Supporting Instruction of Formulaic Sequences Using Videos at Scale

Kyung Je Jo KAIST Daejeon, South Korea kyungjejo@kaist.ac.kr Hyeonggeun Yun DGIST Daegu, South Korea yhg8423@dgist.ac.kr

Juho Kim KAIST Daejeon, South Korea juhokim@kaist.ac.kr

ABSTRACT

To help language learners achieve fluency, instructors often focus on teaching formulaic sequences (FS)—phrases such as idioms or phrasal verbs that are processed, stored, and retrieved holistically. Teaching FS effectively is challenging as it heavily involves instructors' intuition, prior knowledge, and manual efforts to identify a set of FSs with high utility. In this paper, we present FSIST, a tool that supports instructors for video-based instruction of FS. The core idea of FSIST is to utilize videos at scale to build a list of FSs along with videos that include example usages. To evaluate how FSIST can effectively support instructors, we conducted a user study with three English instructors. Results show that the browsing interactions provided in FSIST support instructors to efficiently find parts of videos that show example usages of FSs.

ACM Classification Keywords

H.5.m. Information Interfaces and Presentation (e.g. HCI): Miscellaneous; See http://acm.org/about/class/1998/ for the full list of ACM classifiers. This section is required.

Author Keywords

Formulaic sequence; video learning; language learning

INTRODUCTION

Formulaic sequences (FS)—multi-word phrases (e.g., be able to, in terms of) that are processed, stored, and retrieved holistically—are recognized as an important factor of language fluency as a large amount of natural language produced in daily conversation is found to be formulaic, rehearsed, and prefabricated [4]. Wray claims people belonging to any particular speech community have their preferred ways of saying things, which is reflected in the use of FSs [10]. Prior research estimates that more than half of native speakers' conversations involve FSs [2]. In contrast, non-native speakers not only have difficulties in attaining FSs but also show clear differences against native speakers in the use of FSs [4]. Thus, many instructional materials are focused on teaching FSs.

L@S '19 June 24-25, 2019, Chicago, IL, USA

© 2019 Copyright held by the owner/author(s).

ACM ISBN 978-1-4503-6804-9/19/06.

DOI: https://doi.org/10.1145/3330430.3333671



Figure 1. The overview of Formulaic Sequence List page. It shows the list of FSs which is extracted from large-scale YouTube videos. Initially, the list is sorted by the score of pedagogical value—written as "Total Score" in the figure. The user can browse FSs by applying filtering or sorting interactions based on frequency, mutual information, pedagogical value, or the number of videos.

To support effective teaching of FSs, prior studies have developed lists of FSs which have high utility and pedagogical value to learners [1, 7, 8]. However, there are two limitations to these lists. Firstly, the lists may reflect the language used by limited groups of people depending on the corpus which FSs were derived from. For example, Shin et al. claim that the nature of the corpus they used, the British National Corpus (BNC), largely reflects the colloquial language used by adult British [8]. Secondly, they are often plain lists of multi-word phrases and lack the richness of other useful information such as the context in which FSs are used or example sentences of FSs. Thus, they serve more value to instructors as references rather than direct teaching materials.

We see an opportunity in using videos at scale to combat the challenges above. Videos in foreign languages, such as TV programs or movies, are highly accessible sources of authentic language produced by native speakers. In recent years, the number of videos has explosively increased with the rise of online video platforms, such as YouTube, Netflix, and Vimeo. Videos offer a wide spectrum of contents that are targeted for various groups of interest (e.g., culture, age, language). In addition, videos effectively capture rich contextual information (e.g., gesture, emotion, setting) and a variety of expressions that have high utility as teaching materials. By utilizing videos at scale, we may be able to build a list of FSs that is flexible to provide contents that are appropriate to diverse groups of learners along with video contents that can be readily used as teaching materials.

Permission to make digital or hard copies of part or all of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for third-party components of this work must be honored. For all other uses, contact the owner/author(s).

In this paper, we present Formulaic Sequence Instruction Supporting Tool (FSIST), a tool that utilizes videos at scale to support instructors to effectively identify and create instructional materials for teaching FSs. FSIST consists of two components: 1) Formulaic Sequence List and 2) Video Selection. The Formulaic Sequence List page presents a list of FSs that are extracted from YouTube videos (Figure 1). Instructors can browse and select FSs that meet their needs based on various statistical information such as frequency, mutual information (MI)-a statistical measure of the associations between words in a phrase-, formula teaching worth (FTW)-a measure of the pedagogical value of FS that weighs frequency and MI [1]—, and the number of videos. Video Selection presents a list of example sentences of a FS along with the videos that the FS appears in (Figure 2). Instructors can select and preview parts of videos. They can also edit the length and format of the videos they wish to use as teaching materials.

In a qualitative study with three professional TESOL instructors, we asked participants to carry out a series of browsing and selection tasks as well as to freely explore the tool for usability evaluation. Participants found browsing interactions based on various statistical information of FSs as an efficient way to identify a set of FSs that satisfy their needs. However, they noted video browsing interactions can be improved to better support instructors with varying teaching goals (e.g., daily conversation, presentation, debate).

RELATED WORK

Prior research has introduced a number of approaches to effectively identifying pedagogically meaningful FSs. Many researchers argue that the frequency of FS is an essential criterion in the identification, but relying solely on frequency may not be sufficient as high frequency does not necessarily mean that it has a distinctive function or pedagogical value [1, 8, 7, 10]. To effectively identify FSs, existing studies have considered many factors such as MI [1], grammatical forms [8], and semantics [7, 10]. However, the approaches that consider grammatical forms and semantics of FS are often not suitable for working with large-scale data as they involve a significant amount of manual efforts from instructors to qualitative analyze text. To identify FSs from videos at scale, we use the data-driven approach introduced by Simpson et al. [1] that considers two types of statistical information, frequency and MI, to find pedagogically meaningful FSs.

FORMATIVE STUDY

To understand the challenges in preparing instructional materials for teaching FSs, we conducted interviews with three instructors who teach FSs as their main teaching goal.

Interviews

We interviewed three instructors [P1, P2, P3]. P1 (3 years of experience) posts instructional material on Instagram (with more than 100K followers) and Cake, one of the top-ranked language learning application in South Korea. P2 (7 years of experience) and P3 (8 years of experience) are professional instructors in one of the largest language institutions in South Korea. P2 has also been posting instructional material on Instagram and has more than 100K followers.

Each session took approximately an hour and included a semistructured interview. We asked our participants the following questions: (1) what is the general procedure of identifying FSs?, (2) what are the resources that are commonly used?, (3) how is the pedagogical value of FSs determined?, (4) what kind of information would help in determining the pedagogical value of FSs?, and (5) what is the general procedure of creating the instructional material?

Identification of FS

Rather than relying on the list of FSs that are pre-determined in the existing materials, all instructors preferred to manage a customized list of FSs to their teaching goals. The general procedure of instructors in identifying FSs is two-staged. Firstly, instructors create a list of FS candidates from videos such as movies and dramas. Then, they iterate and finalize the list to ensure all FSs have pedagogical value and high utility.

To create the list of candidates, the instructors said they watch a lot of videos and identify candidates that satisfy their own set of criteria. Frequency and authenticity of a FS were two commonly reported factors. P1 described his identification strategy as follows: "I don't always concentrate, but I have dramas or movies turned on throughout the day. Then, I make notes of phrases that I believe I hear frequently." Then, to ensure that the selected FSs are pedagogically meaningful, instructors finalize the list collaboratively with a group of assistants or other instructors. As the identification process is largely based on their intuition and prior knowledge, instructors wished for statistical evidence to support their judgments. Notably, to better support his judgment, P1 verified the frequency of the selected FSs with his own database of approximately 4,000 video transcripts by simple string matching of phrases.

Creation of Teaching Materials

The three instructors we interviewed used videos as the main teaching materials. For each FS they wished to teach, they prepared multiple videos that cover diverse usages of the FS. Similar to identifying FSs, instructors had their own set of criteria to select videos that meet the desired level of difficulty and offensiveness of language used in the video. To find videos, instructors often used existing websites that support easy browsing of videos based on phrases (e.g., Yarn, YouGlish). Instructors identified two key challenges of existing systems. Firstly, instructors often face difficulties in finding videos that satisfy their needs as the systems do not support efficient browsing interaction. For example, when a user searches for videos that include the FS, "one of the," Yarn displays a list of 3M videos with limited filtering options such as genres, production years, and title. In addition, these systems often do not support customized export in terms of length and output of videos. P3 often desired for videos of different length depending on the complexity of the context of videos to ensure that learners can fully capture the context of the videos, and thus manually downloaded videos and then used a third party application to edit them.

Overall, we observed that the instructors' process of identifying FSs and creating learning materials involved a significant amount of intuition, manual effort, and prior knowledge.

DESIGN GOAL

Inspired by our review of existing materials and interviews with instructors, we identified three key design goals for interfaces to support instructors in identifying, selecting, and teaching FS.

- G1: Provide quick ways of identifying a set of FS candidates
- G2: Provide statistical information of FS to support instructors to make better judgment and justifications on FS
- G3: Provide efficient ways to browse for videos that meet instructors' needs

SYSTEM OVERVIEW

In this section, we describe the components of FSIST: Formulaic Sequence List and Video Selection page. Then, we describe our dataset that enables FSIST.

Formulaic Sequence List

In FSIST, instructors can visually browse a large set of formulaic sequences from the list (Figure 1). Using the list, instructors can easily overview a wide range of FS candidates without having to manually watch individual videos. The list is designed to achieve **G1**: to enable instructors to efficiently identify FSs. Each listed FS shows statistical information such as frequency, MI, FTW, and the number of videos that include the FS. The list is sorted by FTW, by default, and it can be sorted by using any of the aforementioned factors. Instructors can use these factors to find FSs that satisfy their needs. For example, instructors who teach FS to beginners may wish to browse FSs with high FTW values as it means that they are used frequently and have the highest utility.

We show frequency and MI as they are the two key determinants that instructors value in the judgment of FS [1]. FTW is shown as it serves as a composite metric of frequency and MI. Our computation of FTW directly replicates the approach of Simpson et al. [1]. Such information display is added to meet **G2**: to provide statistical information to support a better judgment. In the list, frequency, MI, and FTW are normalized into the scale of 0 to 100 for more intuitive interpretation. Instructors can set filters on each components using two-bar sliders on the top. Statistical information of filtered FSs is shown on the top right. They can also sort FSs by clicking the header row of each column. Instructors can select a FS by clicking a row of the list.

Video Selection Page

After instructors have selected a FS, they can browse videos that include the FS in the Video Selection page (Figure: 2). Instructors can overview various metadata of videos that includes: the video title, the sentence which includes the FS, topics of the video, and the level of difficulty and offensiveness of the language used in the video. Our decision to show these particular components is informed by the existing guideline designed to inform instructors on selecting videos for foreign language teaching [9]. To identify the topics of videos, we use Empath, a tool that reveals topics based on text analysis [3]. We chose to use topics generated by Empath instead of the



Figure 2. In Video Selection page, instructors can overview, select, and edit multiple videos that include a particular FS. Instructors can also browse videos that satisfy their needs through filtering and sorting of topics, difficulty level, and offensiveness level. After instructors select videos, they can specify the format and length of videos.



Figure 3. Instructors can easily preview parts of videos that include usage of a FS by clicking on the title. The video preview appears as an overlay on top of the Video Selection page.

ones generated by video authors as we found large discrepancies in the amount and the specificity of the topics provided by them. Empath analyzes text across 200 human validated categories that include government, business, and technology. To assess the level of difficulty of videos, we considered the difficulty level of the speech FSIST. We used Flesch–Kincaid Grade Level Formula, even though it is not optimized for assessing the difficulty of speech, as it shows a fine-grained score which means the number of years of education generally required to understand a particular text. [5]. To assess the level of offensiveness, we used a publicly available list of offensive words created by Luis von Ahn's group ¹. We used the average occurrences of offensive words per video and then categorized the offensiveness level into low, medium, and high.

To browse for videos that satisfy their needs, instructors can filter and sort videos by topic, level of difficulty and offensiveness. Then, they can also preview parts of videos that include the current FS by clicking on the video title (Figure 3). Instructors can choose as many videos as they wish by checking the boxes on the first column. For those videos that instructors have chosen to use as teaching materials, they can click on "Export Selected Videos" to export videos. Then, they are asked to specify the length and output (e.g., URLs, multiple video files) of the videos. The design of Video Selection is

¹https://www.cs.cmu.edu/ biglou/resources/bad-words.txt

intended to meet G3: to support efficient ways for instructors to select videos of their needs.

Video Dataset

FSIST currently includes 1,447 YouTube videos. To avoid including videos that are inappropriate for educational purposes, we selected videos from voicetube.com, a language learning website which all videos are manually verified.

PRELIMINARY EVALUATION

We ran a qualitative study with three English instructors at our institution to investigate the potential effectiveness of FSIST. Instructors were given a 5-minute tutorial of the interface and asked to freely explore and evaluate FSIST. All instructors found statistical information of FSs presented by the system to be useful in identifying FSs that satisfy their needs. They said the sorting and filtering interactions would enable them to efficiently discover new FSs with stronger justification. However, they wished for more variety in interactions such as filtering FSs by difficulty. Instructors made an observation that FSIST would be most useful for efficient browsing of videos that meet their needs and include pedagogically meaningful FSs. They noted that an easy preview of videos and metadata provided in FSIST would help them save a lot of time finding related videos. However, they expressed that these metadata may not be sufficient. They said that instructors with different teaching goals (e.g., teaching daily conversation, presentation, debate) may wish for different types of metadata such as speaking rate or type of the language used.

FUTURE WORK

Video Browsing Interaction. As an immediate step, we plan to improve video browsing interaction through further user studies. Currently, FSIST supports video browsing based on three elements: topics, difficulty, and offensiveness level. From the preliminary user study, we learned that these elements may not be sufficient in supporting instructors to find appropriate videos for various teaching goals. In future work, we plan to consider other elements such as learner's need, the range of use (e.g., spoken, written), and teachability to develop video browsing interactions that effectively support instructors with diverse teaching goals.

Coverage of Formulaic Sequences. In the current design of FSIST, our approach to identify FSs relied entirely on FTW values. However, as shown in earlier studies, using different identification methodologies can result in discovering a different set of FSs [7, 8]. In addition, there also exist other types of language patterns that serve a similar role as FS in conversation such as "as ~ as" or "not only ~ but also ~." In future work, we plan to explore approaches to expand the coverage of formulaic sequences to include a more variety of FSs that may be useful to learners.

Curriculum Design Tool. Effective language teaching often follows a well-designed curriculum [6]. Designing an effective curriculum is challenging because it not only concerns selection of content (e.g., a set of FSs) that needs to be covered but also in what sequence the content should be presented. One exciting direction is to extend the current system as a tool to support designing an effective curriculum for teaching FSs.

CONCLUSION

This paper presents FSIST, a tool that presents a list of pedagogically useful FSs and videos to support language instructors in teaching FSs. The key idea behind our approach is to utilize large-scale videos to build the list of FSs and videos show the usage of FSs. Our preliminary user study with English instructors shows that FSIST can help them efficiently find videos that include the use of diverse FSs.

ACKNOWLEDGMENTS

This work was supported by Institute for Information & Communications Technology Promotion (IITP) grant funded by the Korean government (MSIT) (R7124-16- 0004, Development of Intelligent Interaction Technology Based on Context Awareness and Human Intention Understanding).

REFERENCES

- Nick C. Ellis and Rita Simpson-Vlach. 2010. An Academic Formulas List: New Methods in Phraseology Research. *Applied Linguistics* 31, 4 (01 2010), 487–512. DOI:http://dx.doi.org/10.1093/applin/amp058
- Britt Erman and Beatrice Warren. 2000. The Idiom Principle and the Open Choice Principle. *Text* 20 (01 2000). DOI: http://dx.doi.org/10.1515/text.1.2000.20.1.29
- Ethan Fast, Binbin Chen, and Michael S. Bernstein. 2016. Empath: Understanding Topic Signals in Large-Scale Text. CoRR abs/1602.06979 (2016). http://arxiv.org/abs/1602.06979
- I. Kecskés and L.R. Horn. 2007. Explorations in Pragmatics: Linguistic, Cognitive, and Intercultural Aspects. Mouton de Gruyter. https://books.google.co.kr/books?id=TfxhAAAAMAAJ
- 5. J.P. Kincaid. 1975. Derivation of New Readability Formulas: (automated Readability Index, Fog Count and Flesch Reading Ease Formula) for Navy Enlisted Personnel. Chief of Naval Technical Training, Naval Air Station Memphis. https://books.google.co.kr/books?id=4tjroQEACAAJ
- Jack C. Richards. 2013. Curriculum Approaches in Language Teaching: Forward, Central, and Backward Design. *RELC Journal* 44, 1 (2013), 5–33. DOI: http://dx.doi.org/10.1177/0033688212473293
- Norbert Schmitt and Ron Martinez. 2012. A Phrasal Expressions List. *Applied Linguistics* 33, 3 (04 2012), 299–320. DOI:http://dx.doi.org/10.1093/applin/ams010
- Dongkwang Shin and Paul Nation. 2007. Beyond single words: the most frequent collocations in spoken English. *ELT Journal* 62, 4 (12 2007), 339–348. DOI: http://dx.doi.org/10.1093/elt/ccm091
- 9. Stempleski Susan, Ed and Arcario Paul, Ed. 1992. Video in Second Language Teaching: Using, Selecting, and Producing Video for the Classroom. (01 1992).
- A. Wray. 2008. Formulaic Language: Pushing the Boundaries. OUP Oxford. https://books.google.co.kr/books?id=zTBnLAAACAAJ