





# Developing an Intelligent Tool for Computer-Assisted Formulaic Sequence Learning from YouTube Videos

#### Dr Phoebe Lin

Department of English
Hong Kong Polytechnic University

phoebe.lin@polyu.edu.hk

https://www.researchgate.net/profile/Phoebe Lin4

# Project information

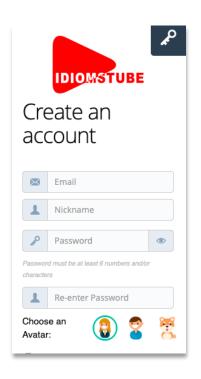
- Title: Developing an intelligent tool for computer-assisted formulaic sequence learning from YouTube videos
- Period: 1 December 2015 31 May 2019
- Ref: EDB(LE)/P&R/EL/164/17
- Keywords:
  - Computer-assisted language learning
  - Second language vocabulary acquisition
  - Internet television/media
  - Formulaic sequences
  - English-as-a-Foreign-Language

# Project objectives

- To develop a computer-assisted FS learning tool that automatically generates a range of pre-learning, knowledge consolidation and revision activities for any English videos that learners watch on YouTube
- 2. To assess the effectiveness of the tool by measuring learners' FS knowledge gains and analysing their usage and performance statistics
- 3. To examine Hong Kong learners' habits and performance when learning FSs from internet television/media
- 4. To inform future plans for local FS instruction and research

# IdiomsTube: Free open access

# To try the app, please create an account at www.idiomstube.com







#### Exposure to English outside classroom

- Exposure to English outside classroom is vital to the learning of English-as-a-Foreign-Language (EFL)
- Researchers have been exploring ways of increasing exposure to English outside the classroom (Nunan & Richards, 2014)
  - E.g. extensive reading, pen-pals, songs, movies, blogs, diaries, etc.

#### English learning through popular media

- Popular media is one of the most preferred and used selfdirected language learning strategies (Chapple & Curtis, 2000; Gieve & Clark, 2005)
- In a survey we conducted in July 2018 with 300 local university students, over 70% reported having a habit of learning English through popular media; 10% of them spend 5+ hours weekly on this activity.
- A longitudinal study by Verspoor et al. (2011) shows that exposure to popular media is effective at enhancing EFL learners' writing skills and vocabulary knowledge

# Popular media in ELT: A long history

Since 1980s, researchers have been exploring ways in which videos and popular media can be used effectively to enrich the English learning experience (Allan, 1985, Candlin et al., 1982; McGovern, 1983)

 In 1990s, satellite television brought new, exciting opportunities to EFL learners (Oxford et al., 1993; Meinhof, 1998)





# ESLprintables .com

Welcome to **ESL Printables**, the website where English Language teachers exchange resources: worksheets, lesson plans, activities, etc.

Our collection is growing every day with the help of many teachers. If you want to download you have to send your own contributions.



Worksheets and activities for te Friends series to English langua learners (kids, teenagers or adu Here you can find printable wor for many levels: beginners, eler intermediate or advanced.

All these worksheets and activit teaching Friends series have be designed by English language t

#### Friends series worksheets

#### <u>Cinema and Television</u> > <u>TV programmes</u> > Friends series



cartoon animated

videos, engaging



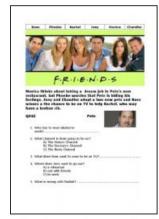
LET'S TALK ABOUT FRIENDS (SPEAKING SERIES 48) NEW



Friends - Tv Series (season 1 - episode 9)



MOVIES - LEARNING ENGLISH THROUGH MOVIES (5 pages) - 5



<u>Friends - Series 3 Episode 21 -</u> Chick and the duck

http://www.eslprintables.com/cinema and television/tv programmes/friends series/

# The age of internet television

 In the new millennium, we entered the age of Internet television and social media (Lin, 2014; Lin & Siyanova, 2014)



#### Internet media offer...

- Greater accessibility
  - Watch anytime anywhere
  - Ubiquitous language learning (Chang et al., 2011)
- More flexible playback options
  - Closed caption/subtitles
  - Replay
  - Play speed (normal, fast, slow)
- Wider choice of videos
  - 300 hours of video are uploaded to YouTube every minute!



# Challenges of internet media as input for English learning

- The number of internet media grows on an exponential scale
- Two new challenges has arisen:
  - 1. Learners face choice overload (Lin, 2014, in preparation)
  - Lack guidance and instruction. No teachers can design meaningful activities fast enough to guide English learning from internet television
- Possible solution? Computer-assisted language learning (CALL)



#### What is IdiomsTube?

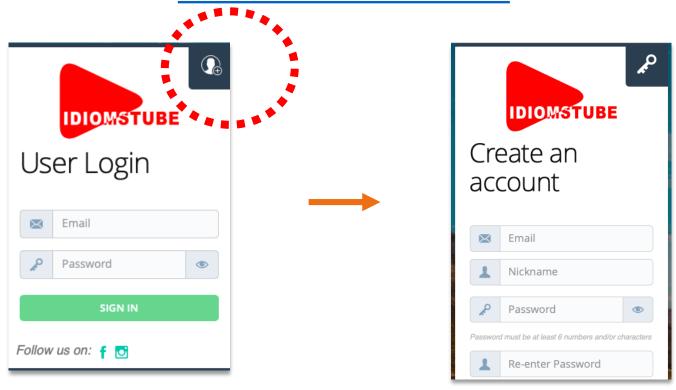


IdiomsTube is a web-based app that allows learners to expand their vocabulary while watching English YouTube videos

- The app was developed with the ultimate aim to enable the use of a big data approach to examine L2 vocabulary acquisition.
- The app focuses on the acquisition of formulaic expressions.

#### To try the app, please create an account at

#### www.idiomstube.com





# What are idiomatic expressions?

- Formulaic language is an umbrella term for all types of fixed or semifixed word combinations, including idioms, speech formulae, conversational routines, proverbs, clichés, sayings, collocations and so on.
- Formulaic expressions are functional units whose meaning can be explained only as functions of habitual usage (Kecskés, 2000).
  - It's raining cats and dogs
  - Here comes X
  - You're unbelievable!
  - Let's hear it for X
  - This paper is divided into X sections, etc.







# What are idiomatic expressions?

- Formulaicity is prevalent in everyday speech and writing.
  - Erman and Warren (2000) found formulaicity expressions account for 59% and 52% of the speech and writing samples they examined.



# Formulaic expressions are VERY difficult for EFL learners

- Formulaic expressions are often situation-bound. EFL learners often lack exposure to a wide enough variety of everyday contexts in English.
- Examples from the newspapers
  - X You missed the order. I have wait so long.
  - X Can you give me some food?
  - X Please check order again and again please
  - X Where is my order?
  - X Faster, faster, faster



- Some errors in the use of formulaic expressions have become fossilised and (almost) impossible to correct
  - X I'll now pass the mic to X

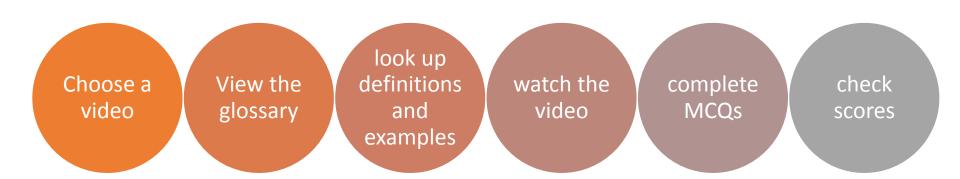
#### Internet media for formulaic language learning

- "No teacher, no classroom can provide the amount, the quality, the variety of language in interesting, meaningful, informative and often amusing contexts that television can" (Vanderplank, 1990, p. 221)
- A survey of a one-hour episode of the British television drama Holby City (series 15 episode 23) reveals that there are 86 dialogues covering a variety of communicative purposes and interlocutor relationships (Lin & Siyanova, 2014)

# IDIOMSTUBE

# Introducing *IdiomsTube*

- IdiomsTube is a web-based app that guides learners to learn idiomatic expressions while watching English YouTube videos
- The app automatically generates vocabulary exercises for any subtitled English video on YouTube





# IdiomsTube: Special features

	Self-directed learning	Learners watch videos they like and track their own progress
		Subtitle analysis
Ć K K	Automated processes	Extraction of expressions Hyperlinks to dictionaries Generation of exercises
	Large database	50,000 idiomatic expressions
16	Enhancements	Teacher's interface Gamification Free!



# How do you use IdiomsTube?

- Can be used at home or in a classroom, on a computer, tablet or smartphone
- Find videos by entering search terms or pasting YouTube URLs
- View the glossary, look up definitions and examples, watch the YouTube video, bookmark items, complete MCQs, check scores
- Check leaderboards, badges, video history and idioms diary
- Flip through flashcards for revision



### What does *IdiomsTube* offer learners?



Everyone can find the videos they like



Everyone can learn in a relaxing environment



- ✓ Just re-watch
- ✓ Auto-scoring



Everyone has complete control over how they want to learn

- ✓ Watch many times
- ✓ Alter playback speed



IdiomsTube fits into everyone's schedules.



Everyone is rewarded for their efforts.

- ✓ Accumulate points
- ✓ Get badges and upgrades
- ✓ Appear on the leaderboards



#### What does IdiomsTube offer Teachers?



Monitor students' performance with ease using the teacher's interface



Enrich classroom teaching

- Build discussions around the videos
- ✓ Suit students' interests



Recommend videos that complement classroom teaching



Assign a new type of exercise that is flexible and fun

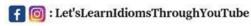
- IdiomsTube, a web-based app
  - Launched in November 2019
  - 1,342 registered users, of which, by inferring from the domains of the registered emails,
    - 8 are officials from the Ministry of Education, Singapore (email contains "moe.edu.sg")
    - 167 are university or school teachers/students/workers (email contains ".edu")
    - 347 from China, Hong Kong and Taiwan, 28 are from Singapore, 4 are from UK. There are also users from Canada, Russia, Vietnam, the Philippines, Germany, Brazil, India and Georgia.

- Invited talks to frontline teachers and students at PolyU library, CUHK English Language Teaching Unit
- Presentations at international conferences in Singapore and Hong Kong
- An English self-learning award scheme for students of Hong Kong Community College and PolyU School of Professional Education and Executive Development (SPEED)

"It's a great app and useful for me and my students which hopefully it will help us a lot on learning vocabs and idioms as well as speaking and listening." (Chaleunsouk Heminthavong, a teacher at National University, Laos)

- Project Instagram and Facebook sites
  - Launched December 2017
  - Published daily posts teaching the use of English idioms (just google!)

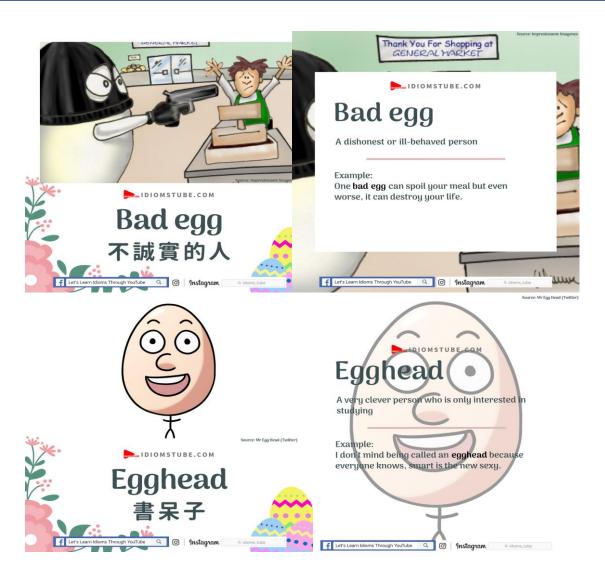






- Instagram:
  - https://www.instagram.com/idioms\_tube/
  - 360 posts, 1,194 followers
- Facebook
  - https://www.facebook.com/IdiomsTube/
  - 302 page likes, 327 followers





- Easter series
- **18-24/4/2019**
- Totals 5,305 likes on Instagram and Facebook

# Evaluating the effectiveness of the app

- A study was conducted to evaluate the effectiveness of the app IdiomsTube
- Research questions:
  - 1. Do the IdiomsTube auto-generated activities facilitate L2 formulaic sequence acquisition?
  - 2. Which input mode (video or reading) is more conducive to L2 formulaic sequence acquisition?
  - 3. What are the factors of L2 formulaic sequence acquisition?

# Research design

- Research question 1:
  - Learning through videos with vs without IdiomsTube automatically generated exercises
  - Between-subject comparison
- Research question 2:
  - Learning through videos vs through reading
  - Within-subject comparison
- Research question 3:
  - A logistic linear regression analysis of all the variables collected in the study

# Research design

А	В	С	D	E		
n=14	n=13	n=14	n=13	n=13		
Pre-test						
	-					
Read text B	Watch Text A	Watch text B	Read Text A	Watch text B		
	Self-guided reflection					
	-					
Watch Text A	Read text B	Read Text A	Watch text B	Watch Text A		
	Self-guided reflection					
Post-test						
Survey						

# Subjects

- 144 Hong Kong university students took the pre-test
- 67 subjects scored below 2 (out of 10) and were randomly assigned into 5 treatment groups
- Age: 19-25, mean=21.3
- No statistically significant difference between the 5 treatment groups in any of the variables measured

Pre-test/post-test: 20 MC questions assessing subjects' knowledge of the following formulaic sequences:

#### **Text A**

dressing-down
fall on deaf ears
smell a rat
away from prying eyes
on the wane

#### **Text B**

out of whack load up on something creep to a halt a recipe for something it happens to the best of us

#### <u>Foil</u>

have a champagne taste on a beer budget talk the talk, walk the walk press someone's buttons run-of-the-mill full of hot air take the bull by the horns have a chip on your shoulder see someone eye to eye have the last laugh don't sweat the small stuff

#### **Materials**

Example of an item in the pre-test/post-test

#### To "fall on deaf ears" means:

- A. no one answers one's phone call
- B. no one accepts one's offer of help
- C. no one listens to one's suggestions or requests
- D. no one believes in a story that one tells
- E. I don't know.

#### Materials

- Texts A and B are authentic texts about two interesting topics, the Olympics and traveller's constipation.
  - https://www.youtube.com/watch?v=I6h9055PQRw
  - https://www.youtube.com/watch?v=jLdZUWhSEIM
- Two versions (video and reading) were prepared for each text. Criteria:

#### **Videos**

- Duration: ~ 3 minutes
- Speech rate: ~140 words per minute

#### Reading passage

• Length: ~500 words

# Findings

- Research question 1: Do the IdiomsTube auto-generated activities facilitate L2 formulaic sequence acquisition?
  - Independent samples t-test results indicate IdiomsTube autogenerated activities make a significant difference in subjects' L2 formulaic sequence knowledge (t=4.64, p<0.001). Groups aided by the activities (mean=3.30, n=54) have significantly more L2 formulaic sequence knowledge gain than the group unaided by the activities (mean=1.15, n=13).

# Findings

- Research question 2: Which input mode (video or reading) is more conducive to L2 formulaic sequence acquisition?
- Videos (mean=3.30, n=54) generated more L2 formulaic sequence knowledge gain than reading (mean=3, n=54).
- Paired-sample t-test indicate no statistically significant difference between the two modes of input (t=1.28, n.s.).

# Survey responses

#### Between video and reading, which one

	Video	Reading
did you prefer for learning English idioms?	76%	24%
has helped you to learn idioms more effectively?	61%	39%
do you do more often in your free time (outside of this study)?	83%	17%
will you choose for improving your English proficiency in the future	69%	31%

n=54

# Survey responses

■ To what extent did you enjoy the activity? (5= Very much 1= None)

	Mean	sd
Seeing the list of idioms to learn before reading or the video	3.59	0.92
Reading the dictionary definitions of idioms	3.63	0.85
Watching the video	4.26	0.78
Reading the passage	3.06	1
Writing the short summary	2.67	0.85
Doing fill-in-the-blanks	3.87	0.78

n=54

Research question 3: What are the factors of L2 formulaic sequence acquisition?

#### correctAns ~.

- testset
- item
- textOrder
- text
- pretest.time
- posttest.time
- inputMode
- dictLookup

- read4Learn
- read4Leisure
- readEnjoy
- view4Learn
- view4Leisure
- viewEnjoy
- enjoyVideo
- enjoyRead
- difficultVideo
- difficultRead

- gender
- vocabSize
- DSEReading
- DSEWriting
- DSEListening
- DSESpeaking
- DSEOverall
- rateReading
- rateWriting
- rateListening

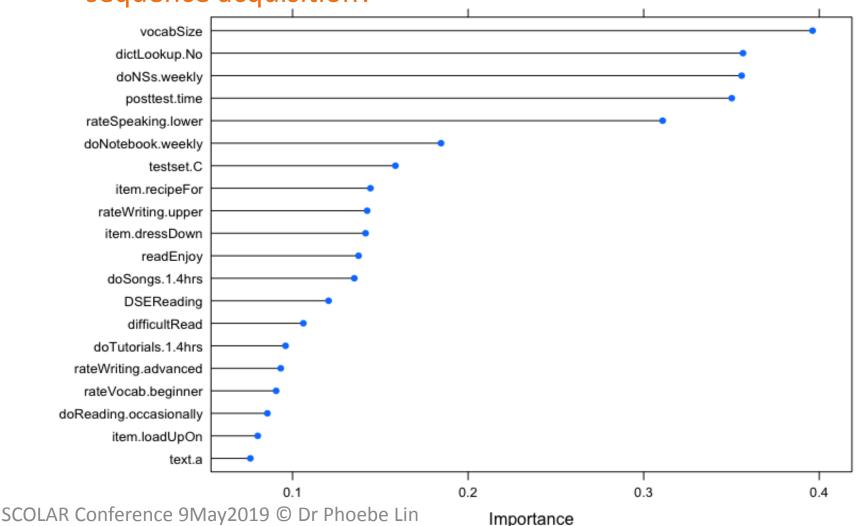
- rateSpeaking
- rateVocab
- doTutorials
- doReading
- doMovies
- doVideos
- doSongs
- doNSs
- doNotebook

 Fitted a penalised generalised linear regression model (i.e. glmnet) on the data using "caret" package in R

# Top 20 most important factors

Research question 3: What are the factors of L2 formulaic

sequence acquisition?



## Conclusion

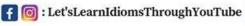
- The project developed an intelligent tool for facilitating L2 formulaic sequence acquisition from YouTube videos
- Achievements since the launch of the app in Nov 2018
  - 1,342 registered users
  - Our daily an-idiom-a-day posts
    - on Instagram: 1,194 followers
    - on Facebook: 327 followers
- A study was conducted to evaluate the effectiveness of the app (n=67)

## Conclusion

- A study was conducted to evaluate the effectiveness of the app (n=67)
  - Learners who learn idioms aided by IdiomsTube's automatically generated activities (mean=3.30, n=54) show significantly higher L2 formulaic sequence knowledge gains than the group unaided by the activities (mean=1.15, n=13), with t=4.64 (p<0.001).
  - Videos generated as much reading L2 formulaic sequence knowledge gains as reading
  - Learners have a clear preference for learning through videos than through reading
  - Vocabulary size is the most important predictor of L2 formulaic sequence acquisition

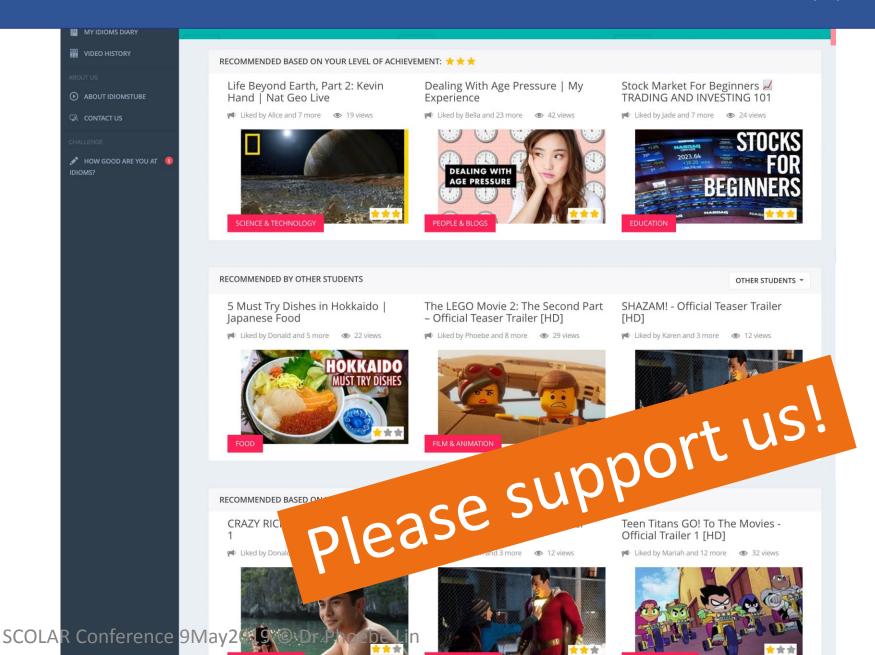
- Please share IdiomsTube (<u>www.idiomstube.com</u>) with your friends, colleagues and students about IdiomsTube
- It's totally free
- Please like our Facebook and Instagram to receive Idiom-of-the-Day posts







## To appear...



# Acknowledgement

We thank The Language Fund, The Standing Committee on Language Education and Research (SCOLAR), for their generous funding support to this research project 'Developing an intelligent tool for computer-assisted formulaic sequence learning from *YouTube* videos' (Ref: EDB(LE)/P&R/EL/164/17)



Questions and comments?

### References

- Aijmer, K. (1996). Conversational routines in English. London; New York: Longman.
- Allan, M. (1985). Teaching English with video. Harlow: Longman.
- Candlin, J., Charles, D., & Willis, J. (1982). Video in English language teaching: An inquiry into the potential uses of video recordings in the teaching of English as a foreign language. Birmingham: University of Aston.
- Chang, C.-C., Tseng, K.-H., & Tseng, J.-S. (2011). Is single or dual channel with different English proficiencies better for English listening comprehension, cognitive load and attitude in ubiquitous learning environment? *Computers & Education*, *57*, 2313-2321.
- Chapple, L., & Curtis, A. (2000). Content-based instruction in Hong Kong: Student responses to film. *System*, *28*(3), 419-433.
- Coulmas, F. (Ed.) (1981). *Conversational routine: Explorations in standardized communication situations and prepatterned speech*. The Hague: Mouton.

## References

- Gieve, S., & Clark, R. (2005). The Chinese approach to learning: Cultural trait or situated response? the case of a self-directed learning programme. *System*, *33*(2), 261-276.
- Hong Kong SAR Census and Statistics Department (2011). Results of the 2011 Hong Kong Population Census. Hong Kong: Hong Kong SAR Census and Statistics Department. Retrieved April 29<sup>th</sup>, 2013 from http://www.census2011.gov.hk.
- Hong Kong SAR Census and Statistics Department (2012). Use of new media. *Thematic household survey report number, 49,* 34-62. Hong Kong: Hong Kong SAR Census and Statistics Department.
- Lin, P. M. S. (2012). Sound evidence: The missing piece of the jigsaw in formulaic language research. *Applied Linguistics*, *33*(3), 342-347.
- Lin, P. M. S. (2014). Investigating the validity of internet television as a resource for acquiring L2 formulaic sequences. *System*, 42(1), 164-176.
- Lin, P. M. S., & Siyanova, A. (2014). Internet television for L2 vocabulary learning. In D. Nunan & J. C. Richards (Eds.), *Language learning beyond the classroom* (pp. 149-158). London: Routledge.

## References

- McGovern, J. (Ed.) (1983). *Video applications in English language teaching*. Oxford: Pergamon.
- Meinhof, U. H. (1998). *Language learning in the age of satellite television*. Oxford: Oxford University Press.
- Nunan, D., & Richards, J. C. (Eds.). (2014). Language learning beyond the classroom. New York: Routledge.
- Oxford, R., Young, P.-O., & Ito, S. (1993). Learning a language by satellite television: What influences student achievement? *System, 21*(1), 31-48.
- Vanderplank, R. (1990). Paying attention to the words: Practical and theoretical problems in watching television programmes with uni-lingual (CEEFAX) subtitles. *System*, *18*(2), 221-234.
- Verspoor, M. H., De Bot, K., & van Rein, E. (2011). English as a foreign language: The role of out-of-school language input. In A. De Houwer & A. Wilton (Eds.), *English in Europe Today: Sociocultural and educational perspectives* (pp. 147-166). Amsterdam: John Benjamins.