Language and Cognition (2026), 18, e9, 1–23
doi:10.1017/langcog.2025.10056

ARTICLE

Habitual code-switching between English and Cantonese: bilingual language control reflected by cognitive load in sight translation

Ut Meng Lei[1,2,3], Ka Lon Chan[3] and Victoria Lai Cheng Lei[1,2,3]

[1]Faculty of Arts and Humanities, University of Macau, Macau SAR, China; [2]Centre for Cognitive and Brain Sciences, University of Macau, Macau SAR, China and [3]Centre for Studies of Translation, Interpreting and Cognition, University of Macau, Macau SAR, China
Corresponding author: Ut Meng Lei; Email: riko.research@gmail.com

(Received 05 February 2024; Revised 12 August 2025; Accepted 03 December 2025)

Abstract
Previous research on bilingual language processing has shed light on language control mechanism behind comprehension and speech production of bilinguals but the commonness or habitualness of code-switched terms in the design of the stimuli is seldom explored. This research attempted to explore the relationship between habitualness of code-switched terms and cognitive load level in Cantonese-English sight translation tasks among native Cantonese speakers in Macao through investigation from both objective and subjective data. The research collected and analysed eye-tracking data, audio data and NASA-TLX data. The results provided partial evidence that Macao native Cantonese speakers tended to experience lower cognitive load in the sight translation task when they were allowed to code switch the words they habitually applied in English; however, the correlation between code-switching and reduction of cognitive load was not significant. The findings suggest the selectivity of code-switching in language output and indicate that different levels of habitualness of the code-switched terms may modulate cognitive load.

Keywords: bilingual language control; cognitive load; habitual code-switching; sight translation

1. Introduction

1.1. Linguistic background of Macao

The official languages of Macao are Chinese and Portuguese, which are widely used in administrative and legal documents. However, the status of different languages of Macao does not align with its official language. Bray and Koo (2004, as cited in Wheeler, 2019, p. 74) describe the language situation of Macao as '[having the] complexities of language policy [that] have all the ingredients of Hong Kong, that is, Cantonese, Putonghua/Mandarin and English, plus the additional ingredient of Portuguese'. As quoted from the

© The Author(s), 2026. Published by Cambridge University Press. This is an Open Access article, distributed under the terms of the Creative Commons Attribution licence (http://creativecommons.org/licenses/by/4.0), which permits unrestricted re-use, distribution and reproduction, provided the original article is properly cited.

https://doi.org/10.1017/langcog.2025.10056 Published online by Cambridge University Press

Census of 2021 (DSEC, 2022), 81% of the population aged 3 and above frequently used Cantonese as their 'usual language', people who used Putonghua (Mandarin) and other Chinese dialects (e.g., Hokkien dialect) accounted for 4.7% and 5.4%, respectively. English was spoken as a usual language by 3.6% of the population. Portuguese, despite its prestigious status as one of the official languages in the territory, was only spoken as a 'usual language' by about 0.6% of the population (DSEC, 2022). In terms of fluency, the population in Macao who were fluent in Cantonese, Putonghua, English and Portuguese accounted for 86.2%, 45%, 22.7%, and 2.3%, respectively. Although English is not the official language in Macao, the historical power of the English language (Ng & Deng, 2017) has long left its trace in Macao, recorded as the medium of contact between English traders and locals in Macao as early as 1637 (Bolton, 2003, p. 126, as cited in Botha, 2013). The infusion of English is commonly noticed in various industries in Macao. For example, adequate English proficiency is seen as an advantage or even a required competence in the workplace regarding the financial, high-technological, international commercial and educational sectors (Cheng & Liu, 2010). Some scholars describe English as a 'de facto working language' (Botha, 2013) or even "de facto official language" (Moody, 2008) in Macao, acclaiming its prestigious status in the territory. In government administration, it was tallied that there were 70% of the government institutions in Macao SAR offered English content on their Internet service in addition to Chinese (Moody, 2008). It shows that prominent identity of English is recognized by the Macao Government. In short, the status of English in different sectors of Macao elucidates the positive outlook of English use.

1.2. Code-switching between Cantonese and English in Macao

Due to the popularity of English in the region, the phenomenon of Cantonese-English code-switching is thus common in Macao. For example, terms such as 'locker', 'final', 'major', 'project', 'professor', 'present' are examples of code-switched terms (hereafter 'CS terms') used in spoken Cantonese among Macao students (Chan, 2015). Despite the acknowledgement of the commonness of code-switching in the daily conversations among Macao residents, the language switching mechanism between Cantonese and English in this context, however, is largely under-explored. This study adopted a sight translation task to investigate the relationship between habitual code-switching and cognitive load level among local Macao residents. Sight translation is a type of simultaneous interpreting that involves oral translation of written source text into target text (Gile, 2009). This mode not only allows the collection of eye-tracking data (i.e., that there will be written source text) and audio data (i.e., from oral rendition of source text) but also enables the observation of how cognitive load level can be affected by the level of habitualness in code-switching particular words (i.e., participants could code-switch certain words habitually in daily language use so that they have difficulty suppressing the English words). This study thus contributes to the illustration of the language control mechanism among the bilinguals that habitually code switch in their daily life.

1.3. Cognitive load and language switching

Cognitive load can be defined as the amount of mental effort or recourses required to perform a cognitive task (e.g., Bóna & Bakti, 2020; Chen, 2017). In the context of

https://doi.org/10.1017/langcog.2025.10056 Published online by Cambridge University Press

language processing, the relationship between cognitive load and switching languages has been widely explored, and experiments with different required tasks for the participants were conducted (Blanco-Elorrieta & Pylkkänen, 2017; Gross & Kaushanskaya, 2015). In a numeral naming task, Meuter and Allport (1999) found that when comparing between trials with language switching with those without language switching, bilinguals responded faster in the latter trial. Costa and Santesteban (2004) revealed similar results with a picture-naming task. While studies such as the above mainly discussed the cued switching conditions in which participants were instructed to use an assigned language in particular trials, experiments with voluntary switching revealed different results. For example, researchers such as de Bruin et al. (2018) and Jevtović et al. (2020) corroborated that voluntarily mixing languages in production is less cognitively demanding than staying in one language. In a magnetoencephalography study, Blanco-Elorrieta and Pylkkänen (2017) found that natural switching in comprehension and production tasks require narrower spread of brain activation than artificial settings. However, switching cost was observed in a study even when participants switched voluntarily (Gollan & Ferreira, 2009). Despite the existing accounts about cognitive load of switching languages, the habitualness of the code-switched terms has largely been underexplored. In the context of Hong Kong, Hui et al. (2022) took into consideration the habitualness of English code-switched terms in a silent reading task of Cantonese monolingual sentences and Cantonese-dominant sentences with code-switched English words. As a Hong Kong local, Hui composed the sentences, and had them rated by students on the language habitualness (i.e., the tendency towards English or Cantonese) of the target words (i.e., the code-switched English terms and their Cantonese equivalents) in the given sentences. Based on eye movement data (e.g., fixation count), the authors found that cognitive load was not present when participants processed the habitually switched English words. In comparison with word-based stimuli such as those in picture-naming task and numeral naming task, Hui et al. (2022) integrated the code-switched terms in the context of sentences, which made the stimuli sensible and natural to local bilinguals, providing a more valid way to investigate the mechanism and the factors behind different levels of cognitive load in processing different languages. Similar to the findings of Hui et al.'s (2022) study, the study of Adamou and Shen (2019) also highlighted the free cost of processing auditory stimuli with mixed languages when the switched words were frequently used in the community. With picture-naming tasks, Kleinman and Gollan's (2016) study found that switching languages in production was not demanding if the accessibility of the switched word is high. The above studies take habitualness of the CS terms into account and provide insights into the perception of code-switched stimuli. However, the code-switching mechanism in relation to modalities involving production awaits elaboration. The process of sight translation, which involves language control in production, should provide more information in terms of how the habitualness of the code-switching terms associate with the cognitive demands for suppressing the non-target language equivalents in oral delivery, complementing the abovementioned studies focusing on code-switching mechanism of perception.

Scholars have schematically outlined the language control mechanism behind language switching (e.g., Green, 2018; Green & Abutalebi, 2013; Green & Li, 2014). The different control states involved in the models illustrate how bilinguals select the language(s) they use in communication. Green and Abutalebi (2013) proposed Adaptive Control Hypothesis (ACH), suggesting three kinds of context (i.e., single-language

https://doi.org/10.1017/langcog.2025.10056 Published online by Cambridge University Press

context, dual-language context, and dense code-switching context) governing the language-switching behaviours of bilinguals: single-language context (i.e., a context where speakers tend to use only one language) and dual-language context (i.e., a context where speakers switch between two languages according to communicative purposes and the languages used by the interlocutors in the context) are hypothesized to be more cognitively demanding than dense code-switching context (i.e., a context where interlocutors share the same language habits and code-switch whenever they feel comfortable). According to the Control Process Model (CPM) (Green, 2018; Green & Li, 2014), staying in one language in the conversation is associated with competitive language control, while switching between languages in an utterance by insertion or alternation is linked to cooperative control. Following Muysken's (2000) definition, insertion features the switch of words or lexical constituents between languages within an utterance, while alternation features a switch in the middle of the sentence, making the sentence start in one language in the first half, and end in another for the second half. The chosen control mode governs the entry of language items into the planning stage of utterance. Referring to the authors' analogy, competitive mode is linked to the situation that the door for the target language is open while the door for the non-target language is locked (i.e., inhibited). In the cooperative mode, on the other hand, the door of the non-target language can be momentarily pushed open (i.e., 'coupled control', Green & Li, 2014, p. 502) or both target-language door and non-target-language door are open for free entering of both languages (i.e., 'open control', Green & Li, 2014, p. 502). As is suggested by Nour et al. (2020), simultaneous interpretation is a context characterized by simultaneous processing of two languages, and the cognitive demands are comparable to or even higher than dual-language contexts common in bilingual language use. As is stated by the authors, while both simultaneous interpretation and dual-language contexts involve processing of two languages, dual-language contexts feature sequential alternation between two languages, whereas simultaneous interpretation involves overlapping language processing (See also Paradis, 1994). The overlaps of languages in simultaneous interpretation reflect the continuous suppression of one language while orally producing another. On the other hand, the allowance of code-switching in production simulates the dense code-switching context and cooperative control state in which individuals can select a proper lexis at the proper position or moment. With the framework of ACH and CPM, this study can explain how specific sight-translation conditions (i.e., allowing habitual code-switching and not allowing habitual code-switching) can trigger different levels of cognitive load.

1.4. Code-switching in interpreting and sight-translation

Language control is considered as a basic requirement in simultaneous interpretation. While interpreters are by default bilinguals, they are supposed to be equipped with higher mastery of language control for prompt and precise delivery of target text (Bialystok et al., 2012; Perani & Abutalebi, 2015). Researchers have a particular interest in the language processing effort of interpreters in language tasks. In the event-related potential study conducted by Proverbio et al. (2004), it was found that interpreters took more time to process input text in the mixed language condition than in the unmixed condition. The N1 and N400 results also signalled larger negativity in mixed condition correspondingly. Sentences featuring a switch from L1 to L2 were proved to be less cognitively demanding than a switch from L2 to L1.

https://doi.org/10.1017/langcog.2025.10056 Published online by Cambridge University Press

With an auditory lexical selection task, Boos et al. (2022) also proved that the non-switch trials were less demanding than the switch trials for the interpreters. They compared subjects with different simultaneous interpretation expertise and found no significant correlation between expertise and cognitive load. In an fNIRS study, Lin et al. (2018) have directly compared code-switching as a translation strategy (referred to as 'nontranslation' in their study) with other two strategies that required constant suppression of the source language. Their result showed that 'nontranslation' led to lower brain activation in brain regions such as left prefrontal cortex and Broca's area, suggesting that 'nontranslation' can serve as an effort-reduction tactic in translation.

Previous translation studies (e.g., Boos et al., 2022; Proverbio et al., 2004), however, have not delineated in detail how cognitive load is modulated by code-switching in oral output, and they lacked consideration of the effect of the habitualness of the code-switched terms in the design of the stimuli. The study of Cheung (2001) investigated the influence of English words and phrases in source text on the production of Cantonese target text in simultaneous interpretation. Situated in the context of Hong Kong, student interpreters who were Cantonese speakers were invited to join the simultaneous interpretation experiment. The researcher prepared an English video-taped speech, and the participants were required to interpret the English source text into Cantonese, in which half of the participants were allowed to code switch. The results indicated that participants who were allowed to code switch performed better than their counterparts. Specifically, they demonstrated a higher frequency of English used in the target text as well as fewer hesitations, pauses, unfinished sentences and meaning errors. The author concluded that code-switching could be used as a strategy to help manage the effort, which guaranteed better completeness, accuracy and fluency in simultaneous interpreting. Although effort has been made to investigate the relationship between code-switching and the cognitive load in interpretation, the measures of cognitive load used in the study only included holistic rating and general cognitive load indicators from verbal data, which were limited in validity and reliability. While the data type of the above study is potentially limited by the mode of simultaneous interpretation, which only captures oral performance, the researchers of this study consider sight translation a valuable mode in the investigation of cognitive load with the allowance to capture eye movement data. Given the association between eye movements and language processing in translation tasks (Seeber, 2013), the related cognitive load can be more adequately analysed together with oral performance if eye-tracking technique is applied. Some eye-tracking matrices such as fixation counts and fixation durations are common cognitive load indicators in interpreting studies. In addition, the applied material in Cheung's (2001) study was derived from a university president's welcoming address spoken by a male Caucasian speaking standard North American accent. The formality makes it hard to judge whether the English words in the video are good representatives of the common English words used in the daily life of Cantonese native speakers in Hong Kong. Although it was claimed that only the English words and phrases commonly used by university students in Hong Kong were used in the speech, Cheung (2001) admitted that it was rather difficult to distinguish which English words are commonly used by Cantonese speakers in their Cantonese conversation, and thus it was subject to the raters' judgement as to decide which English words in the target text are counted as the commonly code-switched terms. Without revealing the details of the commonly code-switched English words in the source text, the interpretation of the results remained dubious. In view of the

https://doi.org/10.1017/langcog.2025.10056 Published online by Cambridge University Press

above matters, well-controlled stimuli and more purposeful experiment design (e.g., specifying the habitual CS terms in the design) should add value to the reliability of research and serve to reveal the rule of habitual code-switching among dense code-switchers.

As is stated, the habitualness of code-switched terms in oral output received limited attention in the research field. This study aims to incorporate the common English code-switched words of Macao's native Cantonese speakers into the English source text, and subsequently investigates how the two different requirements (i.e., allowing code-switching in sight-translation production and not allowing code-switching in sight-translation production) may lead to different levels of cognitive load. The findings should enable a more comprehensive delineation of bilingual language switching mechanism incorporating habitual code-switching. A research question is thus raised to guide the study: To what extent does allowing habitual code-switching in sight translation modulate the level of cognitive load?

2. Methods

2.1. Data collection

Eye tracking data (e.g., total fixation duration, fixation counts) were collected and analysed. According to Hvelplund (2014), long and frequent fixations indicate a high level of processing effort. Likewise, the longer and more frequently the eyes visit a particular area, the higher cognitive load is reflected (Hui et al., 2022). Based on the rationale, mean fixation duration (i.e., average duration of all fixations) on screen was chosen as a parameter for comparisons of cognitive load among baseline task and the two modes of sight translation tasks. All parameters were first averaged per trial for each participant, then averaged across participants for each mode before further analysis.

Besides, cognitive load parameters applied for audio data (i.e., participants' interpretation oral output data) such as total task time, total English syllables, frequency and proportion of fillers, revisions, repetitions, silent pause duration and count and prolonged sounds were collected as objective indicators of cognitive load (e.g., Berthold & Jameson, 1999; Bóna & Bakti, 2020; Inhoff et al., 2011). In addition, this study included total task time (Chmiel & Mazur, 2013) and the mean pause duration related to the habitual English CS terms. In detail, pause duration produced during the sight translation process was identified and calculated as an indicator of cognitive load. Pauses with a duration higher than 300 ms were identified as silent pause. These indicators can be readily identified with the assistance of the open-source software Praat, which is widely used for speech analysis. In particular, we inserted the collected sound waves and make necessary annotations (e.g., the uttered words, the types of disfluencies) with 'Textgrid' right below the window of sound waves, which allowed accurate identification of the timing and duration of each parameter appeared in the audios. After all annotations were made, we output the data as a chart, indexed the parameters, and adjusted it to a format available for calculation. Similar to eye tracking data, for each participant, all parameters were averaged within each trial, then these trial averages were further averaged by mode prior to analysis.

Finally, NASA-Task Load Index (TLX), which allowed participants to self-report their cognitive load, was applied to collect subjective measures of cognitive load for each mode.

https://doi.org/10.1017/langcog.2025.10056 Published online by Cambridge University Press

In addition, a semi-structured interview was conducted to investigate the potential reasons behind the difficulty during the sight translation tasks, and the rationale behind their decision to code switch during the experiment. This step allowed the researchers to see if the participants were aware that they code-switched because of their habit of code-switching particular words in their daily life.

2.2. Materials

For designing the stimuli of the experiments, English code-switched terms that are common in the context of Macao were collected from three different local radio programmes from Teledifusão de Macao, the only Cantonese Chinese radio station in Macao. The three radio programmes are all about contemporary issues of the local Macao society, and are hosted by two to three hosts in a casual conversation manner. In some episodes, the hosts invited guests to the programmes for sharing, and there were also occasions when the hosts recorded their programme or had a live broadcast outside the radio station. Altogether, eight episodes of the three programmes broadcast between March and April 2023, totalling approximately 12 hours, were recorded. The audio recording was processed with machine transcription and manual checking and editing. Consequently, within the 12 hours of recorded radio broadcast, 710 instances of English CS terms were noted. Some of these terms were repeated, resulting in a final list of 339 different English CS terms. Here, the CS terms refer to the use of English lexicons inserted into the middle of the original Cantonese interlocution. These CS terms formed the source of the stimuli.

2.3. Stimuli

The stimuli used in the sight translation experiment consisted of four English paragraphs. The four paragraphs contained phrases and content chosen from the Cantonese radio programmes mentioned above, which were translated and rewritten into English. They were selected either because they contained more English code-switched terms or because the context was suitable for inserting code-switched words. Adaptations from the original meaning were made to ensure coherence. The four paragraphs were of the same length containing 147 words written in 11 sentences. It guaranteed the same average sentence length (i.e., 13.36 words per sentence). Each paragraph contained 17 English CS terms. Finally, readability was normalized by measuring indices, including Flesch Kincaid Grade, Gunning Fog Index, SMOG Index, Coleman-Liau Index, Lexical Density and Type Token Ratio. The readability consensus of all four paragraphs was measured to be suitable for learners of the 6th to 7th Grade, which suggests an average level of difficulty. It ensured that all stimuli were of similar difficulty to avoid biases in the experimental outcome. Table 1 has summarized the parameters that were applied for normalization in the design of the stimuli.

2.4. Participants

There were 30 volunteer participants in this study. Twenty-four of them were females while six of them were males. They were all local Macao residents and spoke Cantonese as their native language, with English either their second language

https://doi.org/10.1017/langcog.2025.10056 Published online by Cambridge University Press

Table 1. Information of the stimuli

	Stimuli 1	Stimuli 2	Stimuli 3	Stimuli 4
Word count	147	147	147	147
Sentence count	11	11	11	11
Average sentence length	13.36	13.36	13.36	13.36
Readability consensus	6th–7th Grade	6th–7th Grade	6th–7th Grade	6th–7th Grade
Lexical density (LD)	53.06	48.98%	49.66%	50.34%
Type/token ratio	0.68	0.61	0.62	0.57
Flesch kincaid grade	4.47	4.55	4.15	4.23
Gunning Fog Index	6.71	5.89	6.43	6.43
SMOG Index	8	8	6.57	6.98
Coleman-Liau Index	5.93	6.76	5.07	6.61
Number of CS terms	17	17	17	17

(56.67%) or third language (43.33%). They were aged between 22 and 30 (M = 25.8, SD = 2.894), and they were either taking their graduate studies or graduated within the last 3 years. All their current or previous majors are English or English-related fields such as Education in English and English studies. According to the results of the demographic questionnaire adapted from Language and Social Background Questionnaire (LSBQ) (Anderson et al., 2018), which will be introduced in the next section, all of the participants considered their Cantonese proficiency to be 5 (native), while their perceived English proficiency ranged between 3 (fine) (36.67%) and 4 (proficient) (63.33%). The participants varied with the age they started learning English (M = 4.43, SD = 2.897). Most participants (94.44%) started learning English from their school. In detail, out of a scale from 0 to 10, from totally not proficient to highly proficient, they reported their proficiency level of English speaking (M = 7.8, SD = 0.847), comprehension (M = 7.93, SD = 0.868), reading (M = 7.73, SD = 0.828) and writing (M = 7.38, SD = 1.129) to be rather proficient. Furthermore, out of a scale from 1 to 5 (with 1 suggesting 'never' and 5 suggesting 'all the time'), they reported that they sometimes used English for speaking (M = 3.4, SD = 0.498), listening (M = 3.167, SD = 0.461), reading (M = 3.467, SD = 0.507) and writing (M = 3.233, SD = 0.568). With regards to the language they used for daily communication, they tended to use Cantonese or Mandarin when communicating with other people in their life, especially with their family members and neighbours. In their daily life, they used English more for writing emails, at school, while using social media and preparing shopping lists and notes. Finally, out of a scale from 1 to 6, with 1 indicating 'not knowing other languages', and 2 to 6 ranging from 'never' to 'all the time', it is worth mentioning that they reported often code-switching when they were using social media (M = 4.7, SD = 0.915) and when they were with their friends (M = 4.33, SD = 884). All participants in this study were informed of the nature and possible consequences of the study and written informed consent was obtained before the experiments according to the experimental protocols approved by the Research Ethics Committee at the University of Macao.

2.5. Instruments

EyeLink 1000 Plus eye tracker with 1000 Hz sampling rate was utilized in the experiment. A head support was set up to make sure the participants can lean their forehead on the rest area to ensure head positioning stability. The LCD screen with

https://doi.org/10.1017/langcog.2025.10056 Published online by Cambridge University Press

1024 × 768 pixel resolution served to demonstrate the stimuli. The viewing distance between the participants and the screen was set at 67 cm (1.75 times the width of the monitor, i.e., 33.5 cm). The stimuli were displayed with the font size of 24 in Times New Roman with 1.5 line spacing. The environment for collecting eye-tracking data was equipped with stable light conditions and temperature, and the noise and vibrations were controlled to minimize the external harm to data quality (e.g., Holmqvist et al., 2011). Areas of interests (AOIs), which are areas selected for analysis, were defined in Data Viewer for data analysis. In the comparison between different modes, the screen and the assigned habitual CS terms were defined as AOIs to enable different overall comparisons and comparisons specifically associated with the CS terms.

Praat was applied to process the verbal data in this study. It is a programme designed for speech analysis that allows users to investigate the features of audio data and label the audio signals.

NASA-TLX was adopted to gain measures from participants' perceived difficulty of translation. The scale was developed Hart and Staveland (1988) and was frequently cited in the subjective measures of mental workload in interpreting. In our design, we adopted the original six subscales highlighting different types of workload, namely *Mental Demand, Physical Demand, Temporal Demand, Effort, Performance*, and *Frustration*, on a rating scale ranging from 0 to 100. The subscales can be specified by the questions centring on the perceived amount of mental and perceptual activities (e.g., thinking, deciding), the perceived amount of physical activity (e.g., controlling, activating), the perceived level of time pressure, the perceived successfulness in accomplishing the goals, the perceived level of required effort for accomplishing and the perceived level of frustrating feelings (e.g., insecurity, discouragement). A bilingual version of the questionnaire with Chinese translation of the original items was provided to the participants.

In order to find out the language background of our participants, Anderson et al.'s (2017) LSBQ was adopted. LSBQ was designed as a reliable and valid measurement tool to quantify bilingualism so as to examine how various aspects of people's language experience make up bilingualism. The original questionnaire includes six main sections that incorporate the respondents' personal demographic information as well as parents' education and occupation, language background, language use behaviour in home, community, daily activity, as well as language switching behaviour. The current study adapted the original design and left out the section of the personal and parents' demographic information as the section includes exposure of private details such as game-playing experience and physical/neurological impairments, and the part asking about the education and occupation level of their parents. Besides, some of the language use behaviours that were found to be redundant were omitted. A bilingual version of the adapted questionnaire with Chinese translation of the original items was provided to the participants.

A Bilingual Switching Questionnaire was adopted from Han et al. (2022), which is adapted from Rodriguez-Fornells et al. (2012) to collect information about the participants' code-switching habit. The questionnaire of the current study includes 12 items from the original questionnaire, which focuses on four aspects: (1) the tendency of switching from L2 to L1, (2) the tendency of switching from L1 to L2, (3) contextual code-switching measuring the switch frequency triggered by contextual cues and (4) unintended code-switching measuring the lack of consciousness of switches. Besides, the two items featuring the types of habitual code-switching,

https://doi.org/10.1017/langcog.2025.10056 Published online by Cambridge University Press

namely inter-sentential code-switching and intra-sentential code-switching, added by Han et al. (2022) were included. Participants were given five options (i.e., 'Never', 'Rarely', 'Sometimes', 'Often' and 'Always'), coded by 1 to 5 points. For easy understanding, all 'L1' and 'L2' in the original questionnaire design were replaced by 'Cantonese' and 'English', respectively, and a bilingual version of the questionnaire (with Chinese translation of the original items) was provided to the participants.

A retrospective questionnaire (hereafter, CS habitualness rating scale) was applied to verify the level of habitualness of the participants regarding the English code-switched words identified in this study. The verification approach was taken reference from Hui et al. (2022), which delivered a 7-point Likert scale questionnaire to the participants to rate whether a code-switched word in a sentence would be used in its code-switched form in L2 (English) or in L1 (Cantonese). A short part of the original excerpts where the English code-switched words were found from the eight episodes of the three radio programmes was listed in the questionnaire. Participants were invited to rate on a 7-point Likert scale ranging from 1, which means 'Strong habitual English use', to 7, signalling 'Strong habitual Cantonese use', at the two ends of the scale. The instances averagely rated 3 or lower were considered as habitual English terms in the context of Macao.

2.5. Procedures

All participants signed a consent form to indicate their consent that their performance in the experiment, including the audio and eye-tracking data, be used in the research. Besides, they filled in the demographic and language background questionnaire before the experiment. The experiment included a practice section and an experiment section. Each participant went through a calibration procedure before experiment section. This step served to collect the geometric characteristics of their eyes, which ensured the accuracy of the parameter calculation. Calibration was executed with 9-point calibration. The participants could proceed to the next stage of experiment only when the reported average error values were less than 0.5 degrees and the maximum error was less than 1 degree. The participants were instructed to try their best to stabilize their head and not move drastically throughout the experiment.

Before the experiment, participants went through the practice section to familiarize themselves with the experiment setting. During the practice section, participants were given four short sentences, in which participants were required to maintain using Cantonese for the entire production in the first half of the trials. Next, they were allowed to code switch in the second half. The four sentences were designed based on the eight episodes where the stimuli of the experiment were derived from but contained no targeted CS terms in the experiment to avoid the possibility of reminding the participants of the CS terms that they applied in daily life. Next, the participants proceeded to the experiment section of English-Cantonese sight translation task, which comprised two main parts: a part where participants were told that they were allowed to code-switch where they found the production of English words natural/hard to translate (hereafter 'CS mode'), and a part where they were asked to maintain using Cantonese (hereafter 'NCS mode') throughout the entire experiment. There were two trials (one paragraph each) for each mode in the experiment section. All trials in the practice section and the experiment section were

https://doi.org/10.1017/langcog.2025.10056 Published online by Cambridge University Press

self-paced. Participants pressed the spacebar of the keyboard that was connected to the computer whenever they finished translating the content on the page, and they can proceed to the next trial. Before each mode, there was an instruction page outlining the requirements of the mode. Participants were instructed to press the spacebar when they finished reading and were ready to begin. In order to direct participants' eye gaze to the beginning of the paragraphs before every trial, a white cross was displayed at the corresponding position against a black background after the button press. Participants were instructed to look at the cross during this stage. The researchers then controlled the monitor to display the paragraph after ensuring that participants' eye gaze was correctly positioned before proceeding. In order to guarantee reliability of the experiment, both the modes and stimuli were counterbalanced across participants.

After the experiment, a semi-structured interview was conducted, where participants were asked about their perception of the difficulty of the tasks in the CS mode and in the NCS mode. For those terms that the participants chose to code-switch during the experiment, the participants were further asked whether the code-switched terms were those they usually use in their life. Finally, the participants finished the NASA-TLX questionnaire. The CS habitualness rating scale was also delivered after the experiment to avoid any priming effect that could potentially influence the language choice during the experiment. Moreover, they finished the Bilingual Switching Questionnaire during their available time in a week after the experiment.

2.6. Data analysis

To ascertain the quality and validity of the eye movement data, stringent criteria were applied, referring to Hvelplund (2014). Three criteria were thus consulted before data analysis: (1) mean fixation duration (MFD), indicating the mean duration of every fixation during a task; (2) gaze time on screen (GTS), reflecting the percentage of total fixation duration in relation to the total task time, (3) gaze sample to fixation percentage (GFP), measuring the ratio of fixations relative to the total quantity of gaze samples. This study adopted the approach of Cui and Zheng (2020) in leaving out invalid data based on the criteria suggested by Hvelplund (2014). The mean and standard deviation of each criterion were calculated for CS mode (MFD: M = 251.82 ms, SD = 39.88 ms; GTS: M = 79.35%, SD = 7.52%; GFP: M = 80.28%, SD = 7.69%) and NCS mode (MFD: M = 257.09 ms, SD = 38.66 ms; GTS: M = 79.74%, SD = 6.79%; GFP: M = 80.54%, SD = 5.69%). In this study, data of each trial lower than one standard deviation below the mean was considered not fulfilling the corresponding criterion. If the data of either mode (i.e., CS mode or NCS mode) are not fulfilling two or more of these criteria (i.e., MFD, GTS, GFP), the corresponding eye-tracking data will be excluded from further analysis. In the end, eye-tracking data of three participants (P06, P13, P23) were deemed invalid and excluded.

In terms of the audio data, silent pauses, filled pauses, revisions, repetitions and prolonged sounds were collected as objective indicators of cognitive load (Bóna & Bakti, 2020). Praat was used to analyse the audio data. Audio recording was input into Praat and was manually analysed for the above parameters. The use of code-switched words by the participants was also used for a detailed analysis of their translation process and their code-switching behaviour.

https://doi.org/10.1017/langcog.2025.10056 Published online by Cambridge University Press

Shapiro–Wilk test was applied to assess the normality of the data before conducting inferential analysis. Parametric tests (e.g., Paired samples t-test, Pearson correlation) and non-parametric tests (e.g., Wilcoxon signed-rank test, Spearman correlation) were applied for normal and non-normal distribution respectively. IBM SPSS Statistics 27.0 was applied for descriptive and inferential statistics.

The data of this study will be provided to the reader(s) upon request.

3. Results

Results of the participants' perception of the habitualness of the CS terms in the stimuli, the eye-tracking data, audio data, as well as NASA-TLX data are presented in the following sections.

3.1. Perception of habitualness of code-switching

In order to guarantee the habitualness of code-switching among chosen terms in the stimuli, we collected data about the perception of the participants regarding the English code-switched words identified in this study. As mentioned above, a 7-point Likert scale questionnaire (CS habitualness rating scale) was used to ask the participants to rate the code-switched words regarding how habitually they used them in English. Each word was demonstrated with a context derived from the corresponding original excerpts from the eight episodes of the radio programmes (e.g., '仲有啲遊玩嘅 tips 可以畀到大家嘅…' [there are also some tips how the game is played for everyone]). The scale ranged from 1, which means 'Strong habitual English use', to 7, signalling 'Strong habitual Cantonese use'. The instances averagely rated 3 or lower were considered as habitual English terms in the context of Macao. The items that received an average rating between 3 and 5 (excluding 3) suggested that the Cantonese and English versions of the same term carried similar weight in Macao, indicating a moderate or inconsistent CS tendency. The items receiving an average rating between 1 and 3 were considered as items of strong habitual English use. According to the results, 58 of the CS words in the stimuli were used habitually in English by the participants. Among all the CS words, the terms that were considered the strongest habitual English use include 'café' (M = 1.37, SD = 0.93), 'hip-hop' (M = 1.37, SD = 1.1), 'story' (M = 1.4, SD = 1.1), 'AI', (M = 1.43, SD = 1.17), 'IG' (M = 1.43, SD = 1.17), 'link' (M = 1.47, SD = 1.17), 'app' (M = 1.47, SD = 1.31), 'Facebook' (M = 1.47, SD = 1.31), 'QR code' (M = 1.5, SD = 0.94) and 'logo' (M = 1.53, SD = 1.22). Besides 'café', 'hip-hop' and 'logo', the remaining seven strongest habitual English terms are all related to technology or social media. Ten terms with moderate CS tendency were identified ($p < .001$, compared with strong habitual English items): 'juicy' (M = 3.17, SD = 1.8), 'artist' (M = 3.2, SD = 1.83), 'romantic' (M = 3.5, SD = 1.55), 'gallery' (M = 3.5, SD = 1.8), 'fake news' (M = 3.6, SD = 1.83), 'overact' (M = 3.63, SD = 2.21), 'spiritual' (M = 3.83, SD = 2.1),), 'single' (M = 3.83, SD = 1.84), 'meditation' (M = 4.03, SD = 2) and 'first love' (M = 4.87, SD = 1.81). With an overall mean lower in value (M = 2.15, SD = 0.67), the internal consistency was confirmed with Cronbach's alpha ($\alpha = 0.969$). Moreover, the participants demonstrated similar language habits by Bilingual Switching Questionnaire. From the five dimensions of switching habits captured in the questionnaire, it can be seen that participants confirmed the habitualness of code-switching in daily life, especially in particular

https://doi.org/10.1017/langcog.2025.10056 Published online by Cambridge University Press

situations, and they switched from Cantonese to English more frequently than from English to Cantonese (L2-to-L1 switch: M = 2.72, SD = 0.55; L1-to-L2 switch: M = 3.31, SD = 0.6; contextual switch: M = 3.81, SD = 0.62). Participants reported a moderate level of awareness of their own code-switching behaviours, indicating they were neither fully aware nor fully unaware of these behaviours in their daily lives (unintended switch: M = 3.28, SD = 0.34). Moreover, they intrasententially switch languages (M = 1.9, SD = 0.66) more frequently than intersententially switch languages (M = 3.03, SD = 0.85).

3.2. Eye tracking data

Comparisons between CS mode and NCS mode were made based on overall eye-tracking data (i.e., defining the whole paragraphs as AOIs) and eye-tracking data specifically associated with CS terms (i.e., defining the assigned CS terms as AOIs) (See Table 2). Data of AOIs were averaged per trial, and then averaged across participants for each mode. Total fixation duration (i.e., the summation of duration across all fixations on the areas of interest) and fixation count (i.e., the total number of fixations falling in the areas of interest) on CS words in the two modes of sight translation tasks were examined, and average fixation duration was included for comparison of overall eye-tracking data. Results of overall eye-tracking data indicate that mean fixation duration, total fixation duration, and fixation count on screen were lower in CS mode than NCS mode. Specifically, difference total fixation duration was significantly higher in CS mode (Z = −2.354, p = .019), while mean fixation duration (t = −1.893, p = .07) and fixation count (Z = 1.898, p = .058) were marginally higher in CS mode. Comparisons focusing on CS terms were made based on the average data associated with the sum of each CS term, the results reveal that CS mode received significantly shorter total fixation duration (Z = −2.595, p = .009) and fewer fixation count (Z = −2.33, p = .02). Additionally, regressions in eye movement, defined as the backward eye movements to earlier text areas (Rayner, 1998) were briefly examined. No statistically significant difference was found between CS and NCS modes in the average regression count per trial for the CS terms (CS mode: M = 0.11, SD = 0.32; NCS mode: M = 0.14, SD = 0.35; p = .123). It is likely that, while

Table 2. Results of eye-tracking data

Measures	CS mode		NCS mode				Effect size
	M	*SD*	*M*	*SD*	*t/Z*	*p*	*d/r*
Overall eye-tracking data							
Mean fixation duration (ms)	258.27	36.39	264	33.01	−1.893 (t)	.070	−0.165 (*d*)
Total fixation duration (ms)	97318.37	21274.22	105028.64	23709.3	−2.354 (Z)	.019	−0.453 (*r*)
Fixation count	383.20	84.42	408	106.11	−1.898 (Z)	.058	−0.365 (*r*)
Eye-tracking data on CS terms							
Total fixation duration (ms)	1391.08	358.46	1601.93	492.68	−2.595 (Z)	.009	−0.499 (*r*)
Fixation count	4.92	1.13	5.52	1.7	−2.33 (Z)	.020	−0.448 (*r*)

https://doi.org/10.1017/langcog.2025.10056 Published online by Cambridge University Press

attempting to complete the task, participants did not read backward but remained focused on the words they were currently processing.

A Mann–Whitney U test was conducted to compare the cognitive load indicated by eye-tracking data related to the strong habitual English terms and the neutral terms. In CS mode, it was found that total fixation count associated with the neutral terms (Mean rank = 46.05) was significantly higher than habitual English terms (Mean rank = 32.51) ($U = 174.5$, $p = .045$). Total fixation duration on neutral terms (Mean rank = 44.50) was longer than on the habitual English terms (Mean rank = 32.78) (marginally significant: $U = 190$, $p = .083$).

In addition, Pearson and Spearman correlation was performed to investigate the relationship between number of English words and cognitive load indicated by eye-tracking data. Pearson correlation was applied to data that met the assumption of normality, while Spearman correlation was used for data that did not follow a normal distribution. However, no significant correlation was found in both modes.

3.3. Audio data

In terms of the comparison of the audio data, there are 8 variables that were found to exhibit statistically significant difference, namely total English words, total English syllables, total Cantonese words, total task time, silent pause count (i.e., the total number of occurrences of silent pauses), mean silent pause duration, percentage of silent pause duration (i.e., duration of silent pause divided by total phonation duration in a speech sample), as well as mean prolonged sound duration (i.e., the average length of each prolonged sound instance). All values were first calculated per trial, then averaged per participant and finally averaged across participants for each mode.

While the first four variables are related to the entire number of words and syllables produced, as well as the time involved in the experiments, silent pause count, mean silent pause duration, silent pause duration percentage and prolonged sound duration produced during the sight translation process were identified and calculated as an indicator of cognitive load. Pauses with a duration higher than 300 ms were identified as silent pause. Sounds (i.e., syllables) lasting longer than 300 ms were considered a prolonged sound in this study. Paired-sample t-test was applied to compare CS mode and NCS mode regarding the total Cantonese Chinese characters and silent pause count. Wilcoxon signed-rank test was adopted to compare the means between groups for the other six variables, including total English words, total English syllables, total task time, mean of silent pause duration, percentage of silent pause duration, as well as mean of prolonged sound duration

With these eight variables, there is a statistically significant difference between the mode of allowing code-switching (CS mode) and the mode not allowing code-switching (NCS mode). As can be referred in Table 3, the voluntary use of code-switching results in a reduction in the production of disfluency markers in sight translation, which enhances the speed of the translation and implies the reduction of the cognitive load. In addition, when compared to NCS mode, participants in CS mode were found to produce significantly fewer words in the entire experiment and completed the translation task in a significantly shorter time.

In terms of the use of code-switched words by the participants, 562 English CS terms were used by the participants in their sight translation. Among those terms,

https://doi.org/10.1017/langcog.2025.10056 Published online by Cambridge University Press

Table 3. Results of audio data

Measures	CS mode		NCS mode				Effect size
	M	*SD*	*M*	*SD*	*t/Z*	*p*	*d/r*
Total English words in the output speech	9.22	5.23	1.33	2.09	−4.500 (Z)	<.001	−0.822 (*r*)
Total English syllables in the output speech	16.20	9.37	2.33	3.74	−4.435 (Z)	<.001	−0.810 (*r*)
Total Cantonese Chinese characters in the output speech	312.98	55.38	332.35	39.62	−3.154 (t)	.004	−0.402 (*d*)
Total task time (min)	1.97	0.39	2.13	0.53	−2.273 (Z)	.023	−0.415 (*r*)
Silent pause count	41.23	11.23	45.17	13.31	−2.472 (t)	.020	−0.320 (*d*)
Mean silent pause duration (s)	0.93	0.31	1.02	0.40	−2.293 (Z)	.022	−0.419 (*r*)
Silent pause duration %	0.30	0.13	0.32	0.15	−2.581 (Z)	.010	−0.471 (*r*)
Mean prolonged sound duration (s)	0.46	0.20	0.52	0.21	−2.232 (Z)	.026	−0.408 (*r*)

475.5 of them (84.61%) were the code-switched terms assigned by us as the habitual CS terms in the stimuli. In the experiment, not all participants made use of the complete CS term that we originally designed for the stimuli: some of them chose to code-switch half of the term and thus the value of 0.5 was recorded because not all participants would use the complete code-switched terms as used in the stimuli. For example, out of the 30 instances of *tiffany blue* from the stimuli, while 11 of the participants (36.67%) maintained the use of *tiffany blue* when they were in the CS mode, three of them (10%) were translated into *tiffany* + 藍 (laam4 – equivalent to 'blue'). Other translations included terms like different forms of the blue colour including 天藍色 (tin1 laam4 sik1 – equivalent to 'sky blue'), 湖水藍 (wu4 seui2 laam4 – equivalent to 'aqua blue'), 湖綠色 (wu4 luk6 sik1 – equivalent to "turquoise"), 寶藍色 (bou2 laam4 sik1, equivalent to 'royal blue') and 鐵芬尼藍/蒂凡尼藍 (tit3 fan1 nei4 laam4/dai3 faan4 nei4 laam4, which is the homophonic translation of 'tiffany' plus the word 'blue'). Another similar instance occurred with *check messages*: while most participants (25 out of 30) made the translation of the term, there were five instances of the use of code-switching, two of which involved the use of the original term *check* message, while the other two of them were translated into *check* + 訊息 (seon3 sik1 – equivalent to 'message'). There was one instance where 'check' and 'message' were used with the Chinese particles 返呢啲 (faan1 ni1 di1 – equivalent to 're-check these'), making the combination of *check* + 返呢啲.

Among the results, one thing worth noticing is that out of the 68 CS terms used in the stimuli, the 10 most frequently used terms by the participants in the sight translation included 'follow' (altogether 22 occurrences), 'IG/Instagram' (21 occurrences), 'Facebook' (21 occurrences), 'QR code' (17 occurrences), 'send' (17 occurrences), 'coupon' (17 occurrences), 'post' (17 occurrences), 'tiffany blue' (16 occurrences), 'logo' (15 occurrences) and 'AI' (14 occurrences). Among these 10 items, the top three items are related to social media, while the other four of them are related to technology. It should be noted the design of the stimuli span across topic, which include but not limited to social media and technology. In addition, participants are not prompted to code-switch or not to code-switch any terms.

https://doi.org/10.1017/langcog.2025.10056 Published online by Cambridge University Press

In order to investigate whether code switching more is significantly associated with decrease of cognitive load, the correlation between number of English words used in oral production and the cognitive load indicators were calculated. Spearman correlation was conducted and correlation was only found between mean prolonged sound duration and the number of English words used ($r_s(30) = -.407$, $p = .026$).

3.4 NASA-TLX data

NASA-TLX data were collected to reflect the self-perceived cognitive load of participants derived from the two modes of cognitive load. Table 4 demonstrated the results of every construct. It can be observed by the means that participants perceived higher level of cognitive load in NCS mode than in CS mode in general. Significant difference was identified in physical demand ($Z = -2.681$, $p = .007$), effort ($Z = -2.616$, $p = .009$), frustration ($t = -2.508$, $p = .018$) and mental demand ($Z = -4.323$, $p = < .001$). Marginally significant difference was identified in temporal demand ($t = -1.893$, $p = .068$). The construct of performance did not show significant difference between the two modes.

4. Discussion

Extensive research on bilingual language processing has contributed insights into the language control mechanism behind comprehension and speech production of bilinguals. The level of cognitive load related to language and code-switching remains to be a heated topic due to the inconsistent results and various interpretations from previous studies. Among the studies investigating language switching and code-switching, only a very limited number of them have shed light on the habitualness of the code-switched terms in the design of the stimuli. This research attempted to explore the influence of habitual code-switching on cognitive load during sight translation task, which not only enabled investigation from triangulated data but also allowed us to see if participants have the tendency to code switch particular habitual CS terms in sight translation to stabilize the cognitive load level.

As is reflected by the results, participants in general had lower cognitive load during sight translation if they were allowed to apply the habitually used code-switched English words in the production. The results of this study align with the studies suggesting that voluntarily mixing language is not cognitively costly (e.g., Blanco-Elorrieta & Pylkkänen, 2017; Jevtović et al., 2020; Lin et al., 2018). Our results also specifically echoed findings of studies that pointed out the low cognitive load of

Table 4. Results of NASA-TLX data

Measures	CS mode sight translation		NCS mode sight translation				Effect size
	M	*SD*	*M*	*SD*	*t/Z*	*p*	*d/r*
Physical demand	31.70	21.13	41.23	25.01	−2.681 (Z)	.007	−0.489 (*r*)
Temporal demand	37.73	25.87	46.40	26.77	−1.893 (t)	.068	−0.329 (*d*)
Effort	56.40	23.66	67.10	20.10	−2.616 (Z)	.009	−0.478 (*r*)
Frustration	33.70	26.33	42.47	20.23	−2.508 (t)	.018	−0.374 (*d*)
Mental demand	42.77	21.38	64.97	20.45	−4.323 (Z)	<.001	−0.789 (*r*)
Performance	70.27	13.16	66.57	19.16	1.045 (t)	.305	0.225 (*d*)

https://doi.org/10.1017/langcog.2025.10056 Published online by Cambridge University Press

processing language stimuli with familiar code-switching patterns. While these studies focused on perception tasks, our research suggested that the tendency also applies to tasks involving production (Adamou & Shen, 2019; Hui et al., 2022; Kleinman & Gollan, 2016). It is worth noting that the stimuli in this study were designed with contexts in sentences, and the target texts were predicted to reflect speech production in real-life context with the habitual English code-switched words. Compared with the experiment design such as picture-naming task and numeral naming task in studies investigating bilingual language processing, the stimuli of this study should provide more accurate and authentic reflection of actual language control mechanism of bilinguals. The results of this study are also in line with Cheung's (2001) study, in which trainee interpreters showed better performance when they were allowed to code-switch. Although Cheung's (2001) study coupled verbal data with the assessment of the interpreting performance from raters, the measurements were overall either binary (e.g., grading sheet of performance featuring only good and bad) or too general (e.g., frequency of hesitation, pauses, and unfinished sentences were provided with a range of '>/< 5' or '>/< 10'). This study stepped further to triangulate the data with eye movements, audio data, and NASA-TLX results, which contributed to providing a more reliable interpretation of the phenomenon, and cast more light on bilingual language control itself. In particular, the NASA-TLX scores, as a subjective measure of cognitive load, were generally consistent with the other objective measures in this study, which aligns with the positive correlation shown in a previous study (Liu et al., 2019). However, the study of Sun and Shreve (2014) has also reported only weak correlation between subjective NASA-TLX scores and objective indicators, suggesting higher quality does not necessarily correspond to lower perceived difficulty. This underscores the importance of incorporating both subjective and objective indicators when examining cognitive load in bilingual language control, in order to achieve a more comprehensive analysis.

Although the results of this study showed a tendency of reduced cognitive load in the mode allowing code-switching, they did not demonstrate an absolute reduction as was reflected by the limited correlation between the objective cognitive load indicators and frequency of code-switching. We thus speculate that interpreters are not entirely flexible in switching between languages; rather, they are constrained by the different accessibility levels of the lexicons in the two languages. As was reflected by the participants in the post-experiment interview, some lexicons of English may have already been integrated into the mental lexicon of Cantonese in the Cantonese-English bilingual community. Such highly habitual English terms were thus treated as a part of daily language use. According to the comparison of eye-tracking data in CS mode between habitual English terms and neutral terms, participants exhibited higher cognitive load when they processed the neutral terms. This indicates that participants probably have the habitual English words at hand, and the Cantonese-translation equivalents of which are less accessible to the extent that they are not of normal use in daily communication. In this case, the participants might not need to exert much cognitive effort on choosing the appropriate language in the sight translation task when they were allowed to code switch as the English terms match their habitual language use. In the cases of processing the neutral terms, the accessibility of both Cantonese and English versions is similar. In other words, no language counterpart is significantly dominating over the other. In this case, participants might have to hesitate in their language choice. The necessity of selecting the

https://doi.org/10.1017/langcog.2025.10056 Published online by Cambridge University Press

word for delivery might cost certain cognitive load, which explains why there was not an absolute reduction in cognitive load even though code-switching was allowed.

The habitualness of the code-switched terms and the use of them in the production during the sight translation process are one of the highlights of this research. Based on the results of both the participants' perception of habitualness of code-switching among terms used in the stimuli, and their use of those terms during the experiments, it has been found that a high proportion of words that were rated to carry "strong habitual English use" were also words that were used more by the participants in the sight translation experiment. For example, among the top ten terms that were rated as 'more strongly habitual English use', five of them were used with a high proportion. Therefore, we can conclude that the CS words assigned in the stimuli were in general effective representatives of the habitual CS words among the participants, as participants considered the terms used in the stimuli are terms that are usually code-switched in the daily life of Macao locals.

Based on the results of the experiment, participants code switched terms related to social media (like 'send', 'IG/Instagram', 'follow', 'post', 'Facebook', 'story') and technology (like 'zoom', 'QR code', 'scan') more. To some extent, it implies that the participants commonly use English CS terms when they are talking about social media. In fact, as reflected by the participants, it is rather hard for the general public to think of the Cantonese equivalents to some of the names of social media platforms (like 'Instagram') or online behaviour related to social media (like 'follow'). In addition, participants tended to code switch terms that are more often expressed in English in their daily life (e.g., 'hip-hop', 'coupon', 'logo', 'channel', 'warm up') or with a shorter syllabic structure in the English CS terms (e.g., 'send', 'AI', 'café'). The above results suggesting the reasons for code-switching align with other studies about code-switching among Cantonese speakers (e.g., Chan, 2011; Lee, 2000), which suggest that Cantonese speakers in Hong Kong and Macao tended to use English CS terms when they refer to ideas related to technology, lifestyle, and western culture. Under such situation, they might just 'quote' the English terms that they learnt or came across first in English in their daily Cantonese speech without even noticing their code-switching or using English (Chan, 2011). Thus, the use of these mental lexicons gives them the ease in their daily communication.

It can be observed in the experiment that when it comes to compound terms, some participants would tend to code-switch only part of them. In the case of 'tiffany blue', for example, participants would tend to use the code-switched term 'tiffany' but translate the second part into its Cantonese equivalent. According to some of our participants in the semi-structured interview, it is not easy for them to immediately think of the Cantonese equivalent of 'tiffany' but for 'blue', as it is a common term in Cantonese, it is easy for them to make use of the Cantonese translation directly. As for the case of *check messages*, where a number of participants code-switched *check* but translated *messages*, we can speculate that participants had the habit of using 'check' in English because of its simple mono-syllabic term in the CS term (cf 檢查 gim2 caa4 in Cantonese), while in the case of 'message', both the English code-switched term and the Cantonese equivalent contain two syllables, which makes using the English CS term without much advantage in convenience, hence the use of half English, half Chinese.

Based on the results, we obtain a clearer picture of the code-switching mechanism in the context of Macao, which appears consistent with the cooperative language control state (coupled control) in bilingual language control mechanism (e.g., Green,

https://doi.org/10.1017/langcog.2025.10056 Published online by Cambridge University Press

2018; Green & Abutalebi, 2013; Green & Li, 2014). In comparison with mandatory code-switching, voluntary code-switching recruits less effort to inhibit the non-target language. Freely switching involves opportunistic planning by which the participants can make use of the most accessible lexicon in mind and incorporate it into the sentence structure of the dominant language. This, however, does not lead to the conclusion that voluntarily switching languages requires no inhibitory control, as suggested by de Bruin et al. (2018). In the control process model developed by Green (2018) and Green and Li (2014), a subtype of cooperative control state named "coupled control" is introduced, highlighting the code-switching behaviour of insertion and alternation. This type of control state stresses the need to select items and construct language. Switching, in this sense, is analogous to temporarily ceding the control on the entering of certain items where appropriateness is fulfilled. In this study, such switching patterns align with the code-switching habits observed among the participants. The cognitive demands involved in this temporary switching may be modulated by the degree of habitualness of English CS terms in Macao. When the English CS terms exhibit a tendency of habitual English use but not reaching a strong level of habitualness, the code-switchers appeared to exert slightly more cognitive effort in choosing the language for oral rendition. In contrast, highly habitual English terms were associated with fewer hesitations and were used more. For example, highly habitual English terms in the context of Macao such as 'café' and 'hip-hop' were used more frequently and triggered less cognitive load than the terms with moderate habitual English use such as 'artist' and 'gallery'. This mode is different from the other subtype of cooperative control state, namely open control, which is characterized by the intertwined morphosyntax of the two languages and the subsequent minimized switching costs. In this study, participants tended to switch from Cantonese to English for concepts habitually expressed in English in their daily lives and commonly shared within the language community. Such temporary ceding of staying in one language was likely motivated by the aim of reducing cognitive load and maintaining naturalness within the bilingual context of Macao. To conceptualize the phenomenon following the metaphor of CPM, when the participants tried to 'push open' the door of the English counterparts of the neutral terms, there could be some resistance derived from the Cantonese equivalents, which prevented absolute reduction of cognitive load even when code-switching was allowed. However, further research with a more specific focus will be necessary to verify this pattern.

Several implications are provided here based on the results of the study. First, the importance of triangulating data with both objective and subjective data should be taken into consideration in empirical studies attempting to investigate or examine the bilingual switching mechanism. Data triangulation not only allows more comprehensive and reliable interpretation of the language application patterns of bilinguals but also provides opportunities to look into some overlooked phenomena that are worth researching themselves, e.g., the misalignment between self-reported cognitive load and objectively measured cognitive load could be accounted by participants' misperception of language switching habit to some extent. Moreover, the current study has paved the way for future studies focusing on more detailed delineation of bilingual control mechanism pertaining to a specific cultural background. For example, studies can dig into how the potentially different levels of cognitive load associate with the habitualness levels of different code-switched English words and patterns among the participants in a speech community.

https://doi.org/10.1017/langcog.2025.10056 Published online by Cambridge University Press

Lastly, several limitations of this study should be pointed out. First, this study adopted a sight translation task as a window to observe the language control pattern of bilinguals that habitually switch languages in their daily life, which allowed us to investigate the relationship between habitualness level of the CS terms and the corresponding cognitive load level in code-switching behaviours. However, the code-switching patterns and the associated cognitive load levels could be different from natural production. Future studies can consider tasks adhering to natural production such as mocking conversation and storytelling based on daily topics. Second, the participants in this study had a background of English major, which could induce bias in the results. This is to be solved by recruiting general public and applying more natural production task in the future. Third, terms such as 'IG' may not be as 'translatable' as other terms in the stimuli, which potentially diverse the translation difficulty of the targeted CS terms. A more prudent procedure of selecting the and categorizing the CS terms is needed in the future. Lastly, the studies relied on behavioural data and subjective data to reflect cognitive load, which did not reveal brain activities. Contrast of cognitive load level in different modes and the association between cognitive load and code-switching frequency could be more significant in brain activation than in behaviours. For future studies, neuroimaging techniques such as EEG and fMRI can provide a more comprehensive and convincing delineation of the code-switching phenomenon.

5. Conclusion

This study revealed the relationship between habitualness of CS terms and cognitive load level in Cantonese-English sight translation tasks among native Cantonese speakers in Macao. With eye-tracking data, audio data and NASA-TLX data, we provided partial evidence that the native Cantonese speakers in Macao tended to experience lower cognitive load when translating words they habitually used in English rather than in Cantonese. Voluntary code-switching does not necessarily indicate an absolute reduction of cognitive load; instead, the habitualness level of CS terms may have a modulating effect on the level of cognitive load in code-switching. Such patterns appear to be compatible with the cooperative control state and coupled control mode in the realm of bilingual language control mechanism. Anchored in the sociolinguistic context of Macao, the findings contribute to our understanding of bilingual language control by demonstrating how the habitualness of code-switched terms may influence the selective use of code-switching in language output.

Funding statement. This work was supported by University of Macau grants (MYRG-CRG2024-00047-ICI, MYRG-GRG2023-00239-FAH, CPG2023-00016-FAH, MYRG2022-00200-FAH, CRG2021-00001-ICI).

Data availability statement. Access to the data supporting this study is restricted due to its potential future use, but the data are available from the corresponding author upon reasonable request and subsequent evaluation.

References

Adamou, E., & Shen, X. R. (2019). There are no language switching costs when codeswitching is frequent. *International Journal of Bilingualism*, 23(1), 53–70. https://doi.org/10.1177/1367006917709094.

Anderson, J. A. E., Mak, L., Keyvani Chahi, A., & Bialystok, E. (2018). The language and social background questionnaire: Assessing degree of bilingualism in a diverse population. *Behavior Research Methods*, 50(1), 250–263. https://doi.org/10.3758/s13428-017-0867-9.

https://doi.org/10.1017/langcog.2025.10056 Published online by Cambridge University Press

Berthold, A., & Jameson, A. (1999). Interpreting symptoms of cognitive load in speech input. In *UM99 user modeling* (pp. 235–244). Springer. https://doi.org/10.1007/978-3-7091-2490-1_23.

Bialystok, E., Craik, F. I., & Luk, G. (2012). Bilingualism: Consequences for mind and brain. *Trends in Cognitive Science*, 16(4), 240–250. https://doi.org/10.1016/j.tics.2012.03.001.

Blanco-Elorrieta, E., & Pylkkänen, L. (2017). Bilingual language switching in the laboratory versus in the wild: The spatiotemporal dynamics of adaptive language control. *Journal of Neuroscience*, 37(37), 9022–9036. https://doi.org/10.1523/JNEUROSCI.0553-17.2017.

Bóna, J., & Bakti, M. (2020). The effect of cognitive load on temporal and disfluency patterns of speech. *Target*, 32(3), 482–506. https://doi.org/10.1075/target.19041.bon.

Boos, M., Kobi, M., Elmer, S., & Jancke, L. (2022). Tracking lexical access and code switching in multilingual participants with different degrees of simultaneous interpretation expertise. *European Journal of Neuroscience*, 56(6), 4869–4888. https://doi.org/10.1111/ejn.15786.

Botha, W. (2013). English-medium instruction at a university in Macau: Policy and realities. *World Englishes*, 32(4), 461–475. https://doi.org/10.1111/weng.12053.

Bolton, K. (2003). *Chinese Englishes: A sociolinguistic history*. Cambridge University Press.

Chan, B. H.-S. (2011). English in Hong Kong Cantopop: Code-switching, pop culture and the local identity of Hong Kong Chinese. In J. Lee & A. Moody (Eds.), *English in Asian popular culture* (pp. 35–57). Hong Kong University Press.

Chan, B. H.-S. (2015). A local voice of Macau: Traditional characters, code-switching and written Cantonese in an internet forum. *Global Chinese*, 1(2), 281–310. https://doi.org/10.1515/glochi-2015-1013.

Chen, S. (2017). The construct of cognitive load in interpreting and its measurement. *Perspectives*, 25(4), 640–657. https://doi.org/10.1080/0907676x.2016.1278026.

Cheung, A. K.-F. (2001). Code-mixing and simultaneous interpretation training. *Interpreters' Newsletter*, 11, 57–62. http://hdl.handle.net/10077/2447.

Cheng, X. H., & Liu, X. B. (2010). 澳門三語流通與中文的健康發展. [The healthy development of the common use of the three languages and Chinese.]. In X. H. Cheng (Ed.), *The selection of studies in humanities and social sciences of Macau: Language and translation* (pp. 611–622). Social Sciences Academic Press.

Chmiel, A., & Mazur, I. (2013). Eye tracking sight translation performed by trainee interpreters. In C. Way, S. Vandepitte, R. Meylaertes, & M. Bartłomiejczyk (Eds.), *Tracks and treks in translation studies: Selected papers from the EST congress, Leuven 2010* (pp. 189–205). John Benjamins Publishing Company. https://doi.org/10.1075/btl.108.10chm.

Costa, A., & Santesteban, M. (2004). Lexical access in bilingual speech production: Evidence from language switching in highly proficient bilinguals and L2 learners. *Journal of Memory and Language*, 50(4), 491–511. https://doi.org/10.1016/j.jml.2004.02.002.

Cui, Y. X., & Zheng, B. H. (2020). Consultation behaviour with online resources in English-Chinese translation: An eye-tracking, screen-recording and retrospective study. *Perspectives*, 29(5), 740–760. https://doi.org/10.1080/0907676x.2020.1760899.

de Bruin, A., Samuel, A. G., & Duñabeitia, J. A. (2018). Voluntary language switching: When and why do bilinguals switch between their languages? *Journal of Memory and Language*, 103, 28–43. https://doi.org/10.1016/j.jml.2018.07.005.

DSEC. (2022). Detailed results of 2021 population census (revised version). https://www.dsec.gov.mo/getAttachment/6cb29f2f-524a-488f-aed3-4d7207bb109e/E_CEN_PUB_2021_Y.aspx.

Gile, D. (2009). *Basic concepts and models for interpreter and translator training*. John Benjamins Publishing Company.

Gollan, T. H., & Ferreira, V. S. (2009). Should I stay or should I switch? A cost-benefit analysis of voluntary language switching in young and aging bilinguals. *Journal of Experimental Psychology Learning Memory and Cognition*, 35(3), 640–665. https://doi.org/10.1037/a0014981.

Green, D. W. (2018). Language control and code-switching. *Language*, 3(2). https://doi.org/10.3390/languages3020008.

Green, D. W., & Abutalebi, J. (2013). Language control in bilinguals: The adaptive control hypothesis. *Journal of Cognitive Psychology*, 25(5), 515–530. https://doi.org/10.1080/20445911.2013.796377.

Green, D. W., & Li, W. (2014). A control process model of code-switching. *Language, Cognition and Neuroscience*, 29(4), 499–511. https://doi.org/10.1080/23273798.2014.882515.

https://doi.org/10.1017/langcog.2025.10056 Published online by Cambridge University Press

Gross, M., & Kaushanskaya, M. (2015). Voluntary language switching in English–Spanish bilingual children. *Journal of Cognitive Psychology*, 27(8), 992–1013. https://doi.org/10.1080/20445911.2015.1074242.

Han, X., Li, W., & Filippi, R. (2022). The effects of habitual code-switching in bilingual language production on cognitive control. *Bilingualism: Language and Cognition*, 25(5), 869–889. https://doi.org/10.1017/s1366728922000244.

Hart, S. G., & Staveland, L. E. (1988). Development of NASA-TLX (task load index): Results of empirical and theoretical research. In P. A. Hancock & N. Meshkati (Eds.), *Human mental workload* (pp. 139–183). North-Holland.

Holmqvist, K., Nyström, M., Andersson, R., Dewhurst, R., Jarodzka, H., & Van de Weijer, J. (2011). *Eye tracking: A comprehensive guide to methods and measures*. Oxford University Press.

Hui, N.-Y., Fong, M. C.-M., & Wang, W. S. (2022). Bilingual prefabs: No switching cost was found in Cantonese–English habitual code-switching in Hong Kong. *Language*, 7(3). https://doi.org/10.3390/languages7030198.

Hvelplund, K. (2014). Eye tracking and the translation process: Reflections on the analysis and interpretation of eye-tracking data. *MonTI. Monografías de Traducción e Interpretación*, 201–223. https://doi.org/10.6035/MonTI.2014.ne1.6.

Inhoff, A. W., Solomon, M., Radach, R., & Seymour, B. A. (2011). Temporal dynamics of the eye–voice span and eye movement control during oral reading. *Journal of Cognitive Psychology*, 23(5), 543–558. https://doi.org/10.1080/20445911.2011.546782.

Jevtović, M., Duñabeitia, J. A., & de Bruin, A. (2020). How do bilinguals switch between languages in different interactional contexts? A comparison between voluntary and mandatory language switching. *Bilingualism: Language and Cognition*, 23(2), 401–413. https://doi.org/10.1017/s1366728919000191.

Kleinman, D., & Gollan, T. H. (2016). Speaking two languages for the price of one: Bypassing language control mechanisms via accessibility-driven switches. *Psychological Science*, 27(5), 700–714. https://doi.org/10.1177/0956797616634633.

Lee, M. (2000). Code-switching in media texts: Its implications on society and culture in postcolonial Hong Kong. In D. Li, A. Lin, & W.-K. Tsang (Eds.), *Language and education in postcolonial Hong Kong* (pp. 95–127). The Linguistic Society of Hong Kong.

Lin, X., Lei, V. L. C., Li, D., & Yuan, Z. (2018). Which is more costly in Chinese to English simultaneous interpreting, "pairing" or "transphrasing"? Evidence from an fNIRS neuroimaging study. *Neurophotonics*, 5(2), 025010. https://doi.org/10.1117/1.NPh.5.2.025010.

Liu, Y., Zheng, B., & Zhou, H. (2019). Measuring the difficulty of text translation: The combination of text-focused and translator-oriented approaches. *Target: International Journal of Translation Studies*, 31(1), 125–149.

Ng, S., & Deng, F. (2017, August 22). Language and Power. *Oxford Research Encyclopedia of Communication*. Retrieved 12 Jan. 2026, from https://oxfordre.com/communication/view/10.1093/acrefore/9780190228613.001.0001/acrefore-9780190228613-e-436.

Meuter, R. F. I., & Allport, A. (1999). Bilingual language switching in naming: Asymmetrical costs of language selection. *Journal of Memory and Language*, 40(1), 25–40. https://doi.org/10.1006/jmla.1998.2602.

Muysken, P. (2000). *Bilingual speech: A typology of code-mixing*. Cambridge University Press.

Moody, A. (2008). Macau English: Status, functions and forms. *English Today*, 24(3), 3–15. https://doi.org/10.1017/S0266078408000242.

Nour, S., Struys, E., & Stengers, H. (2020). Adaptive control in interpreters: Assessing the impact of training and experience on working memory. *Bilingualism: Language and Cognition*, 23(4), 772–779. https://doi.org/10.1017/s1366728920000127.

Paradis, M. (1994). Toward a neurolinguistic theory of simultaneous translation: The framework. *International Journal of PsychoLinguistics*, 10, 319–335.

Perani, D., & Abutalebi, J. (2015). Bilingualism, dementia, cognitive and neural reserve. *Current Opinion in Neurology*, 28(6), 618–625. https://journals.lww.com/co-neurology/fulltext/2015/12000/bilingualism,_dementia,_cognitive_and_neural.12.aspx.

Proverbio, A. M., Leoni, G., & Zani, A. (2004). Language switching mechanisms in simultaneous interpreters: An ERP study. *Neuropsychologia*, 42(12), 1636–1656. https://doi.org/10.1016/j.neuropsychologia.2004.04.013.

Rayner, K. (1998). Eye movements in reading and information processing: 20 years of research. *Psychological Bulletin*, 124(3), 372–422. https://doi.org/10.1037/0033-2909.124.3.372.

https://doi.org/10.1017/langcog.2025.10056 Published online by Cambridge University Press

Rodriguez-Fornells, A., Kramer, U. M., Lorenzo-Seva, U., Festman, J., & Munte, T. F. (2012). Self-assessment of individual differences in language switching. *Frontiors in Psychology*, 2(388), 1–12. https://doi.org/10.3389/fpsyg.2011.00388.

Seeber, K. (2013). Cognitive load in simultaneous interpreting: Measures and methods. *Target*, 25(1), 18–32.

Sun, S., & Shreve, G. M. (2014). Measuring translation difficulty. *Target: International Journal of Translation Studies*, 26(1), 98–127. https://doi.org/10.1075/target.26.1.04sun.

Wheeler, J. (2019). Macau or Macao? – A case study in the fluidity of how languages interact in Macau SAR. *Journal of English as an International Language*, 14(1), 73–91.

Cite this article: Lei, U. M., Chan, K. L., & Lei, V. L. C. (2026). Habitual code-switching between English and Cantonese: bilingual language control reflected by cognitive load in sight translation, *Language and Cognition*, 18, e9, 1–23. https://doi.org/10.1017/langcog.2025.10056

https://doi.org/10.1017/langcog.2025.10056 Published online by Cambridge University Press