# The KISTEC Search Interface の概要：KIT Speaking Test Corpus の普及に向けて

田中 悠介 (熊本学園大学) *
瀬戸口 彩花 (都城市立明和小学校)
近 大志 (富山大学)
神澤 克徳 (京都工芸繊維大学)

# Promoting the Use of the KIT Speaking Test Corpus through the KISTEC Search Interface

Yusuke Tanaka (Kumamoto Gakuen University)
Ayaka Setoguchi (Meiwa Elementary School, Miyakonojo)
Taishi Chika (University of Toyama)
Katsunori Kanzawa (Kyoto Institute of Technology)

**要旨**

本稿では、KIT Speaking Test Corpus（KISTEC）の検索および分析を支援する Web インターフェースである The KISTEC Search Interface を紹介する。KISTEC は、2018 年度に京都工芸繊維大学で実施された英語スピーキングテストを受験した 1 年次学生のうち、コーパス構築へのデータ利用に同意した 574 名の解答の書き起こし約 29 万語から構成され、発話の特徴を表すタグや学習者属性に加え、解答別の得点、受験者単位の総合得点、TOEIC L&R スコアなどの情報が付与されている。本インターフェースは、KWIC 検索、コロケーション分析、語彙リスト作成、特徴語抽出、N-gram 分析に対応し、設問、性別、各種得点による絞り込みと CSV 出力が可能である。これらの機能を紹介するとともに、研究および教育における活用可能性と、検索結果を解釈する際の留意点について論じる。

## 1. はじめに

The KIT Speaking Test Corpus（以下、KISTEC）は、京都工芸繊維大学で開発・実施された英語スピーキングテストの解答音声を書き起こした学習者コーパスである (Kanzawa et al. n.d.)。KISTEC の大きな特徴は、各解答に、主として課題内容の達成度を評価する Task Achievement（TA）と、主として発話の伝達の効果性を評価する Task Delivery（TD）の得点が付与されているほか、受験者単位のスピーキングテスト得点や TOEIC L&R スコアなどのメタデータも利用できる点にある。これにより、言語形式の出現傾向を記述するだけでなく、特定の語彙や表現、非流暢性現象とスピーキング評価との関係を、よりきめ細かく検討することができる。

* yu-tanaka@kumagaku.ac.jp

しかし、KISTEC は主としてテキストファイルで配布されており、従来は利用者がデータをダウンロードし、目的に応じて加工したうえで検索する必要があった。そこで著者らは、Web ブラウザ上で KISTEC を検索・分析することを可能にする The KISTEC Search Interface を開発し、公開した (Tanaka et al. n.d.)。本インターフェースは、一般的なコーパス検索機能に加え、KISTEC 固有の評価情報や学習者属性に基づいてデータを絞り込む機能を 1 つの環境に統合している。

本稿は、Tanaka et al. (2026) で報告した内容を、言語資源の公開と利活用促進という観点から再構成するものである。以下では、まず KISTEC の構成と評価情報を整理し、次にインターフェースの設計と主要機能を紹介する。そのうえで、研究および教育における活用可能性と、検索結果を解釈する際の留意点を述べる。

## 2. KISTEC の概要

KISTEC には、2018 年度に京都工芸繊維大学で独自に開発・実施された英語スピーキングテスト（KIT Speaking Test）を受験した 1 年次学生のうち、コーパス構築へのデータ利用に同意した 574 名分の解答が収録されている。総語数は約 290,000 語、総発話時間は約 74 時間である。ただし、音声データはプライバシー保護の観点から公開されていない。受験者の約 97% は日本人であるが、留学生も含まれる。そのため、本稿では収録者全体を指す場合、「英語学習者」または「受験者」と表記する (Kanzawa et al. n.d.)。

KIT Speaking Test は、CBT 方式で実施されるスピーキングテストであり、3 つのバージョンが用意され、それぞれ 9 つの設問から構成される。各解答は、課題で求められた内容をどの程度達成しているかを表す TA と、発話内容がどの程度効果的に伝達されているかを表す TD の 2 つの観点から評価される。各観点について、ネイティブスピーカーとノンネイティブスピーカーの評価者各 1 名が 0–5 点で評定し、原則として両者の評定値の平均が解答ごとの得点として記録される。このため、TA および TD の得点は、0 から 5 までの 0.5 点刻みの 11 段階となる(1)。これらとは別に、項目応答理論に基づいてバージョン間で比較可能な受験者単位の TA および TD の能力値が推定されている。TA rank と TD rank は、それぞれの能力値の順位に基づいて受験者を 20% ずつ 5 群に分け、1–5 のランクを付与したものである。また、Overall score は 0–100 点で表され、課題達成を重視するテスト設計を反映して、TA に 80%、TD に 20% の重みを付して算出される (Kanzawa et al. n.d.)。

表 1 に示すように、TA 得点と TD 得点は解答ごとに付与されるのに対し、TA rank と TD rank は受験者単位で算出され、両者では尺度も異なる。後述する The KISTEC Search Interface の「TA」「TD」フィルターおよび KWIC 検索結果に表示される値は、各解答に付与された 0–5 点の TA 得点と TD 得点である。

KISTEC の書き起こしには、フィラー、繰り返し、自己訂正など、発話上の特徴を表す 16 種類のタグが付与されている。例えば、フィラーは <F>uh</F>、繰り返しは <R>I</R>、

(1) 2 名の評価者による評定値の差が 2 点以上である場合には、シニアレーターが再評定を行い、その評定値を得点として採用している。

表 1　KISTEC に含まれる主なスピーキング評価情報

| 項目 | 単位 | 尺度 | 内容 |
|---|---|---|---|
| TA 得点・TD 得点 | 各解答 | 0–5（0.5 刻み） | 2 名の評価者による評定の平均値 |
| TA rank・TD rank | 各受験者 | 1–5（1 刻み） | 項目応答理論による等化得点をもとに算出した 5 段階のランク |
| Overall score | 各受験者 | 0–100（1 刻み） | TA と TD を統合した等化得点 |

自己訂正は <SC>I am</SC> のように表される。これらのアノテーションを利用することで、語彙や表現だけでなく、発話の流暢性や産出過程に関わる特徴についても検討できる。実際、KISTEC は、フィラーの使用頻度と流暢性評価との関係の分析 (Tanaka et al. 2025) や、学習者英語における非流暢性を自動検出する BERT モデルの性能評価 (Skidmore and Moore 2023) に利用されている。

このようなタグと表 1 に示したテストスコアに加え、性別、国籍、テストバージョン、TOEIC L&R の総合スコア、Listening スコア、Reading スコアなどの情報も収録されている。このように、発話データにアノテーション、評価情報、学習者属性が統合されていることが、KISTEC の言語資源としての特徴である。

## 3.　The KISTEC Search Interface の設計

The KISTEC Search Interface は、KISTEC に特化した Web ベースの検索・分析環境である。利用者は、ソフトウェアをローカル環境にインストールしたり、コーパスファイルを分析ツールに読み込んだりすることなく、Web サイト上で検索や分析を行うことができる。本稿では、2026 年 8 月時点で公開されているバージョンの仕様を扱う。

本インターフェースの設計上の特徴は、語彙や表現の検索機能と、スピーキング評価情報や学習者属性に基づく絞り込み機能を統合している点にある。AntConc (Anthony 2005) に代表される一般的なコーパス分析ツールでも、KWIC 検索や語彙頻度分析は可能であるが、KISTEC に付与された評価得点や学習者属性に基づいて解答を絞り込むには、通常、利用者側でデータを加工する必要がある。本インターフェースでは、Overall score、解答別の TA 得点と TD 得点、TOEIC L&R の総合スコア、Listening スコア、Reading スコアについて下限と上限を指定できるほか、性別や設問（Q1–Q9）による絞り込みも可能である。

主要な分析機能は、KWIC 検索、コロケーション分析、語彙リスト作成、特徴語抽出、N-gram 分析の 5 種類である。各機能の結果は CSV 形式で出力できるため、画面上での探索を起点として、Excel、R、SPSS などを用いた後続分析へ移行できる。CSV 出力は検索結果の保存と再利用を容易にするが、分析の再現性を高めるためには、出力ファイルとともに検索語、検索モード、フィルター条件、利用日なども記録することが望ましい。

## 4. 主要機能

### 4.1 KWIC 検索

KWIC 検索では、検索語の各出現例を前後の文脈とともに一覧表示する。例えば、*go* を検索すると、*go to*、*go shopping*、*go there* などの出現例から、学習者による *go* の用法や周辺語との結びつきを確認できる（図 1）。検索結果には、左文脈、検索語、右文脈に加え、当該解答の TA 得点および TD 得点が表示される。

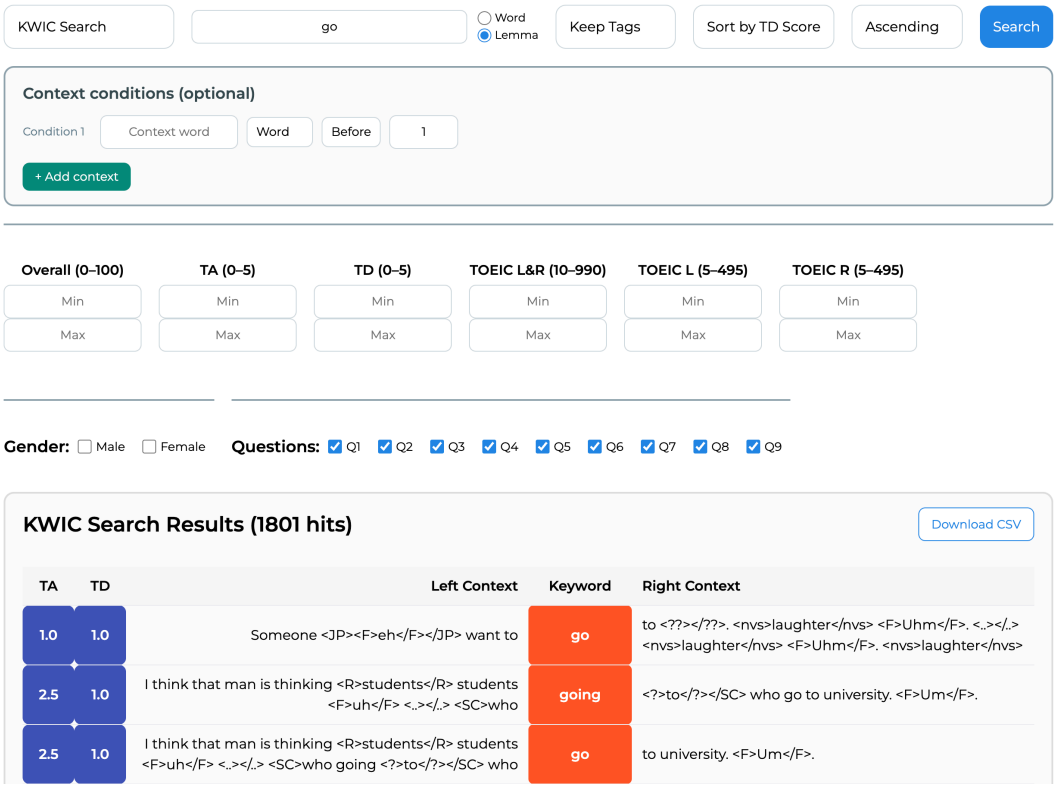


図 1　検索語を *go* とした KWIC 検索の例

検索モードは、語形検索とレンマ検索から選択できる。レンマ検索では、AntBNC Lemma List（ver. 004）(Anthony n.d.) を用い、例えば *go* を指定すると、*go*、*goes*、*going*、*went* をまとめて検索できる。また、タグの表示・非表示を切り替えられるほか、TA 得点または TD 得点を基準とした並べ替えと、昇順・降順の指定が可能である。

さらに、“Context conditions (optional)” 欄では、検索語の前後の特定の位置に現れる語を検索条件として指定できる。例えば、検索語を *go*、右 1 語目を *to* と指定すると、*go to* を含む用例に限定して検索できる。また、右 1 語目を *to*、右 2 語目を *university* と指定すると、*go to university* を含む用例が抽出される。このように、複数の文脈条件を指定した場合には、すべての条件を満たす用例が抽出される。この機能により、単語単位の検索に加え、特定の連語を対象とした検索も可能である。

### 4.2　コロケーション分析

コロケーション分析では、検索語の左 3 語および右 3 語の範囲について、位置ごとに共起語を確認できる。図 2 は、*at* を検索語とし、出現頻度順に結果を表示した例である。Left 1 では *good* が 44 回と最も多く出現しており、KISTEC では *at* の直前に *good* が頻繁に現れることが分かる。分析単位は語形とレンマから選択でき、“Top N collocates” 欄では、表示する共起語数を指定できる。

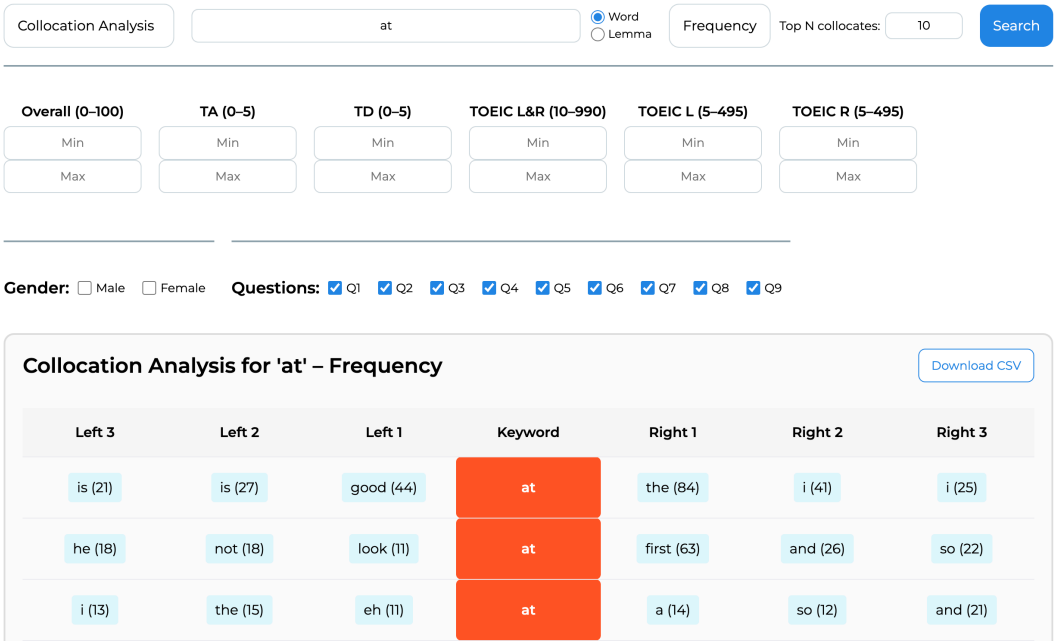


図 2　検索語を *at* としたコロケーション分析の例（出現頻度）

出現頻度に加え、*t*-score を基準とした表示も可能である（図 3）。*t*-score は、実際の共起頻度と、検索語および共起語の出現頻度から算出される期待共起頻度との差に基づき、共起関係の強さを評価する指標である。例えば、図 2 では、*at* の直前に現れる *look* と *eh* はいずれも 11 回であり、出現頻度は同じである。一方、図 3 では、*look* の *t*-score は 3.27、*eh* は −0.26 となっている。このように、共起頻度が同じ場合でも、各語の出現頻度から算出される期待共起頻度が異なるため、*t*-score には差が生じ得る。本機能は、KISTEC 内で頻繁に生起する共起パターンだけでなく、検索語との結びつきが相対的に強い語を探索するために利用できる。ただし、ある共起が学習者英語に特徴的であるか、あるいは不自然な組み合わせであるかを判断するには、母語話者コーパスなどの参照資料との比較が必要である。

### 4.3　語彙リスト作成

語彙リスト作成機能では、指定された範囲のデータに含まれる語形を、出現頻度の高い順に一覧化する。図 4 のようにコーパス全体の頻度表を作成することも、フィルターを組み合わせて、特定の設問、得点範囲、性別に限定した頻度表を作成することもできる。図 4 に示すコーパス全体の語彙リストでは、*i* が 13,109 回と最も多く出現し、次いで *is* が 10,908 回となって

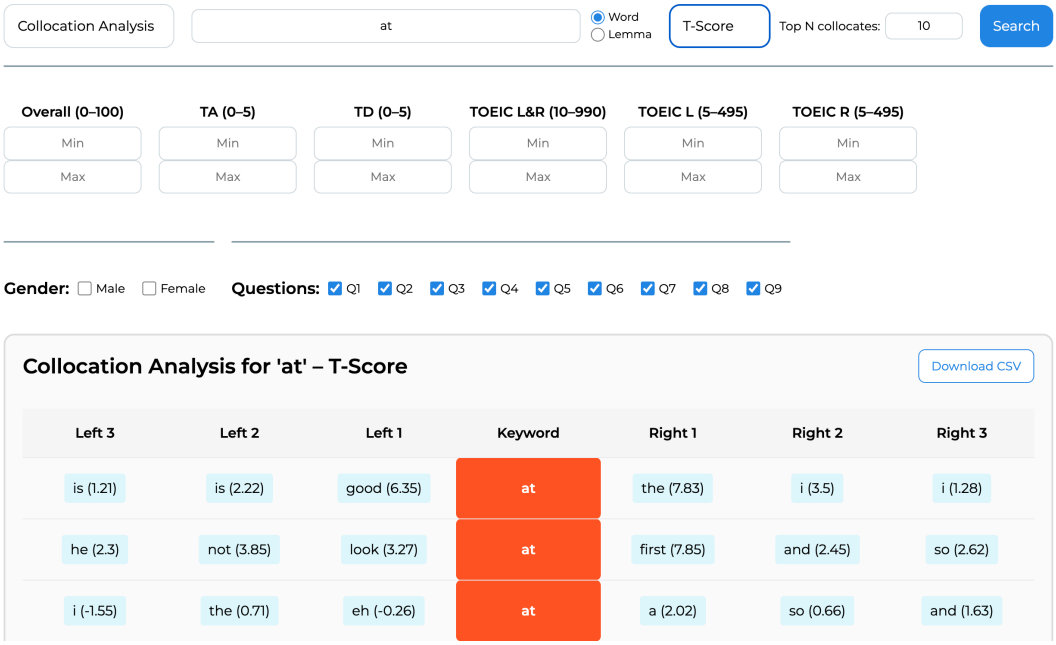


図 3　検索語を *at* としたコロケーション分析の例（*t*-score）

いる。例えば、TD 得点の高い解答と低い解答について語彙リストを別々に作成し、両者を比較することが可能である。

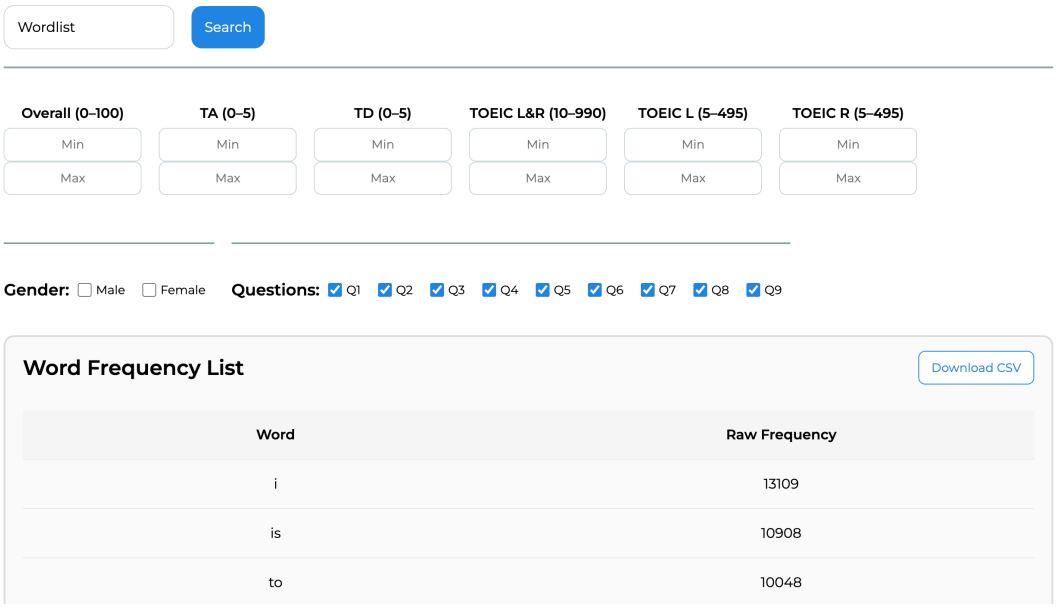


図 4　語彙リスト作成機能の出力例

頻度表は、対象データの語彙的な特徴を概観するうえで有用である。一方、設問番号によって課題形式や求められる内容が異なるため、得点層間の差を解釈する際には、比較対象となる設問番号を揃える必要がある。ただし、同一の設問番号でもテストバージョンによって具体的

な課題内容は異なるため、この点にも留意する必要がある。

### 4.4　特徴語抽出

特徴語抽出機能では、Overall score に基づいて受験者を Low、Mid、High の 3 群に分け、選択した群で相対的に特徴的な語を TF–IDF（Term Frequency–Inverse Document Frequency）により抽出する。各群の得点範囲は利用者が設定できる。図 5 は、0–40 点を Low、41–70 点を Mid、71–100 点を High としたうえで、Low 群の上位語を表示した例である。*eh*、*uh*、*ah* などのフィラーが上位に現れており、低得点群の発話にどのような語や表現が特徴的に現れるかを検討する手掛かりが得られる。

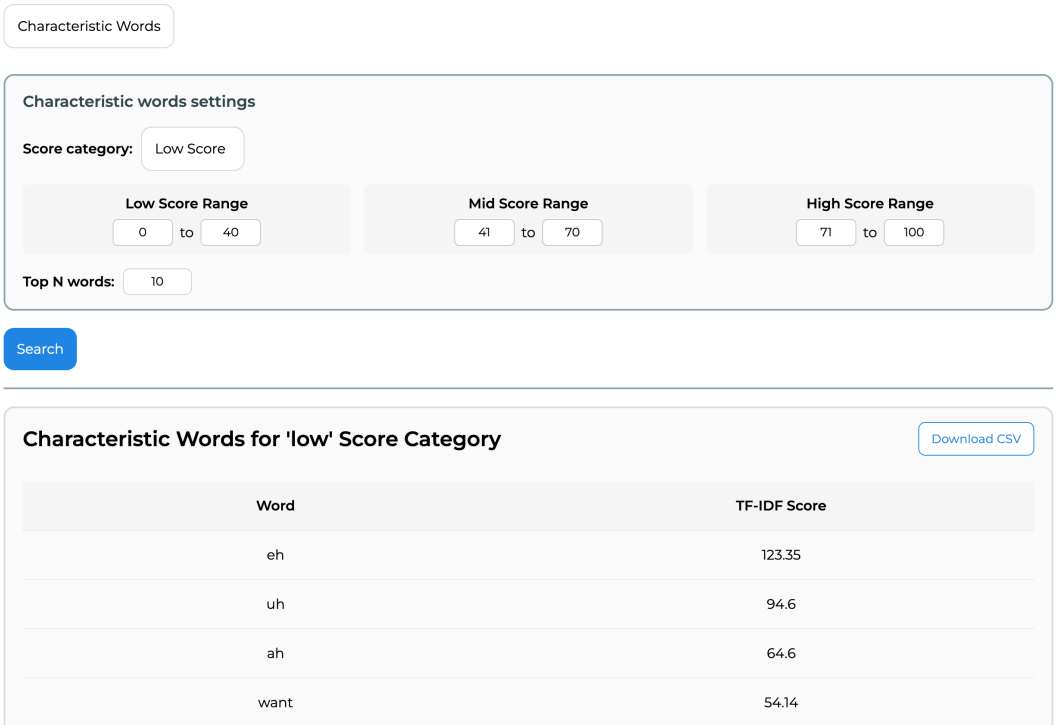


図 5　Low 得点群に特徴的に現れる語の抽出例

TF は対象群における語の出現頻度を反映し、IDF は複数群に広く現れる語の重みを小さくする。そのため、ある群で頻繁に用いられ、他群では相対的に一般的でない語が上位になりやすい。この機能は、詳細に検討する語の候補を探索するためのものであり、群間差の統計的有意性を直接検定するものではない。また、得点群ごとの人数、総語数、設問の分布が結果に影響する可能性があるため、抽出された語については、KWIC 検索によって具体的な用例を確認するとともに、必要に応じて外部ツールによる統計分析を行う必要がある。

### 4.5　N-gram 分析

N-gram 分析では、連続する $N$ 語のまとまりを出現頻度順に抽出する。例えば、$N = 2$ を指定した図 6 では、*want to* が 2,531 回、*i think* が 1,676 回出現しており、これらの語の連鎖が KISTEC で頻繁に用いられていることが分かる。利用者は、“N-gram length (N)” 欄で $N$

の値を、“Max results” 欄で表示件数を指定できる。

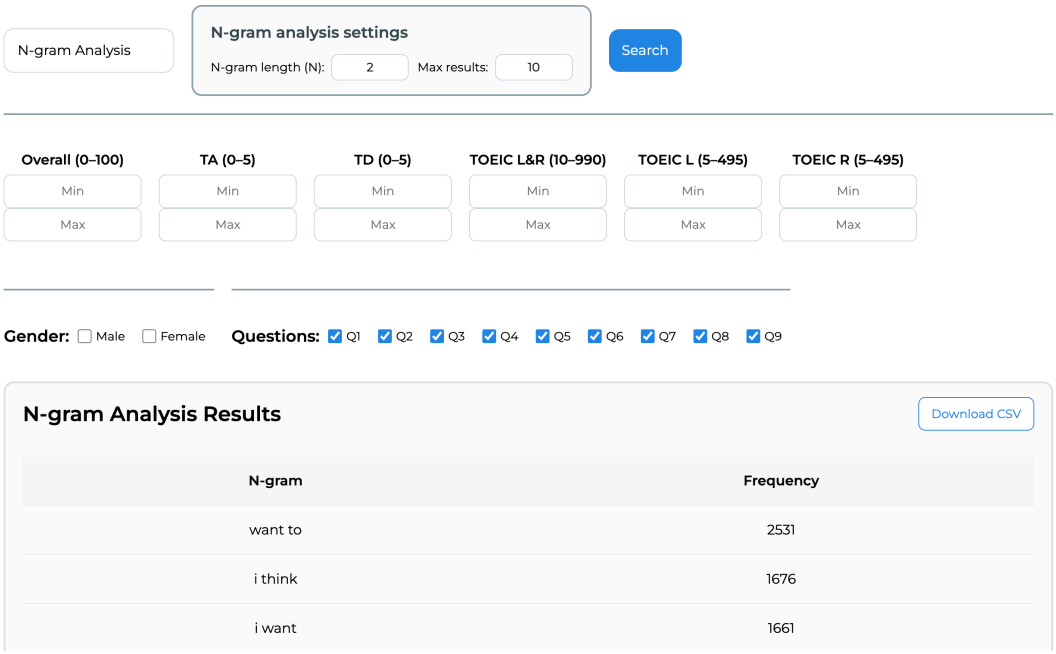


図 6　N-gram 分析の例（2-gram）

この機能は、個別語の頻度だけでは捉えにくい定型表現や構文パターンを探索する際に有用であり、得点層や設問ごとの使用傾向を比較することもできる。ただし、高頻度の N-gram には機能語の連鎖や、設問内容によって強く誘発された表現も含まれ得る。そのため、得点群や学習者集団の特徴として解釈する際には、設問別の分布と具体的な用例を確認する必要がある。

### 4.6　フィルタリングと CSV 出力

ここまでに紹介した機能のうち、特徴語抽出を除く各機能では、Overall score、解答別の TA 得点と TD 得点、TOEIC L&R の総合スコア、Listening スコア、Reading スコアについて、下限と上限の一方または両方を指定できる。さらに、性別や設問を指定することで、分析対象となる解答を絞り込むことができる。利用者は、各分析画面のフィルター欄に得点の下限や上限を入力し、必要に応じて性別や設問を選択したうえで分析を実行する。例えば、特定の設問に対する解答のみを対象としたり、TA 得点が一定以上の解答を抽出したり、TOEIC L&R スコアが特定の範囲にある受験者の解答を比較したりすることが可能である。

各分析結果は、分析の実行後に結果表示画面の “Download CSV” をクリックすることで、CSV 形式でダウンロードできる。KWIC 検索の CSV には、左文脈、検索語、右文脈、解答全文、TA 得点および TD 得点、Overall score などが含まれる（図 7）。これにより、検索された用例を評価情報と結びつけたまま保存し、外部ツールによる頻度集計や統計分析などの後続分析に利用することができる。なお、CSV に含まれる項目は使用する機能によって異なる。

| | A | B | C | D | E | F | G |
|---|---|---|---|---|---|---|---|
| 1 | left_context | keyword | right_context | full_sentence | TA | TD | Overall |
| 2 | Someone <JP | go | to <??></??> | Someone <JP | 1 | 1 | 24 |
| 3 | I think that m | going | <?>to</?></ | I think that m | 2.5 | 1 | 50 |
| 4 | I think that m | go | to university. | I think that m | 2.5 | 1 | 50 |
| 5 | <F>Eh</F>, | go | </?></SC> t | <F>Eh</F>, | 1 | 1.5 | 28 |
| 6 | The bicycle is | go | to the</SC> | The bicycle is | 1.5 | 1.5 | 31 |
| 7 | The bicycle is | go | to <nvs>laug | The bicycle is | 1.5 | 1.5 | 31 |
| 8 | <TO><F>Hm | going | to travel by ai | <TO><F>Hm | 1 | 1.5 | 26 |

図 7　KWIC 検索結果を CSV 形式で出力した例

## 5.　活用可能性と利用上の留意点

### 5.1　研究における利用

研究面では、語彙や表現の使用傾向を解答別の評価得点と関連づけて探索できる点が重要である。例えば、特定の語や定型表現が TA 得点または TD 得点の異なる解答でどのように用いられているかを比較したり、フィラーや自己訂正などのタグを含む用例がどのような文脈に現れるかを KWIC 検索によって確認したりすることができる。また、同一の設問番号に限定したうえで得点範囲ごとの語彙リストや N-gram を比較することで、課題形式や制限時間の違いをある程度抑えながら、評価得点との関連が示唆される語や表現を探索することも可能である。

ただし、本インターフェースが提供するのは、主として探索的、記述的な分析のための機能であり、得られた傾向のみから語や表現と評価得点との関係を直接検証することはできない。両者の関連を統計的に検証する場合には、設問や受験者によるばらつきなどを考慮した統計分析が必要となる。CSV 出力は、本インターフェース上で得られた傾向を、こうしたより詳細な分析へとつなげるためのデータ抽出手段として利用できる。

### 5.2　教育における利用

教育面では、特定の設問や得点層における実際の解答例を検索し、学習者が用いる語彙や表現、発話上の特徴を検討することができる。例えば、同じ検索語を含む用例を TD 得点順に並べることで、検索語そのものの使われ方だけでなく、その周辺の表現やフィラー、自己訂正などを、得点の異なる解答間で比較できる。このような用例は、得点の高い解答を模範例、得点の低い解答を誤用例と単純にみなすのではなく、課題の達成度や発話の伝達の効果性が解答間でどのように異なるかを検討するための教材として活用することができる。

一方、KISTEC に収録された解答には、文法的な誤り、不完全な発話、聞き取りが困難な箇所なども含まれる。そのため、教育利用に際しては、検索結果をそのまま模範的な用例として提示するのではなく、設問、得点、前後の文脈などを確認したうえで、教師が教育目的に応じて用例を選定し、必要に応じて注釈を加えることが求められる。

### 5.3 データと機能の制約

KISTEC は、一大学において一年度に実施された、1 年次学生を対象とする時間制限付きの CBT 方式スピーキングテストで得られた発話から構成される。したがって、検索結果に見られる傾向を、日本人英語学習者全般や高習熟度の英語学習者を中心とする集団、日常会話や対話場面にそのまま一般化することはできない。特に、設問内容や制限時間は、語彙選択や発話量に加え、ポーズやフィラーの出現傾向にも影響を及ぼす可能性がある。

また、現在のインターフェースは書き起こしを対象とするテキストベースの分析環境であり、音声の再生には対応していない。音声データはプライバシー保護の観点から一般公開されていないため、韻律や発話速度、正確なポーズ時間などの音声情報を検索結果から直接検討することはできない。さらに、結果の可視化、検索条件の自動記録、利用可能なメタデータを用いた絞り込み機能の拡充などは、今後の改善課題である。

## 6. おわりに

本稿では、KISTEC の検索および分析を支援する The KISTEC Search Interface の設計と主要機能を紹介した。本インターフェースは、KWIC 検索、コロケーション分析、語彙リスト作成、特徴語抽出、N-gram 分析を Web ブラウザ上で提供し、設問、性別、解答別の TA 得点と TD 得点、Overall score、TOEIC L&R スコアによる絞り込みと CSV 出力を可能にする。これにより、コーパスデータの取得や整形、検索に伴う技術的な障壁を低減するとともに、言語使用とスピーキング評価との関係を探索するための環境を提供する。

一方、本インターフェースから得られる検索結果や集計結果は、それ自体で因果関係や統計的有意差を示すものではない。しかし、検索条件を適切に設定し、具体的な用例の確認や外部ツールによる統計分析と組み合わせることで、学習者英語、スピーキング評価、英語教育に関する研究課題の発見や具体化に活用することができる。今後は、検索結果の可視化、検索条件やデータ版情報の記録、メタデータによる絞り込み機能の拡充などを進め、KISTEC の利用可能性をさらに高めていく必要がある。


## 謝　辞

本稿は、Tanaka et al. (2026) に基づくものであり、『英語コーパス研究』編集委員会より、出典を明記したうえで内容を再利用する許可を得た。記して感謝申し上げる。また、本研究の一部は JSPS 科研費 22K00736 の助成を受けた。


## 文　献


Katsunori Kanzawa, Yuichiro Kobayashi, Jaeho Lee, Haruhiko Mitsunaga, Masayuki Mori, Yusuke Tanaka, and Taishi Chika (n.d.). *The KIT Speaking Test Corpus (KISTEC)*. `https://kitstcorpus.jp/`.

Yusuke Tanaka, Ayaka Setoguchi, Taishi Chika, and Katsunori Kanzawa (n.d.). *Search Interface for the KIT Speaking Test Corpus*. `https://www.kistecsearch.org/`.

Yusuke Tanaka, Ayaka Setoguchi, Taishi Chika, and Katsunori Kanzawa (2026). “The

KISTEC Search Interface: A web-based interface for analyzing the KIT Speaking Test Corpus." *English Corpus Studies*, 33, pp. 161–170.

Yusuke Tanaka, Ayaka Setoguchi, Taishi Chika, and Katsunori Kanzawa (2025). "Moderate use of fillers can enhance fluency assessments." Bethany Lacy, R. Paul Lege, and Malcolm Swanson (Eds.), *Moving JALT into the Future: Opportunity, Diversity, and Excellence*. JALT. pp. 97–104.

Lucy Skidmore, and Roger K. Moore (2023). "BERT models for spoken learner English disfluency detection." *Proceedings of the 9th Workshop on Speech and Language Technology in Education (SLaTE 2023)*, pp. 91–92.

Laurence Anthony (2005). "AntConc: Design and development of a freeware corpus analysis toolkit for the technical writing classroom." *2005 IEEE International Professional Communication Conference Proceedings*, pp. 729–737.

Laurence Anthony (n.d.). *AntBNC Lemma List (ver. 004)*. `https://www.laurenceanthony.net/software/antconc/`.

**関連 URL**

| | |
|---|---|
| The KISTEC Search Interface | `https://www.kistecsearch.org/` |
| The KIT Speaking Test Corpus | `https://kitstcorpus.jp/` |