring but not necessarily contiguous words with a cohesive meaning or function that is not easily discernible by decoding the individual words alone. (p. 304)

つまり，ここで対象とされているのは，学習，教育上の重要度から，個別の単語の意味のみでは認識が困難な語彙項目である。この定義に従い，主たる規準 3 項目，補助的規準 3 項目を基に，British National Corpus（BNC）を用いて項目の抽出が行われた。コーパス・コンコーダンサーを用いた項目候補の抽出，著者による質的選定作業を経て，最終的に 505 項目による PHRASE List が作成された。完成したリストには，各項目について，一般的話し言葉，一般的書き言葉，学術的書き言葉の 3 分類における重要度が 3 段階で表示されている。上位 10 項目を例に挙げると，1. have to，2. there is / are，3. such as，4. going to (FUTURE)，5. of course，6. a few，7. at least，8. such a(n)，9. I mean，10. a lot のようになる。最初の項目，have to を例にとって，その重要度を見ると，一般的話し言葉では最も重要とされる 3，一般的書き言葉では 2，学術的書き言葉では 1 となる。

本研究と同様に，口語での定型表現をリスト化する研究として，Shin and Nation（2007）がある。Martinez and Schmitt（2013）と同じ BNC を基礎コーパスとして用いているが，大きく異なる点は，Shin and Nation（2007）では，BNC 全体の 10％程度となる，話し言葉コーパス，約 1,000 万語のみ

が用いられていることである。基礎コーパスで頻出する，1,000語を軸語（pivot word）として用い，共起する語（collocate）の情報から話し言葉コロケーションを抽出した。その際，1,000万語につき30回以上の頻度で出現するなど6つの規準が設定された。結果として，4,698項目のコロケーションが抽出された。完成したリストの項目は，著者らが指摘するように，イギリス英語の独特の表現が多く含まれること，大部分は成人話者の発話であること，改まった文体ではなく，日常会話的であるという特徴を持つ。上位10項目を例に挙げると，1. you know，2. I think，3. a bit，4. used to，5. as well，6. a lot of，7. {No.} pounds，8. thank you，9. {No.} years，10. in fact のようになり，7はイギリス英語の特徴を持つことが分かる。このような特徴を持ちながらも，BNCにおける頻度情報では，本コロケーションリストの上位308項目は，単語リスト上位2,000語レベルに相当する頻度であり，定型表現が教育，学習の対象とすべき重要な項目であることが示された。

上記で概観した研究以外にも，最頻出の句動詞150項目を抽出したPHaVE List（Garnier & Schmitt, 2015）など教育的観点から作成された定型表現リストも存在する。本研究では，外国語学習環境で英語を学ぶ学習者を対象として，頻繁に用いられる基礎的口語表現リストの作成を目指している。そのため，学術場面を想定したリスト（Ackermann & Chen, 2013; Simpson-Vlach & Ellis, 2010），構成性が低い項目を含むもの（Martinez & Schmitt, 2010），頻出語を軸語とするコロケーション・リスト（Shin & Nation, 2007）とは重複する項目が含まれる可能性はあるものの，そのいずれとも目的が異なっている。

3．英語基礎口語定型表現リスト・教材の作成

本研究で作成する英語口語定型表現リスト（Basic Spoken Formula List）は，外国語学習環境の学習者を対象として想定している。具体的には，日本における英語学習のように，教室内外での目標言語インプットの量が不

足する環境の初中級の学習者を対象とする。また，抽出する定型表現は，先行研究に見られる学術場面ではなく，日常生活場面でのコミュニケーションで頻出し，対人コミュニケーションにおいて有用性が高いものを目指す。次に，本研究で作成する英語基礎口語定型表現リスト，教材，それぞれの作成手順を述べる。

3.1　英語基礎口語定型表現リスト（Basic Spoken Formula List）の作成

下記に本研究におけるリスト作成の過程を，基礎コーパスの構築，基礎口語定型表現の選出の 2 段階に分けて説明する。

3.1.1　基礎コーパス American Sitcom Corpus の構築

本研究の目的に照らすと，日常の様々な場面における話し言葉のコーパスが基礎コーパスとして必要となる。先行研究で用いられた BNC の話し言葉コーパス，現代アメリカ英語コーパス（Corpus of Contemporary American English, COCA）の話し言葉モジュールの利用なども候補となるが，場面の統制が難しいこと，教材を作成する際に定型表現が用いられた文脈の確認が難しいことなどから，別途，コーパスを作成することとした。しかし，独自に話し言葉を収集してコーパス化するには，時間と費用の課題がある。そのため，本研究では，日常における様々な対人コミュニケーション場面が再現されていると考えられる TV ドラマ作品の脚本によりコーパスを構築することとした（American Sitcom Corpus）。また，その際，アメリカの TV ドラマでプライム・タイム（一般的に 19–23 時）に放送され，放送コード上，卑語を多く含まないと考えられるシチュエーション・コメディー作品を用いることとした。同様の例として，規模は異なるが，COCA にあるアメリカ・ソープオペラ・コーパス（SOAP），映画コーパス（Movies），テレビ・コーパス（TV Corpus）がある（Davies, 2021）。このうち，TV Corpus（Davies, 2021）を用例検索の補助手段として用いることとした。

American Sitcom Corpus（以下，ASC）の作成にあたり，様々なファン

サイト（例，Fandom 2021; Big Bang Theory Transcripts, 2019）に公開されている脚本データを収集し，放送期ごとにテキストファイルを作成した。対象としたのは，2005年から2016年に放送された18作品である（付録1）。脚本データ収集の結果，約525万語の総語数（token）から成るコーパスが完成した。なお，異なり語数（type）は，66,447語であった。表1に，口語の特徴を明らかにするために，レマ化を行っていない，縮約形も1語と数える形式での本コーパスの頻度順上位50語を示す。

表1．American Sitcom Corpus 頻度順上位50語

順位	単語	頻度	割合	順位	単語	頻度	割合
1	you	181,775	3.46%	26	for	36,296	0.69%
2	I	174,658	3.33%	27	go	34,792	0.66%
3	be	169,345	3.23%	28	not	34,364	0.65%
4	the	137,097	2.61%	29	your	34,077	0.65%
5	a	126,824	2.42%	30	so	33,569	0.64%
6	to	120,605	2.30%	31	it's	33,396	0.64%
7	and	84,601	1.61%	32	don't	30,923	0.59%
8	that	69,150	1.32%	33	with	30,282	0.58%
9	it	66,350	1.26%	34	but	28,049	0.53%
10	of	57,640	1.10%	35	like	27,939	0.53%
11	do	53,947	1.03%	36	all	27,201	0.52%
12	have	50,943	0.97%	37	yeah	24,333	0.46%
13	this	50,092	0.95%	38	okay	24,053	0.46%
14	in	49,285	0.94%	39	you're	23,733	0.45%
15	my	46,918	0.89%	40	think	22,764	0.43%
16	me	45,792	0.87%	41	right	21,940	0.42%
17	I'm	45,427	0.87%	42	out	21,554	0.41%
18	get	45,008	0.86%	43	about	21,215	0.40%
19	what	43,243	0.82%	44	well	20,703	0.39%
20	no	38,177	0.73%	45	want	19,754	0.38%
21	we	37,946	0.72%	46	up	19,584	0.37%
22	oh	37,857	0.72%	47	that's	18,791	0.36%
23	know	37,509	0.71%	48	here	18,689	0.36%
24	just	36,878	0.70%	49	say	18,541	0.35%
25	on	36,665	0.70%	50	at	18,283	0.35%

本コーパスの最大の特徴は，話し言葉，書き言葉，ジャンルを均衡化したコーパス（例，COCA, BNC）とは異なり，「ここで今」（here and now）行われている会話を反映しているため，you（3.46%），I（3.33%）の使用割合が高いことである。例えば，現代アメリカ英語コーパス（COCA）頻度上位 5,000 語のリストでは，I は頻度順 11 位で割合が 0.93%，you は頻度順 14 位で 0.92%である（Davies, 2011）。本研究の ASC における，この特徴は，同じく英語使用圏の TV ドラマ脚本を使用した約 3 億語の TV Corpus（Davies, 2021）における特徴，I（3.75%），you（3.73%）とも一致するものである。ASC でも，その縮約形である，I'm（0.87%），you're（0.45%）を I, you にそれぞれ加えれば，その割合は，さらに高いもの（you ＝3.91%，I＝4.20%）となる。その他にも，表 1 で確認できるように，一人称の目的格（me），所有格（my），二人称の所有格（your）の頻度が高いことも確認できる。TV ドラマでは，様々な場面が登場するが二者間で行われる，現在進行中の会話が多いと考えられる。そのため，ASC は，基本的な対人コミュニケーション場面における，基礎口語定型表現を抽出する基礎コーパスとして，適切なものと判断できる。

3.1.2　基礎口語定型表現の選出

基礎口語定型表現の選出には，（1）複数語句の抽出，（2）基礎口語定型表現としての抽出規準の作成，（3）抽出規準を用いた取捨選択の 3 段階を経ることとした。

ASC の情報検索，分析には，CasualConc（Imao, 2019）を使用した。CasualConc は Mac OS で用いることができるコーパス・コンコーダンス・ソフトウェアである。はじめに，2 語から 7 語の n-gram（複数語句）リストを作成した。その概要を確認したところ，6 語，7 語では頻度が高い用例は多くなく，そのほとんどが 5 語に含まれることから分析の対象は 2 語から 5 語の複数語句とした。

次に，複数語句から基礎口語表現を抽出するための規準を作成することとした。定型表現の抽出に複数の規準によるチェックリストを用いること

は，先行研究でも行われているが，抽出する定型表現の目的により，規準は様々である（例，Shin & Nation, 2007; Wray & Namba, 2006）。本研究は，日常的な対人コミュニケーション場面で用いることのできる基礎的口語定型表現の抽出を目的とするため，下記の4つの規準で確認することとした。

（1）2語以上であること（2－5語）
（2）100万語で標準化した頻度が10回以上であること（more than 10 word-per-million）
（3）語用論的まとまり（pragmatic integrity）があること
（4）意味は，原義通りではない不透明なものよりも，透明性の高い，原義通りであること（transparent rather than opaque, literal rather than non-literal）

このうち，（1）の規準を用いたのは既に述べた通りである。（2）は，Simpson-Vlach and Ellis（2010）と同じで，Shin and Nation（2007）よりは厳しい基準となる。（3）は，本研究で対象とする日常的な口頭コミュニケーションを考慮しての規準である。語用論的まとまりの例としては，談話標識（discourse marking）の you know, if you know what I mean，フェイスやポライトネスに関わる表現の do you think, do you want me to，曖昧さや近似を表す表現の a couple of, or something like that などが挙げられる（O'Keeffe, McCarthy, & Carter, 2007）。（4）については，透明性の低い表現（例，a piece of work で「扱いにくい人」を表す）の重要性を否定するものではない。しかし，本研究では，日常的な口頭コミュニケーション場面における基礎的定型表現の抽出を目的としており，Adolphs and Carter（2013）が指摘するように，「母語話者の直感でも即座には気づくことが難しく，概して潜在意識的に用いられる一般的な定型表現」（p. 27）が対象であるため，原義通りの透明性があることを重視した。

最初の2つの規準により，候補となる項目数は，2語が10,202項目，3語が4,377項目，4語が717項目，5語が119項目の計15,415項目となっ

た。次に語用論的まとまりを持つものを残すために，Shin and Nation（2007）のコロケーションに含まれる独立した語彙項目（例，credit card, birthday party，video game）を除外した。さらに，時間（all day，an hour, all night），空間（out of here，right here，over there）に関わる項目を除外した。ただし, 除外した項目も他の語との結合により, 最終的な項目に残ったものもある（例，get out of here における out of here）。なお，本定型表現リストは，文法知識が十分に発達していない初級段階からも，意味を持つまとまりとして受容，産出できる定型表現項目を抽出する目的を持つため，複数語句の抽出にあたっては，派生形を 1 つの見出しとしてまとめるレマ化の作業は行っていない。

上記の作業と並行して，重複する項目を確認する作業および異なる綴りで同じ語からなる定型表現の見出し項目を統合する作業を手作業で行った。重複する項目の例として，最頻出の 2 語表現，you know を挙げる。you know は単独で談話標識として用いられることもあるが，本リストの項目には，その他にも下記の例がある。

you know what，do you know，you know how，but you know，
well you know, did you know, you know it, how do you know,
do you know what，you know why，you know what I mean,
because (cause) you know，you know who，let you know,
you know maybe，you know what that，you know what it's,
if you know，as you know，you know me,

これらは，you know の定型表現族（formula family）として，その情報をまとめた。異なる綴りを見出し項目として統合した例には，it would be / it'd be, what is this / what's this, I just want to / I just wanna などが含まれる。また，その頻度情報も合計して順位を表示することとした。さらに，見出し項目となる定型表現の左右に頻繁に共起する語がある場合は，それぞれ L-collocate，R-collocate として情報を補った。以上の作業の結果，最終的

に2,585項目の定型表現を抽出し，英語基礎口語定型表現リストを作成した。表2に上位50項目の見出し，標準頻度（100万語あたりの頻度，word per million）を示す。また，付録2として上位51-200項目を示した。

表2に示したのは，2,585項目中のわずか50項目であるが，限られた項目数でも英語基礎口語定型表現リストの特徴をいくつか指摘することができる。中でも特徴的なのは，前節でも指摘したように，典型的に二者間で展開される会話を前提とした，Iとyouを含む，対人コミュニケーション

表2．基礎口語定型表現上位50項目

順位	項目	標準頻度	順位	項目	標準頻度
1	you know	2890.76	26	you guys	790.80
2	I don't	2560.97	27	did you	761.28
3	are you	2068.99	28	I'm sorry	729.32
4	this is	1916.83	29	need to	729.13
5	have to	1826.69	30	you know what	707.32
6	do you	1608.86	31	I didn't	698.31
7	want to	1569.74	32	what are you	697.92
8	I know	1426.48	33	it's not	677.80
9	I have	1339.05	34	a lot	663.89
10	all right	1291.86	35	oh my god	659.37
11	no no	1259.90	36	talk about	647.56
12	I think	1222.67	37	I want	646.40
13	I'm gonna / going to	1147.56	38	for you	621.02
14	going to	1101.59	39	you think	607.67
15	I mean	1067.18	40	I thought	578.15
16	a little	1059.85	41	and then	565.74
17	come on	989.16	42	I love	529.08
18	if you	964.90	43	look at	520.68
19	thank you	955.31	44	we have	499.18
20	I can't	954.56	45	talk to	496.99
21	I got	914.89	46	one of	483.58
22	you want	907.89	47	I need	483.39
23	I don't know	853.78	48	let me	478.31
24	you guys	837.80	49	I had	472.49
25	I can	801.89	50	go to	471.61

を成立させるための定型表現の割合の高さである。上位 50 項目でも，およそ半数が I，you，we，me を含んでいる。教室指導会話コーパス，教科書コーパスと Biber et al.（1999）で用いられた Longman コーパスの会話領域のデータの比較を行った，Biber et al.（2004）と同様に，指示表現（referential expression）よりも，態度表現（stance expression），談話構成表現（discourse organizing expression）が多くの割合で含まれることが確認できる。

次に指摘しておかなければならないのは，より語数の少ない項目が拡張して，より語数の多い項目を形成する，語彙連結（lexical bundle）としての特徴である。頻度順，1 位の 2 語項目の you know に 16 項目の定型表現族が含まれることは前述したが，例えば，12 位の I think にも，次のような拡張表現がある。

> I think we，but I think，I think you're，I think that，I think we're，I think we should，I think I'm gonna / going to，I think I know，I think I can，I think you should，I think so，I think I have，I think it's a，I think we can，

この特徴は，すでに Biber et al.（1999）でも指摘されており，単なる拡張（例，I think so）以外にも，異なる語彙連結同士の統合（例，I think we should，I think I know）もその特徴である。

次に指摘できる特徴は，動詞句を形成する準法助動詞（semi-modal verb）や句動詞である。上位 50 項目では，have to，need to の準法助動詞，look at，talk about，talk to など動詞と前置詞の組み合わせで他動詞の役割を果たす句動詞が確認できる。その他，17 位の句動詞 come on は，Biber et al.（1999）でも領域を問わず，最も頻出する句動詞であることが報告されており，基礎コーパスとして用いた ASC でもある行動を強く促す働き（例，Come on，let me help you.），移動を呼びかける働き（例，Come on，let's get out of here.）が多く確認できる。

上記の他にも，上位50項目では，複数の機能を果たす定型表現（例，a little，all right，a lot），談話指標（例，I mean，you know，you know what），謝罪，同情，感謝などの感情表現（例，I'm sorry，thank you）などの特徴も確認できる。複数の機能を果たす例として挙げた，a little には，副詞句で数量詞（例，Take this and relax a little.），副詞句で減少詞（例，Let's catch up a little bit.），形容詞句で数量詞（例，So a little change of plans.）などの用法が確認できる。談話指標の例として挙げた3項目も，より複雑な文を構成する一部になるなど，談話指標以外の働きが確認できる。頻度の高い定型表現は，見出しの項目としては1つであっても，その機能は多義的であることは，教材作成の段階においても注意が必要であると言える。

上述した特徴を持つ，英語基礎口語定型表現リストは，Biber et al.（1999）で指摘された，「様々な状況において，様々な話者によって頻繁に用いられる語彙的構成単位（lexical building blocks）」（p. 991）にあたる。特に，基礎コーパス ASC は，TV ドラマで異なる話者間のコミュニケーションとして展開される「ここで今」行われている会話であるため，本研究で抽出した基礎口語定型表現の多くは，「対人コミュニケーションを成立させる語彙的構成単位」（lexical building blocks for interpersonal communication）と見なすことができる。

3.2　英語基礎口語定型表現リストに基づく教材作成

次に，完成した英語基礎口語定型表現リスト（Basic Spoken Formula List）に基づいた教材を作成することとした。言語学習教材作成の過程について，Jolly and Bolitho（2011）では，（1）教材作成のニーズの同定（identification），（2）ニーズに関する言語知識，技能，機能，意味の探索（exploration），（3）教材の状況的実現（contextual realization），（4）教材の教育方法的実現（pedagogical realization），（5）教材の物理的生成（physical production），（6）教材の使用と学習・指導目標に合わせた評価（evaluation）の6段階を想定している。また，これらの段階を直線的に進むこともあるが，各段階を行き来する動的な過程にもなりうることを指摘

している。

本研究においては，(1) の学習上のニーズについては，対人コミュニケーションを円滑に進めるための基礎口語定型表現知識として想定している。また，(2) とは基礎コーパスの構築，探索を経て，(3) は基礎口語定型表現リストの作成として行なっている。次は，基礎口語定型表現リストとそれに至る段階を見直しながら，見出し項目ごとに例文を付与する (4) の教育方法的実現の段階となる。なお，(5) 教材の物理的生成，(6) 教材の使用と評価については，本稿公表以降の作業となる。

本研究の目的と上記の教材作成の過程を踏まえ，教材の目的は，口頭での対人コミュニケーションを円滑に進めるために，TV ドラマのセリフのような例文で基礎口語定型表現，会話表現を学ぶこととした。また，対象は，初中級段階の英語学習者（中学生，高校生，大学生，一般）とした。例文には，ドラマのように登場人物を 5 名設定した（Sara，Kent，Rui，Aoi，Kai）。教材は，下記の原則にしたがって基礎口語定型表現リストの見出し項目に例文を付与することとした。例文の作成には，英語母語話者であり，米国出身の応用言語学・外国語教育学研究者 1 名の協力を得た。

(1) 基礎コーパス ASC の情報に基づいたものとする，その際，コーパス・コンコーダンサー（CasualConc）を用いて情報を検索する
(2) 口語体とする
(3) 学習者の興味を引く，楽しいものとする
(4) 言語機能，可能な場合は，語用論的意味を持たせる
(5) 現実の会話と同じように，ある程度までは無作法，失礼な発言を許容する
(6) 幅広い年代の話者によって用いられるものとする（暴力的な表現，卑語の使用を避ける）

完成した例文の有用性を検証することを目的に，表 3 に項目と例文を挙

げる。例文の抽出が意図的にならないよう，頻度順100項目ごとに，計26文を取り上げる。

表3．英語基礎口語定型表現リスト例文例

順位	項目	例文
0001	*you know*	I'm probably wrong, but, **you know**, I just want to make sure.
0101	*do you think*	Why **do you think** it's my fault?
0201	*don't care*	Fine! I **don't care**! I **don't care**! I never wanted that job in the first place!
0301	*I promise*	If you follow that itinerary to the letter, **I promise** you, you will not be disappointed.
0401	*how can*	**How can** I ever thank you?
0501	*there is no*	You're the only one I can count on. **There is no** one else.
0601	*good morning*	**Good morning**, pumpkin, and **good morning**, you lovely lady.
0701	*you're gonna have*	**You're gonna have** to change all the plans you made.
0801	*you know why*	I haven't been sleeping well. **You know why**? Because you're always snoring.
0901	*I don't think so*	You come to me asking for money, and you treat me like this? **I don't think so**.
1001	*just give me*	**Just give me** one quick sec to confer with my associate.
1101	*for the record*	**For the record**, I tried to get something done about this two years ago.
1201	*every night*	You've been out **every night** this week.
1301	*it's nothing*	I feel a little light-headed. All right, well, I'm sure **it's nothing**, you know.
1401	*just kidding*	Here's dinner. Breast of frog on a bed of wild mice. **Just kidding**. It's chicken.

順位	項目	例文
1501	*was thinking about*	I just wanted to let you know that I **was thinking about** you.
1601	*I can't believe this*	**I can't believe this** is happening.
1701	*I want to go*	I don't think **I want to go** out with you anymore.
1801	*started to*	So you sign? – I do. My brother's deaf, and I **started to** learn, and fell in love with the language, I guess.
1901	*bowl of*	I've been waiting two months for that **bowl of** ice cream.
2001	*how's that*	Hey, remember that time at the lake I introduced you to Aoi? **How's that** going?
2101	*I don't know what you're*	**I don't know what you're** talking about, Rui.
2201	*running for*	Are you theoretically interested in **running for** office?
2301	*I feel terrible*	**I feel terrible**, but we tried everything.
2401	*I never said*	Okay, maybe **I never said** it out loud, but I was thinking about it.
2501	*many things*	I'm a wreck. There are **many things** seriously wrong with me.

例文の作成にあたり，特定の言語機能の導入を前提としていなかったが，結果的に表３に見られるように，言語の働きが想像できる，語用論的なまとまりを持つ例文が多くなっている。言語機能としては，感謝（例，How can I ever thank you?），警告・助言（例，You're gonna have to change all the plans you made.），懇願（例，Just give me one quick sec to confer with my associate.），非難（例，You've been out every night this week.），後悔・謝罪（例，I feel terrible, but we tried everything.）などが確認できる。この理由として，定型表現によっては，その表現自体が語用論的なまとまりを持つ，語用論的定型表現（pragmatic routine）である項目が多く含まれ

ていることが指摘できる。例えば，it's nothing は，表3では自分の「頭がふらふらする（light-headed）」症状が，「何でもない」の原義通りの意味で用いられているが，その他にも，自分の行為に対して相手に感謝された際に謙遜する丁寧な応答として用いられることもある。また，you're gonna have は表3の例文で確認できるように，to を補って，「（あなたは）～しなければならなくなる」の意味で警告，助言として機能する。前節で確認したように，基礎口語定型表現自体が，対人コミュニケーションを成立させるための構成単位（building blocks）として機能していることを示している。

その他の特徴として，つなぎ言葉（filler）の you know，談話指標の for the record「はっきり言うと」，約束型の発話内行為である I promise など，会話を成立させる構成単位としての定型表現使用が確認できる。一方で，表3では，すでに学習者にとって馴染みがあると考えられる挨拶表現（good morning）や bowl of といった単純な名詞句を構成する項目，例文も含まれている。本研究の4つの規準を使用した定型表現選出は，著者1名で行なったため，表3に示した以外の同様の項目については，例文の有用性を含めて，再検討の必要がある。なお，本研究で作成した英語基礎口語定型表現リストと例文については，プロジェクト用のウェブサイト（https://sites.google.com/view/02241112/home）で公開している。

4．今後の課題

本稿では，初中級の英語学習者を対象に，日常生活場面での対人コミュニケーションにおいて有用性が高い，基礎口語定型表現リストとリストに基づく教材作成の過程，結果について報告した。完成したリストは，結果として「対人コミュニケーションを成立させる語彙的構成単位」（lexical building blocks for interpersonal communication）を多く含むものとなった。また，定型表現項目を用いた教材は，語用論的まとまりを持ち，多様な言語機能を含む例文集となった。今後は，下記に挙げる課題を検討し，

学習や指導の目標に合わせたリストの整理，教材の完成を目指す段階となる。

まず，現段階での基礎口語定型表現リストおよび教材は，補足情報（日本語訳，英語音声）がなく，項目の頻度のみによる見出し順などの課題を抱えた暫定的なものである。実用時には，当然，学習・指導目標に合わせた修正や再編集の作業が必要になる。最終的にリスト，教材とも学習者にとって実用性のあるものにするには，定型表現項目の妥当性や有用性について，先行研究のように専門家を含めた評価，あるいは学習者による評価のような実証的な検証作業も重要となる。

第一に，挙げられる課題は，見出し項目の再検討である。例えば，本リストでは，Shin and Nation（2007）では除外されている，I don't，are you，do you などの基礎的な 2 語項目についてもリストに含めている。このうち，I don't については，本リスト内において，それを含む 30 以上の見出し項目が定型表現族として存在し，そのほとんどが，I don't に続けて展開，拡張する定型表現となっている。しかし，are you の定型表現族については，are you＋形容詞，are you＋～ing，are you going (to) [FUTURE]，疑問詞＋are you＋形容詞，疑問詞＋are you＋～ing など多くの文構造が含まれるため，1 つの項目としてまとめることの妥当性の検証や例文を用いた体系的な教材作成時の配慮が必要になると考えられる。こうした煩雑さを避けるには，意味的なまとまりがない 2 語項目を削除することも可能性としてあるが，you know，I mean，you guys など 2 語項目のみで独立して頻繁に用いられる項目を考慮すると，その判断は容易ではない。

次にリスト内の項目数の課題があげられる。最終的に，本リストの見出しは，2,585 項目となり，そのまま学習対象とするには項目数が非常に多い。現時点では，定型表現項目の抽出作業は，著者 1 名で行なっている。そのため，上記にあげた，専門家や学習者を対象とした実証的な検証作業による整理が重要であると考えられる。また，その際には，教材としての見出し項目の類型化，体系化や言語機能情報などの付加による整理などの改善が考えられる。さらに，そうした情報を付加した後に，田中（2021）

で報告したようなモバイル端末アプリなど学習システムでの応用も検討できる。

学習教材としては，Jolly and Bolitho（2011）に示された，（5）教材の物理的生成（Physical production），（6）教材の使用と学習・指導目標に合わせた評価（Evaluation）の2段階については今後の作業となる。今後は，上記に挙げた課題の解決を検討して，英語口語基礎定型表現リスト（Basic Spoken Formula List）および教材の完成とオンライン，オフラインによる公開を進めることとしたい。

謝辞

本研究は科研費（19K00854）の助成を受けたものである。

英語基礎口語表現リスト例文教材の作成にあたっては，北海学園大学人文学部，米坂スザンヌ教授にご協力いただきました。心より感謝申し上げます。

引用文献

Ackermann, K., & Chen, Y. H. (2013). Developing the academic collocation list (ACL) - A corpus-driven and expert-judged approach. *Journal of English for Academic Purposes, 12*(4), 235–247. https://doi.org/10.1016/j.jeap.2013.08.002

Adolphs, S., & Carter, R. (2013). *Spoken corpus linguistics: From monomodal to multimodal*. Routledge.

Benson, M., Benson, E., & Ilson, R. (1986, 1997, 2010). *The BBI combinatory dictionary of English*. John Benjamins Publishing Company.

Bardovi-Harlig, K. (2012). Formulas, routines, and conventional expressions in pragmatics research. *Annual Review of Applied Linguistics, 32*, 206–227. https://doi.org/10.1017/s0267190512000086

Biber, D., Johansson, S., Leech, G., Conrad, S., & Finegan, E. (1999). *Longman grammar of spoken and written English*. Longman.

Big Bang Theory Transcripts. (2009). Big Bang Theory Transcripts. Retrieved December 1, 2016, from https://bigbangtrans.wordpress.com/

Browne, C. (2014). A new general service list: The better mousetrap we've been

looking for? *Vocabulary Learning and Instruction, 3*(2). 1–10. https://doi.org/10.7820/vli.v03.2.browne

Davies, M. (2011). *Word frequency: Based on one billion word COCA corpus*. Word frequency data. https://www.wordfrequency.info

Davies, M. (2021). *Full-text data from English-Corpora.org: Billions of words of downloadable data*. Full-text corpus data. https://www.corpusdata.org/

Erman, B., & Warren, B. (2000). The idiom principle and the open choice principle. *Text-Interdisciplinary Journal for the Study of Discourse, 20*(1), 29–62. https://doi.org/10.1515/text.1.2000.20.1.29

Fandom Home Page. (2021). Fandom. https://www.fandom.com/

Foster, P. (2001). Rules and routines: a consideration of their role in the task-based language production of native and non-native speakers. In M. Bygate, P. Skehan, & M. Swain (Eds.), *Researching pedagogical tasks: second language learning, teaching and testing* (pp.75–93). Longman.

Garnier, M., & Schmitt, N. (2014). The PHaVE List: A pedagogical list of phrasal verbs and their most frequent meaning senses. *Language Teaching Research, 19*(6), 645–666. https://doi.org/10.1177/1362168814559798

Hsu, J. T., & Chiu, C. Y. (2008). Lexical collocations and their relation to speaking proficiency of college EFL learners in Taiwan. *The Asian EFL Journal, 10*(1), 181–204.

Imao, Y. (2019). CasualConc (Version 1.9.8) [Computer software] https://sites.google.com/site/casualconc/home

Jolly, D. & Bolitho, R. (2011). A framework for materials writing. In B. Tomlinson, (Ed.), *Materials development in language teaching*. Cambridge University Press.

Keshavarz, M. H., & Salimi, H. (2007). Collocational competence and cloze test performance: a study of Iranian EFL learners. *International Journal of Applied Linguistics, 17*(1), 81–92. https://doi.org/10.1111/j.1473–4192.2007.00134.x

Longman Dictionary of Contemporary English (Fifth edition). (2009). Pearson Education Limited.

Mackin, R., & Cowie, A. P. (1983). *Oxford dictionary of current idiomatic English*. Oxford University Press.

Martinez, R., & Schmitt, N. (2012). A Phrasal Expressions List. *Applied*

Linguistics, 33(3), 299–320. https://doi.org/10.1093/applin/ams010

O'Keeffe, A., McCarthy, M., & Carter, R. (2007). *From corpus to classroom: Language use and language teaching*. Cambridge University Press.

Schmitt, N. (2010). *Researching vocabulary: A vocabulary research manual*. Palgrave Macmillan.

Schmitt, N., & Schmitt, D. (2020). *Vocabulary in language teaching* (2nd ed.). Cambridge University Press.

Shin, D., & Nation, P. (2007). Beyond single words: the most frequent collocations in spoken English. *ELT Journal, 62*(4), 339-348. https://doi.org/10.1093/elt/ccm091

Simpson-Vlach, R., & Ellis, N. C. (2010). An academic formulas list: New methods in phraseology research. *Applied Linguistics, 31* (4), 487–512. https://doi.org/10.1093/applin/amp058

West, M. (1953). *A general service list of English words*. Longman, Green & Co.

Wood, D. (2006). Uses and functions of formulaic sequences in second language speech: An exploration of the foundation of fluency. *Canadian Modern Language Review, 63*(1), 13–33.

Wray, A. (2002). *Formulaic language and the lexicon*. Cambridge University Press.

Wray, A. (2017). Formulaic sequences as a regulatory mechanism for cognitive perturbations during the achievement of social goals. *Topics in Cognitive Science, 9*(3), 569–587. https://doi.org/10.1111/tops.12257

Wray, A. & Namba, K. (2003). Use of formulaic language by a Japanese-English bilingual child: A practical approach to data analysis. *Japanese Journal for Multilingualism and Multiculturalism, 9*(1), 24–51.

井上永幸・赤野一郎（2018）.『ウィズダム英和辞典』（第4版）. 三省堂. 東京.

勝俣詮吉郎（1939, 1958, 1995）.『研究社英和活用大辞典』研究社. 東京.

神田乃武・南田恒太郎（1909）.『英和雙解熟語大辞典』有朋堂書店. 東京.

大学英語教育学会基本語改訂特別委員会（2016）.『大学英語教育学会基本語リスト　新JACET8000』桐原書店. 東京.

田中洋也（2015）.「定型表現に基づく英語アニメーション教材：語用論的能力育成を目指して」『北海学園大学学園論集』第164・165号, 1–23.

田中洋也（2020）.「日本の世界遺産を伝える英語語彙・定型表現：英語の学びの視点から」北海学園大学人文学部世界遺産研究班（編）.『世界遺産とは何

か―様々な「物語」を読み解く』（188-214 頁）．マイナビ出版．東京．
田中洋也（2021）．「電子ポートフォリオ連携型英語語彙学習アプリ・フレーズ学習モジュールの開発：長期自律外国語学習支援を目指して」『人文論集』第 70 号，1-23.
東京外国語大学投野由紀夫研究室（2020）．「CEFR-J Wordlist Version 1.6」Retrieved from https://www.cefr-j.org/download.html
南出康世（2014）．『ジーニアス英和辞典』（第 5 版）．大修館書店．東京．
八木克正・井上亜依（2013）．『英語定型表現研究』開拓社．東京．

付録 1．American Sitcom Corpus に用いた作品

作品名	放送期
The Big Bang Theory	Season 1 – 8
Modern Family	Season 1 – 6
Parks and Recreation	Season 2 – 6
Baby Daddy	Season 1 – 4
Jessie	Season 1 – 2
Girl Meets World	Season 1 – 2
The Middle	Season 3 – 6
It's always sunny in Philadelphia	Season 6 – 10
2 Broke Girls	Season 1 – 4
Bad Teacher	Season 1
Silicon Valley	Season 1 – 2
Switched at Birth	Season 1 – 4
Raising Hope	Season 1 – 4
Community	Season 1 – 5
Happy Endings	Season 1 – 3
Melissa & Joey	Season 2 – 4
Mike & Molly	Season 1 – 5
How I Met Your Mother	Season 6 – 9

付録2．英語口語基礎定型表現リスト(General Spoken Formula List)上位51-200

51	that is	82	I want to	114	I can't believe
52	I could	83	can you	115	we could
53	what do you	84	you need	116	what happened
54	you're going to / gonna	85	oh yeah	117	come to
		86	would you	118	we need
55	you can't / cannot	87	the only	119	you have to
56	of course	88	that's not	120	try to
57	I'm just	89	the best	121	how to
58	I guess	90	get to	122	you mean
59	the way (that)	91	why don't	123	the other
60	we can	92	I like	124	go back
61	don't you	93	have you	125	I'm not going to / gonna
62	you want to	94	the first		
63	you get	95	you could	126	the last
64	I get	96	the one	127	why don't you
65	can I	97	you go	128	there is
66	a lot of	98	get out	129	you see
67	we're gonna / going to / are going to	99	we should	130	I see
		100	give me	131	what's going on / what is going on
68	oh no	101	got it		
69	I would	102	do you think	132	you look
70	see you	103	let's go	133	you got to / gotta
71	so much	104	I should	134	at least
72	tell me	105	I feel	135	look like
73	what are you doing	106	I don't want to / wanna	136	oh god
74	it's just			137	that's why
75	you should	107	I don't think	138	do you want
76	I don't want	108	you didn't	139	how about
77	I have to	109	all of	140	I've got
78	supposed to	110	that's what	141	I need to
79	back to	111	used to	142	there's no
80	the same	112	excuse me	143	don't worry
81	think about	113	each other	144	yeah (/yes), but

145	what if
146	at all
147	you're right
148	no one
149	I'm sure
150	time to
151	I hope
152	listen to
153	sound like
154	find out
155	what about
156	get out of
157	why are you
158	I found / find
159	break up
160	can we
161	you don't have
162	I love you
163	just like

164	more than
165	do you know
166	how much
167	I hate
168	it's okay
169	you need to
170	go out
171	out there
172	part of
173	about to
174	there are
175	here we
176	get back
177	we have to
178	the next
179	think of
180	I thought you
181	come in (here)
182	I told you

183	a little bit
184	get in
185	do you have
186	I don't have
187	I would like to / I'd like to
188	could you
189	I was just
190	we need to
191	take care (of)
192	come up (with)
193	back in
194	thanks for
195	what's up (with)
196	go on
197	come out (of)
198	I don't know what
199	I wanted to
200	I had to