Intelligent CALL: Individualizing Learning Using Natural Language Generation

John Blake(✉)

University of Aizu, Aizuwakamatsu, Japan
jblake@u-aizu.ac.jp

Abstract. This chapter describes the theoretical underpinning, development and evaluation of an online natural language generation app, the Question Generator. This app individualizes language learning by creating interrogative statements from declarative statements using a natural language generation pipeline, enabling learners to create their own individualized practice activities. Learners can discover inductively how negation and auxiliary verbs are used in questions. A classroom observation of learners using the web app was conducted with junior high school students and university sophomores. Both groups were engaged and stayed on task with minimum supervision. The individualization appeared to motivate learners as they input sentences that were of interest to them. Learners were observed to be particularly active when working in pairs. The Question Generator is the first online tool that enables learners to generate questions based on user input, and thus breaks new ground in the growing set of intelligent CALL tools.

Keywords: Computer-assisted language learning · Natural language generation · Interrogative statements · Question formation · Individualized learning · Discovery learning

1 Introduction

Teachers of English strive to meet the individual needs of learners. When teaching individuals or small groups, teachers are more able to tailor the lesson content and materials to meet the needs of each learner. However, providing tailored content becomes much more difficult as group size and group diversity increase. In mixed-level classes, class teachers may individualize materials by providing different content and/or different tasks to learners. This level of dedication is laudable, but it requires a significant time commitment from teachers.

The problem, in short, is how to personalize materials to suit the needs of each learner. This may involve providing the: (1) same content (e.g. a reading or listening text) with different tasks, (2) different content with the same task, or (3) different content and different tasks. All three options involve additional work by the materials developer. For paper-based materials, teachers most likely need to prepare and print the materials prior to class. However, for resources that are available online and open-access, teachers and students simply need access to a wifi-enabled device. The development of online

A. W. B. TSO et al. (Eds.): HKAECT 2022, ECTY, pp. 3–18, 2023.
https://doi.org/10.1007/978-981-19-9217-9_1

resources that adapt to user needs and levels takes significant programming ability and an equally significant time cost. However, by harnessing resources that have been developed and deployed online, teachers and students can use them with little or no preparatory work.

The golden chalice of materials development is on-demand generation of individualized or personalized materials that meet the needs and expectations of learners. This involves creating a software program that can output personalized materials for each user. By harnessing natural language generation (NLG), materials may be created from user input. Using NLG provides more control over the language than using authentic texts. NLG software can pitch the materials generated accurately to the level of the learner based on the parameters and vocabulary sets specified in the software.

This chapter shows how natural language generation can be used to create materials related to the aims, level and interests of learners. An early prototype of a web-based NLG application, the Question Generator, serves as the vehicle to illustrate how NLG can be used to provide learners with relevant examples.

The remainder of this chapter is organized as follows. The importance of exposure to language and the necessity to notice specific language features are discussed in the following section. The trend to adopting more interactive and individualized learning approaches is then discussed in the subsequent section. The move from computer-assisted language learning (CALL) to intelligent CALL is described next. The focus is then narrowed to NLG, one domain within the broader field of natural language processing. Having established the theoretical underpinning for adopting NLG for pedagogic purposes, the Question Generator is then introduced in detail. The design, development and deployment of this online tool is described in a (relatively) non-technical manner. The method and results of the classroom observation of pilot testing with high school and university students are detailed. The extent to which the Question Generator meets the needs of the learners during the classroom observation is reported. The chapter finishes with a prediction that intelligent CALL is set to be a game-changer in materials development in the same way that online dictionaries and machine translation created a sea change in the way that language learners read and learn online.

2 Background

Children acquire fluency in the languages to which they are exposed. Krashen explains this using the concept of comprehensible input (Krashen 1982), asserting that vocabulary is acquired through reading (Krashen 1989). According to Piaget (1956), the breadth and depth of the language proficiency of children ranges widely depending on the types of interactions they are involved in. Likewise, adult language learners require exposure to learn a new language. Adults raised in monolingual environments with little interaction with speakers of the target language tend not to develop productive fluency. This is particularly true in countries, such as Russia and Japan, where there is little need or opportunity to use English in daily life. In short, exposure to input-rich language environments provides opportunities to learn the language; conversely, with no exposure, there is no opportunity to learn (Griffiths and Soruç 2018). The importance of exposure underpins the approaches in many language learning textbooks and second

language learning classrooms. Every classroom and every coursebook includes the target language even if rubrics and explanations are provided in a different language. A widely-used approach in second language teaching is for the teacher and/or the teaching materials to systematically expose learners to target language items, such as vocabulary sets or grammatical structures (Spratt et al. 2005). Nowadays, the target language items are commonly contextualized in textbooks using authentic and semi-authentic written texts (Little 2014).

When target language is presented in context, learners may be given comprehension questions to guide them to understand the gist of the whole text and specific meanings of key parts of the text (Hondo 2015). Despite skimming and scanning the text to answer the comprehension questions, learners may not have paid much attention to the target language itself. In communicative language teaching, learner attention is often drawn specifically to the target items by first focusing on its meaning, then its structural form, pronunciation and appropriacy of use (Isaacs 2009). Simply put, learners are guided to notice the target language. This noticing is said to give them the opportunity to learn the language. Consider how you would say *thank you* in a language that is extinct. Would you not have to see or hear it, before you could learn it? However, with some extinct languages there are no audio or textual records, and so these languages are destined to remain extinct. The noticing hypothesis proposed by Schmidt (1990) claims that noticing language features is a necessary precursor to learning the language features. In short, if learners are exposed to a particular language feature in context, but do not pay attention to that feature, the learners will not learn the language feature. However, if learners consciously notice the target feature, then there is the possibility that they may learn that feature. In a nutshell, failing to notice rules out learning. The noticing hypothesis has received criticism because of the lack of theoretical basis and the difficulty to test the hypothesis (Truscott 1998). Despite this, the day-to-day activities of many language teachers appear to focus learners on target features to help students learn those features.

If we accept the proposition that it is necessary to get learners to notice language features to help them learn or acquire those language features, then the dilemma for the teacher is how to get the learners to notice the target features most effectively and most efficiently. Most textbooks for language learning focus on the written form of language features. Even when a word or structure is presented in a listening text, students are frequently asked to examine the written transcript. For example, when helping learners understand how to formulate closed questions, learners need to be exposed to the various question forms. Students hearing (or reading) *"Are you okay?"* might be able to generalize the form to *"Are you ready?"* but may overgeneralize to *"Are you understand?"* if they failed to notice that the auxiliary verb *be* is used to create closed questions when the predicate is an adjective. Learners need to notice how different auxiliary and modal auxiliary verbs are used to be able to ask grammatically-correct questions with different main verbs in different tenses. This is usually systematically dealt with in course materials, often with the help of mainstream textbooks. Textbooks invariably introduce the more commonly-used tenses, such as *present simple* and *past simple* before the less-frequently used and more complex tenses, such as *past perfect* and *future progressive* (Aitken 2021).

3 Learning and the Individual

Textbooks used in language learning classrooms reflect the fashionable orthodoxies at the time of press (Harwood 2014). Many language schools base their curriculum on textbooks and many teachers base their classes on the approach and activities contained in the textbooks. Language teaching approaches and methods have changed over the years. Grammar translation dominated language learning for most of the last two millennia (Chang 2011). However, in the last century multiple new teaching approaches were introduced including direct method, audiolingual, communicative and post-communicative.

A general trend that can be identified among these approaches is the move from lockstep classroom teaching in grammar translation to a focus on interaction (Long and Porter 1985) in communicative and post-communicative approaches. Although learning can occur without teaching, and teaching can occur without learning; our hope as teachers is that there is some correlation between our teaching and student learning. In an effort to increase the amount of learning, students are encouraged to more actively participate. In language classes, active participation tends to be achieved by encouraging students to work together through the use of pairwork and groupwork. Teachers set students tasks or activities to work on with their partner(s). These tasks may be classed as information or communication gap activities (Nakahama et al. 2001). Language labs gained in popularity in the 1950s and 60s (Alexander 2007). Students in language labs tended to interact with audio materials, which often involved shadowing or responding to audio prompts (Hamada 2019). The use of technology in the language labs enabled learners to practise at different levels simply by selecting materials that were pitched at different levels. There was a strong movement to self-access centers in the 1980s and 90s, which again changed the type of interaction. Learners were now being encouraged to study autonomously without regular access to a teacher (Benson 2001). Although pairwork and groupwork is still encouraged, most learners availing themselves of these centers interact with online resources. The Centre for Independent Language Learning website[1] developed by Andy Morrall at Hong Kong Polytechnic University was one of the first online resources that provided learners with multiple client-side interactive activities.

Ubiquitous access to the internet, affordable wifi-enabled devices and a relatively tech-savvy generation of language learners who grew up playing online games, relying on search engines and checking social media sites has again altered the language learning technoscape (Appadurai 1996). The internet is a rich source of input with an infinite stream of reading and listening texts. Learners who need to check or look up information on grammar or the meanings of words are spoilt for choice.

The advent of intelligent CALL gives both teachers and learners the opportunity to generate their own practice materials on demand. This obviates the need for materials developers to produce a vast bank of materials from which materials for each individual learner can be selected. By combining the power of artificial intelligence with genre-specific templates based on corpus studies, tailormade texts can be generated on demand. Until now, texts have been categorized into the authentic or non-authentic

[1] Https://elc.polyu.edu.hk/cill/.

dichotomy; but perhaps the human-created or computer-created will become a more important dichotomy in years to come as the number of computer-created texts increases.

Students learning languages in classes tend to be grouped together by age and/or ability. In most cases, the whole class studies the same language points using the same material. For example, Japanese students attending junior high school are taught the same content in the same manner with the same material regardless of the interests and ability of each student. Japanese school students are relatively homogenous, having followed the same school curriculum taught using government-approved textbooks. Many classrooms worldwide have far lower levels of linguistic, cultural, and social homogeneity. Individualized learning is achieved by matching the content, mode and pace of learning to each learner. Although the Japanese junior high school students are likely to share the same mother tongue and many cultural values, each student has different learning styles, learning preferences, interests and language competences. Individualized learning aims to tailor learning to each learner profile. For example, materials may be offered at slightly different levels with learners able to self-select the level appropriate for themselves, which enables learning to be individualized based on language competence. In an attempt to cater to learning style preferences, learners may be provided with materials that reflect their preferences. To facilitate individualized learning, technology is often used. By harnessing both machine learning and computer-assisted language learning, learning may be personalized to each learner.

4 Intelligent CALL

Computer-assisted language learning (CALL) made inroads into the classroom in the 1980s, but took off in the 2000s. Nowadays, increasingly more language learners are using their mobile phones to access learning materials online. The increase in ownership of mobile devices and the ubiquity of wireless connections to the internet resulted in a new subdomain of research and practice, namely mobile-assisted language learning (Chinnery 2006). The mobile-first approach (Mullins 2015) in which web apps and webpages are designed first to function on the smaller viewport of mobile devices is now standard practice for website developers.

One of the domains gaining popularity recently under the umbrella of CALL is intelligent CALL (iCALL) (Volodina et al. 2012), which harnesses natural language processing for language learning purposes (Gamper and Knapp 2002). However, it should be noted that iCALL is over three decades old. A summary of the research and practices of iCALL was published back in the late 1980s (Bailin et al. 1989). However, the high cost of computers, limited access to the internet and the technological knowledge hurdle were barriers that hindered its transition to mainstream language learning. The combination of recent breakthroughs in machine learning combined with increasing accessibility to the internet are a driving force for iCALL.

Language learning applications harnessing iCALL have been created to help learners with pronunciation (Boitsova et al. 2018; Bogach et al. 2021; Qian et al. 2018), grammar (Ward et al. 2019; Purgina et al. 2020) and writing (Chukharev-Hudilainen 2019; Chukharev-Hudilainen and Saricaoglu 2016). All of these learning applications use natural language processing to search texts for patterns, but none of them uses natural language generation.

One natural language generation tool that is utilized in classrooms, but was not developed for language learning is a Conversational User Interface (CUI). There are two types of CUI: chatbots and voice assistants. The early version is the chatbot. Chatbots engage in conversations with humans (or other chatbots) by generating responses to declarative statements and answers to interrogative statements. Chatbots are designed to simulate human-like conversation using text messages and have been used in CALL (Wik and Hjalmarsson 2009). Chatbots use artificial intelligence when generating responses and answers. More recently, voice assistants communicate via voice user interfaces, such as Alexa (Amazon), Google assistant (Google) and Siri (Apple). Open access to natural language processing (NLP) platforms, such as Dialogflow (Google)[2], allow educators to tailor Conversational User Interfaces for language learning. Such CUIs can be designed using a combination of general modules offered by the platform and specific modules that relate to the needs of the learners. Voice-driven artificial intelligence assistants (e.g. Alexa) can be trained to interact with learners in a variety of contexts, which can provide language learners with valuable fluency practice (Dizon 2017; Dizon and Tang 2020). However, because of the nature of machine learning, it is difficult to understand how the speech is generated. Natural language generated by rule-based parsing is easier to control and can provide learners of English with practice activities tailored to their specific interests and needs.

Data-driven learning (Boulton 2017; Hadley 2002) is one way in which learners can be exposed to particular language features. The most common source of data is from a corpus (e.g. Braun 2007; Chujo et al. 2012) which may be accessed using a concordancer, such as the popular open-access free-of-charge downloadable AntConc (Anthony 2022) or via an online subscription platform, such as Sketch Engine (Kilgarriff et al. 2014).

A novel way to provide learners with data is by creating texts automatically using natural language generation (NLG) pipelines. According to McDonald (2010, p. 148), natural language generation started to flourish in the 1980s as a subfield of computational linguistics. One of the first types of NLG was the creation of random sentence generators, which were developed to check grammars.

There are two main approaches to NLG: rule-based and probabilistic (Blake 2020). With a rule-based approach the software developer has greater control over the syntax and lexis used in the generated text, which may help the developer pitch the level of the text to its target users more easily. With a probabilistic approach, machine learning is used and so the training data greatly impacts the generated text. Any biases or idiosyncrasies in the training data are likely to be reflected in the texts produced. The software developer has less control over the output produced by machine learning approaches. Machine learning, including deep learning, are so-called black box approaches, since it is not possible to see or understand exactly how the system works (Rubin 2019).

Language awareness raising activities help learners notice linguistic features. The practices of many language teachers are founded on the principle that noticing is a precursor to learning (Ellis and Mifka-Profozic 2013). NLG is a scalable method to provide controlled individualized practice that help language learners notice language features. NLG for language learning is a relatively new innovation. This nascent application of pedagogic NLG is firmly situated on the cutting-edge of intelligent CALL.

[2] Https://cloud.google.com/dialogflow/docs/.

5 Question Generator

The inspiration for this tool stemmed from a controlled practice activity used in English language classrooms, in which learners create as many questions as possible based on a given sentence. This allows students the choice of selecting which question to ask, yet given the limited number of choices, teachers can anticipate the types of mistakes that learners make. For example, using the input sentence in (1) and limiting the question types to open questions, two different types of questions (i.e. subject and predicate) may be asked. Subject questions involve replacing the grammatical subject with an interrogative pronoun, while predicate (or object) questions use inversion and require an appropriate auxiliary verb.

(1) I woke up on the sofa at seven o'clock this morning.
Students can select from four different interrogative pronouns (i.e. who, where, what and when) to create questions such as those shown in examples (2) to (5).
(2) Who woke up on the sofa at seven o'clock this morning?
(3) Where did you wake up at seven o'clock this morning?
(4) What time did you wake up on the sofa this morning?
(5) When did you wake up on the sofa at seven o'clock?

Examples (6) to (9) show some of the common mistakes made by Japanese learners of English. In example (6) the tense of the main verb should be past simple and not present simple. In example (7) the auxiliary verb *did* is omitted. Tense is a problem again in example (8) as both the auxiliary and main verb carry past tense. The auxiliary continues to cause problems in example (9), but this time the issue is using the incorrect auxiliary verb and incorrect tense.

(6) Who wake up on the sofa?
(7) Where you wake up this morning?
(8) What time did you woke up on the sofa this morning?
(9) When are you wake up on the sofa this morning?

Practice activities like this can be personalized by allowing learners to supply the initial statement. This provides learners with the freedom to select the topic, length and complexity of the declarative statement, which in turn affects the questions.

All question types present challenges for learners of English. The three main question forms are open-ended, closed-ended and tag questions. Open-ended questions can be subdivided into questions that ask about the subject or the predicate, each of which adhere to different grammatical rules. Many learners of English struggle to create grammatically accurate questions. There are numerous reasons for this, including the necessity to manipulate the syntax of a declarative statement into the appropriate form for an interrogative statement. Different types of verbs (e.g. stative, dynamic) behave in different manners. Tense, voice, aspect and modality of the finite verb phrase affect the syntax used in questions.

Prototype proof-of-concept natural language generation tools were created by computer science majors in the University of Aizu enrolled in a credit-bearing elective course

on computational linguistics. The course followed a Content and Language Integrated Learning (CLIL) approach with the two-fold aim of developing both content knowledge and language proficiency. In terms of content, learners were required to analyze linguistic patterns occurring in different language systems (e.g. phonemic, syntactic) and master powerful search expressions used in programming called regular expressions. A number of prototype web apps were created. These including a trend description generator, a closed question responder and a very basic question generator. The trend description generator automatically creates a descriptive text from data values, which may also be used to create a graph or bar chart. The closed question responder provides positive and negative short answers to any closed question. A number of initial question generators were created, each of which focused on one type of question. These prototypes informed the development of a more sophisticated question generator that is the focus of this section.

Having established the feasibility of turning the inspiration of a question generation tool into a web application, the next step was to create a high-fidelity prototype to help Japanese learners of English learn the form of various question types by providing unlimited examples.

The Question Generator is the first release of a natural language generation web app. A technical description of the design and development of the Question Generator is available in Vu and Blake (2021). Although many NLP researchers work on problems related to question answering, almost all focus on extracting answers to questions from texts. However, this intelligent CALL tool generates questions rather than answers. Although there is a substantial amount of research on question generation presented at top-tier computational linguistics conferences, there is a lack of operational deployments. Apart from releases from our laboratory, no online question generation tools were discovered at the time of writing. An extensive search of the literature revealed no reports of any system specifically created to generate questions based on an input sentence for language learning purposes. Systems that rely heavily on neural networks to create sentences may provide authentic examples, but the complexity and vocabulary of the generated questions is not graded to the level of the input sentence, potentially making the questions too difficult to understand. The Question Generator therefore breaks new ground by being the first online tool to generate questions from user input to help learners of English.

The Question Generator creates interrogative statements from declarative statements using natural language generation. This is achieved through an extensive set of transformation rules supplemented by machine learning to generate appropriate questions. This system enables learners to create their own individualized controlled practice activities. Learners can input a sentence and the web app automatically generates three types of questions: closed-ended, open-ended and tag.

By comparing the input and output, learners can raise their language awareness of syntax either inductively or deductively (Shaffer 1989). For example, through interacting with the web app, learners can discover how negation and auxiliary verbs are used in tag questions, and induce the rules that govern the syntax. When used deductively, learners can create questions based on an input sentence, and then compare the questions that they have created with those created by the Question Generator. They can then analyze and evaluate differences between the questions to determine which (if any) are inaccurate.

To use this web app, users simply need to navigate to the website, input a declarative statement and select the required question type. Learners can select to generate closed-ended, open-ended and/or tag questions. Figure 1 shows a screenshot of the graphical user interface of the web app for a very short simple sentence taken from the popular coursebook Headway Beginner.

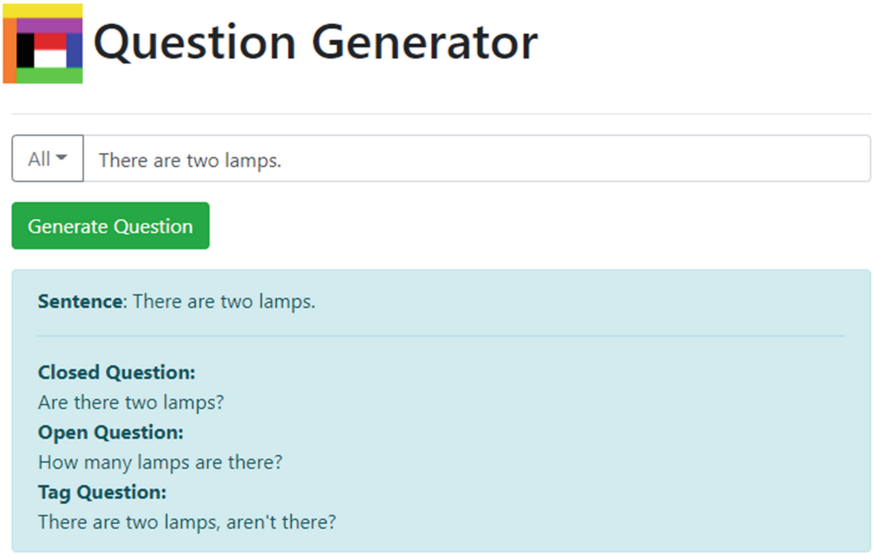

Fig. 1. Screenshot of the output for "There are two lamps."

Figure 2 shows how the Question Generator deals with a slightly more complex sentence. In this case, the grammatical subject comprises two entities, namely Mary and Tom. This needs to be resolved to the plural pronoun *they* to create a tag question.

Figure 3 shows the result for a simple sentence that contains multiple adverbials.

In addition to the necessity to identify time and location adverbials in order to determine the most appropriate interrogative pronoun, the fronting of the time adverbial *Next week* poses another challenge. In the question forms, time adverbials are usually placed at the end of the clause rather than before the interrogative pronoun. However, it should be noted that in the third open question, a rather formal and marked question is generated due to the fronting of the prepositional phrase *From Canada.* A more appropriate question form would be: *Where will my aunt be flying to next week?*

Automatic question generation is technically challenging for a number of reasons. Two particularly notable issues were the selection of syntax and pronouns (interrogative and personal). The primary substantial challenge is to be able to create the correct syntax for the question form being constructed. Transformation rules were created to map each of the three question forms to generic text inputs. A detailed and rather technical explanation of the way in which transformation rules are applied is given in Vu and Blake (2021). The transformation from declarative to interrogative statement is exacerbated by

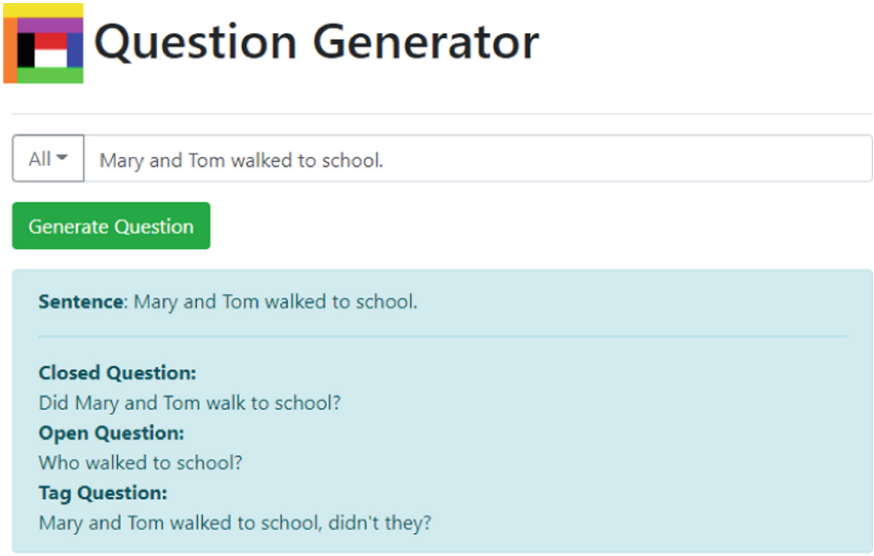

Fig. 2. Screenshot of the output for "Mary and Tom walked to school."

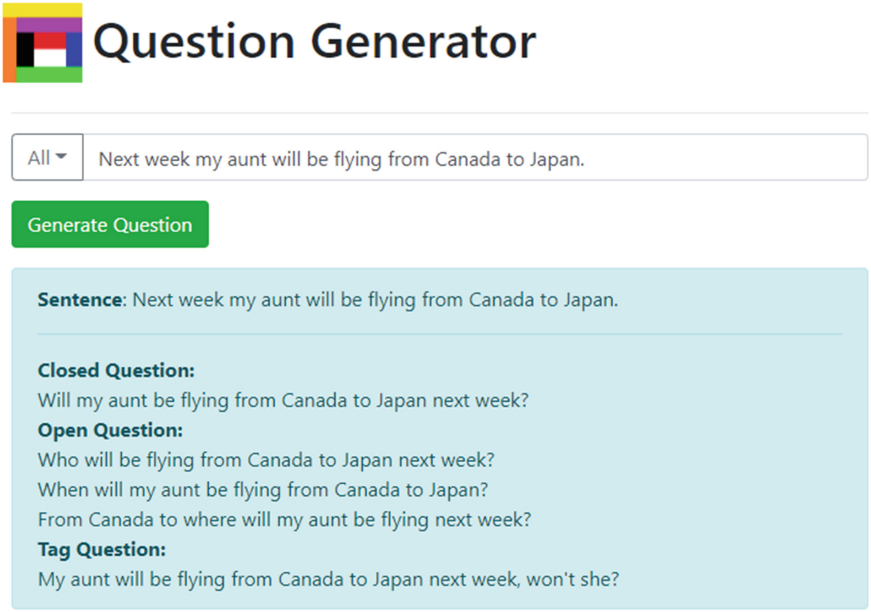

Fig. 3. Screenshot of the output for "Next week my aunt will be flying from Canada to Japan."

having to deal with tense, voice and aspect, each of which impacts the form of the question. The automatic selection of appropriate interrogative pronouns is also problematic. There is a plethora of research on pronoun resolution, in which the pronoun is mapped to

the entity it represents. However, there is a paucity of research on pronoun generation in which entities are mapped to pronouns. In the English language classroom, teachers can provide general guidelines to learners to select appropriate interrogative pronouns, such as to use *where* for places, *who* for people and *what* for things. Students are generally able to differentiate between places, people and things. However, for software, words are just a series of letters, referred to as a string. In order for the software to classify strings as places, people or things, a combination of extensive lists and machine learning is needed. To provide an example of the complexity of the task, a three of the many rules are given below.

(10) When the head word of an adverbial phrase is on the adverbs of time list, use *when.*
(11) When the head word of a noun phrase is an entity of the type geo-political, use *where.*
(12) When the distance from the head word vector to the vector human in the semantic space is less than the distance to the vector object, use *who.*

The rule shown in (10) is straightforward and may be thought of in practical terms as when expressions like *yesterday* and *last month* are used, questions using *when* are created. The rule in (11) means that the pronoun *where* is assigned when names of regions, such as Ohio, the United States and the Himalayas are detected. The rule shown in (12) is rather complex, but involves predicting the likelihood that the head word is human using machine learning. The accuracy rate of the interrogative pronoun selection is highest when based on the presence of a single word or phrase, lower when based on identification of entity type and the lowest when based on the decision of a machine learning classifier.

When creating open or closed questions, there is no need to replace entities with pronouns, but in order to create sentence tags, the pronoun for the grammatical subject needs to be determined. This raises a number of issues that need to be overcome. The number of a noun needs to be established so plural nouns may be assigned the pronoun *they*. This was achieved in part by handcrafting subject-pronoun rules based the noun phrases (NPs) and subject-verb agreement.

In cases where the subject is a proper noun, named entity recognition is used to identify names used for people. Machine learning is harnessed to classify given names by gender based on a training set of 7000 Western names and 500 Japanese names. Other proper nouns that are not categorized as names of people and then compared for gender based on an extensive list of genderized nouns. Common nouns are classified using a complex algorithm harnessing word embedding, semantic representation and machine learning (see Vu and Blake 2021 for more details).

6 Classroom Observation of Pilot Testing

Pilot testing is used to verify whether a software program or system works under typical operating conditions. Generally, a group of real-world end users are asked to try out a program or system (Rossett and Schafer 2012). In our case, we also observed participants using the web app. The Question Generator was trialed during the development period

with small focus groups of students known to the developers. The purpose of these focus groups was to enhance the usability of the tool by observing how learners interact with the tool with the view to improving the user interface and user experience. However, to get a fuller picture of whether the web app meets user needs, two different age groups of users were observed using the Question Generator. The first was a cohort of junior high school children (n = 30; male = 14; mean age = 15, CEFR[3] = A1) and the second was a class of university sophomores (n = 48; male = 44; mean age = 20, CEFR = A2 and B1), majoring in computer science and engineering. Each focus group of between 6 and 10 spent one teaching period (50 min) to try out the Question Generator, and share their feedback. There were two prepared questions for each focus group:

1. What did you like or dislike about the Question Generator?
2. How can we improve it?

To draw out more information for the participants, probing and clarification questions were asked to gain a better understanding of what we should improve in the next release.

During the classroom observation, the learners accessed the Question Generator via either a desktop computer installed in the language laboratory in which the study was being conducted or via their own wifi-enabled device. The junior high school children shared workstations and worked in pairs. The university students were given the option to work alone or in pairs. Around half of the university students opted to work in pairs while the remaining students worked individually.

Both the junior high school students and university students appeared to be engaged and on task during the study period. The class teacher introduced the Question Generator in a similar manner in both studies by showing how to access the web app. Students worked with minimum supervision in both studies.

Despite the differences in the number of years having studied English, both groups had difficulty formulating syntactically correct questions in English. This web app thus meets a need in terms of appropriacy of content as both school-age students and university learners struggled at times to create grammatically-accurate questions for the input sentences. Both groups had a reasonable mastery of commonly taught and highly frequent question forms, such as: *"What is your name? What time is it? Can you help me?",* but struggled with longer question forms.

The usage of the web app differed slightly between the two cohorts of students. The schoolchildren entered shorter sentences on more concrete topics, such as "I went to school yesterday" while university learners entered more complex and more abstract sentences, many of which were copied from online sources. The junior high school students tended to input sentences that they created related to their interests but also input sentences that were copied from their English textbooks.

The tutor noted that in both groups learners were particularly active when working in pairs on one monitor. This may be due to increased opportunity for interaction and the students desire to or willingness to communicate with their classmate. However, this dynamic is unlikely to be directly related to the web app itself.

[3] Common European Framework of Reference for Languages.

The individualization appeared to motivate learners as they were able to input sentences that were of interest to them. Participants in both classroom observations smiled and joked while entering statements for the Question Generator. Although enjoyment may not correlate directly to learning, students are likely to spend more time on tasks that are enjoyable (Fielding 2020). The ability to choose topics, vocabulary and grammatical structures gave learners a wide range of autonomy and added the possibility for them to energize their learning. The junior high school students, in particular, appeared to draw on in-jokes when entering statements. Although both groups in this study were courteous; with less well-behaved groups, it could be necessary for teachers to check that learners are not inputting statements portraying their peers in a negative light.

Based on feedback from the learners, a colorization feature is planned. Different colours may be used for words based on whether or not the words are present or absent in the original statement, or their form has been altered. When implemented this feature can help learners understand which elements in the question are borrowed, altered or not present in the original declarative statement.

7 Conclusion

Using NLG enables users to create language learning materials on demand based on user input. This obliviates the need to prepare inordinate amounts of materials to meet the needs of learners. The generation of single statements for pedagogic purposes is already possible as the Question Generator has shown. Although many genres of writing, including scientific research articles, have been automatically generated using NLG, none of these controlled the language based on sophistication and suitability for language learning. Other natural language generation tools currently under development, such as the trend description generator, involve the generation of a paragraph that serves as a prototypical example of this text type. This type of language generation has to be based on insights gained from specialist corpora to ensure the appropriacy of the rhetorical organisation and lexical realization.

Natural language generation simplifies the process of individualizing materials for language learners as the decisions and parameters are built into the software prior to the materials being created. By utilizing existing libraries and lexical lists, texts can be individualized to learners. In the case of the Question Generator, learners determine the content of the question by inputting a prompt sentence whereas in the trend description generator, learners input or select data values (e.g. time periods and numerical values) that can be used to create a graph or bar chart, and concomitantly an accompanying textual description of the visual. Future tools could ask users to input details regarding various parameters, such as language level, learning styles and preferences. These details may be used to inform the mode, manner and medium of texts generated.

The Question Generator is the first online pedagogic tool that enables learners to generate questions based on user input, and thus breaks new ground in the growing set of intelligent CALL tools. However, the selection of an appropriate hosting platform needs to be ironed out before this is released for general use. There are currently three versions of this software, each with slightly different levels of accuracy. The version with the lowest accuracy is deployed on a free plan on Heroku, which is a platform

that allows developers to run applications in the cloud. Although this is free, access is limited and once the threshold is reached, the web app is not accessible. In this version few machine learning functions are integrated to reduce the processing time. The second most accurate version is deployed on a pay-for-use platform, Amazon Web Services (AWS). The running costs, however, rise substantially as the number of users increases. More machine learning functions are included, which results in higher accuracy than for the free version on Heroku. The most accurate version runs offline, and is able to use machine learning with a time delay of only a few seconds. Thus, our current dilemma is the trade-off between running cost and performance. Two ways under consideration to support the running costs of the web app are using an advertising model and/or a subscription model for regular users, which could allow guest visitors to use the web app free of charge.

Although natural language processing pipelines can generate language (questions in our case) on demand, sharing such pipelines with a wider audience is a challenge. Intelligent CALL is set to radically impact online language learning, but the true power of iCALL will probably only be available once a suitable hosting platform model has been created. Video sharing platforms, such as YouTube, were a game changer for video creators. Let's hope that an entrepreneur sets up a similar such platform for iCALL tool developers to share their creations freely.

References

Aitken, R.: Teaching Tenses. Intrinsic Books Ltd. (2021)

Alexander, C.: Language labs: an overview of the trends. Teach. Engl. Technol. ***7***(3) (2007)

Anthony, L.: AntConc (Version 4.0.11) [Computer Software]. Tokyo, Japan: Waseda University (2022). https://www.laurenceanthony.net/software/antconc/

Appadurai, A.: Modernity at Large: Cultural Dimensions of Globalization. University of Minnesota Press, Minneapolis (1996)

Bailin, A., et al.: A bibliography of intelligent computer-assisted language instruction. Comput. Humanit. **23**, 85–90 (1989)

Benson, P.: Teaching and Researching Autonomy in Language Learning. Longman, London (2001)

Blake, J.: Intelligent CALL: using pattern matching to learn English. In: Kruk, M., Peterson, M. (eds.) New Technological Applications for Foreign and Second Language Learning and Teaching, pp.1–23. IGI Global, Hershey (2020). https://doi.org/10.4018/978-1-7998-2591-3

Bogach, N., et al.: Speech processing for language learning: a practical approach to computer-assisted pronunciation teaching. Electronics **10**(3), 235 (2021). https://doi.org/10.3390/electronics10030235

Boitsova, E., et al.: StudyIntonation courseware kit for EFL prosody teaching. In: Proceedings of 9th International Conference on Speech Prosody 2018, pp. 413–417 (2018)

Boulton, A.: Data-driven learning and language pedagogy. In: Thorne, S.L., May, S. (eds.) Language, Education and Technology. ELE, pp. 181–192. Springer, Cham (2017). https://doi.org/10.1007/978-3-319-02237-6_15

Braun, S.: Integrating corpus work into secondary education: from data-driven learning to needs-driven corpora. ReCALL **19**(3), 307–328 (2007). https://doi.org/10.1017/S0958344007000535

Chang, S.C.: A contrastive study of grammar translation method and communicative approach in teaching English grammar. Engl. Lang. Teach. **4**(2), 13 (2011)

Chinnery, G.M.: Going to the MALL: mobile assisted language learning. Lang. Learn.Technol. **10**(1), 9–16 (2006). 10125/44040

Chujo, K., Anthony, L., Oghigian, K., Uchibori, A.: Paper-based, computer-based, and combined data-driven learning using a web-based concordancer. Lang. Educ. Asia **3**(2), 132–145 (2012). https://doi.org/10.5746/LEiA/12/V3/I2/A02/Chujo_Anthony_Oghigian_Uchibori

Chukharev-Hudilainen, E.: Empowering automated writing evaluation with keystroke logging. In: Lindgren, E., Sullivan, K. (eds.) Observing Writing: Insights from Keystroke Logging and Handwriting, pp. 125–142. Brill (2019). https://doi.org/10.1163/9789004392526_007

Chukharev-Hudilainen, E., Saricaoglu, A.: Causal discourse analyzer: Improving automated feedback on academic ESL writing. Comput. Assist. Lang. Learn. **29**(3), 494–516 (2016)

Dizon, G.: Using intelligent personal assistants for second language learning: a case study of Alexa. TESOL J. **8**(4), 811–830 (2017). https://doi.org/10.1002/tesj.353

Dizon, G., Tang, D.: Intelligent personal assistants for autonomous second language learning: an investigation of Alexa. JALTCALL J. **16**(2), 107–120 (2020)

Ellis, R., Mifka-Profozic, N.: Recasts, uptake, and noticing. In: Bergsleithner, J.M., Frota, S.N., Yoshioka, J.M. (eds.) Noticing and Second Language Acquisition: Studies in Honor of Richard Schmidt, pp. 61–80. National Foreign Language Resource Center University of Hawaii, Honolulu (2013)

Fielding, R.: Language teaching in monolingual policy settings: teacher views of successful language learning and effective language programmes. Lang. Learn. J. **50**, 344–359 (2020). https://doi.org/10.1080/09571736.2020.1762711

Gamper, J., Knapp, J.: A review of intelligent CALL systems. Comput. Assist. Lang. Learn. **15**(4), 329–342 (2002). https://doi.org/10.1076/call.15.4.329.8270

Griffiths, C., Soruç, A.: Chapter 3 Learning as an adult. In: Burns, A., Richards, J.: (eds.) The Cambridge Guide to Learning English as a Second Language, pp. 27–34. Cambridge University Press, Cambridge (2018). https://doi.org/10.1017/97810090247 61.005

Hadley, G.: An introduction to data-driven learning. RELC J. **33**(2), 99–124 (2002)

Hamada, Y.: Shadowing: what is it? How to use it. Where will it go?. RELC J. **50**(3), 386–393 (2019). https://doi.org/10.1177/0033688218771380

Harwood, N. (ed.): English Language Teaching Textbooks Content, Consumption, Production. Palgrave Macmillan, London (2014). https://doi.org/10.1057/9781137276285

Hondo, J.: Teaching English grammar in context: the timing of form-focused intervention. In: Christison, M., Christian, D., Duff, P.A., Spada, N. (eds.) Teaching and Learning English Grammar, pp. 58–73. Routledge, Abingdon-on-Thames (2015)

Isaacs, T.: Integrating form and meaning in L2 pronunciation instruction. TESL Canada J. **27**(1), 1–12 (2009). https://doi.org/10.18806/tesl.v27i1.1034

Kilgarriff, A., et al.: The sketch engine: ten years on. Lexicography **1**(1), 7–36 (2014). https://doi.org/10.1007/s40607-014-0009-9

Krashen, S.: Principles and Practice in Second Language Acquisition. Pergamon Press, New York (1982)

Krashen, S.: We acquire vocabulary and spelling by reading: additional evidence for the input hypothesis. Mod. Lang. J. **73**(4), 440–464 (1989). https://doi.org/10.1111/j.1540-4781.1989.tb05325.x

Little, D.: Responding authentically to authentic texts: a problem for self-access language learning? In: Voller, P., Benson, P. (eds.) Autonomy and Independence in Language Learning, pp. 225–236. Routledge, Abingdon-on-Thames (2014)

Long, M.H., Porter, P.A.: Group work, interlanguage talk, and second language acquisition. TESOL Q. **19**(2), 207–228 (1985). https://doi.org/10.2307/3586827

Mullins, C.: Responsive, mobile app, mobile first: untangling the UX design web in practical experience. In: Proceedings of the 33rd Annual International Conference on the Design of Communication, vol. 22, pp. 1–6 (July 2015). https://dx.doi.org/https://doi.org/10.1145/2775441.2775478

Nakahama, Y., Tyler, A., Van Lier, L.: Negotiation of meaning in conversational and information gap activities: a comparative discourse analysis. TESOL Q. **35**(3), 377–405 (2001). https://doi.org/10.2307/3588028

Goertzel, B., Pennachin, C., Geisweiller, N.: Natural Language Generation. In: Engineering General Intelligence, Part 2. ATM, vol. 6, pp. 487–495. Atlantis Press, Paris (2014). https://doi.org/10.2991/978-94-6239-030-0_28

Piaget, J.: The Language and Thought of a Child. Harcourt Brace and Co., New York (1956)

Purgina, M., Mozgovoy, M., Blake, J.: WordBricks: mobile technology and visual grammar formalism for gamification of natural language grammar acquisition. J. Educ. Comput. Res. **58**(1), 126–159 (2020). https://doi.org/10.1177/0735633119833010

Qian, M., Chukharev-Hudilainen, E., Levis, J.: A system for adaptive high-variability segmental perceptual training: implementation, effectiveness, transfer. Lang. Learn. Technol. **22**(1), 69–96 (2018). 10125/44582

Rossett, A., Schafer, L.: Job Aids and Performance Support: Moving from Knowledge in the Classroom to Knowledge Everywhere. John Wiley & Sons (2012)

Rudin, C.: Stop explaining black box machine learning models for high stakes decisions and use interpretable models instead. Nat. Mach. Intell. **1**(5), 206–215 (2019). https://doi.org/10.1038/s42256-019-0048-x

Schmidt, R.: The role of consciousness in second language learning. Appl. Linguis. **11**(2), 129–158 (1990). https://doi.org/10.1093/applin/11.2.129

Shaffer, C.: A comparison of inductive and deductive approaches to teaching foreign languages. Mod. Lang. J. **73**(4), 395–403 (1989). https://doi.org/10.2307/326874

Spratt, M., Pulverness, A., Williams, M.: The TKT Course. Cambridge University Press, Cambridge (2005)

Truscott, J.: Noticing in second language acquisition: a critical review. Second. Lang. Res. **14**(2), 103–135 (1998). https://doi.org/10.1191/026765898674803209

Volodina, E., Borin, L., Lofsson, H., Arnbjörnsdóttir, B., Leifsson, G. Ö.: Waste not; want not: towards a system architecture for ICALL based on NLP component re-use. In: Proceedings of the SLTC 2012 Workshop on NLP for CALL. Linköping University Electronic Press (2012)

Vu, D.T., Blake, J.: Design and development of a question generator for learners of English. In: Roy, D., Fragulis, G., Cantu Campos, H.A. (eds.) Proceedings of the 3rd ETLTC International Conference on Educational Technology, Language and Technical Communication. SHS Web Conference, vol. 102, p. 01011 (2021). https://doi.org/10.1051/shsconf/202110201011

Ward, M., Mozgovoy, M., Purgina, M.: Can WordBricks make learning Irish more engaging for students? Int. J. Game-Based Learn. **9**(2), 20–39 (2019). https://doi.org/10.4018/IJGBL.2019040102

Wik, P., Hjalmarsson, A.: Embodied conversational agents in computer assisted language learning. Speech Commun. **51**(10), 1024–1037 (2009). https://doi.org/10.1016/j.specom.2009.05.006