

「論文」

A Learner Corpus-Based Study of L1 Effects on L2 English Auxiliary Verb Use —The Case of *Will*—

Laurence NEWBERY-PAYTON

Abstract

This study conducts Contrastive Interlanguage Analysis of written essays by L1 Chinese and L1 Japanese learners of English contained in the International Corpus Network of Asian Learners of English (ICNALE). The two groups of learners are compared with each other and with native speakers regarding their use of the modal auxiliary *will*. Consideration of relevant characteristics of Chinese and Japanese suggests that L1 Chinese learners will overuse *will* due to functional similarities with the Chinese modal auxiliary *hui*, whereas this trend is not predicted to occur among L1 Japanese learners. Analysis of the corpus data reveals that Chinese L1 learners do overuse *will* at lower proficiency levels, providing evidence for crosslinguistic influence. In contrast, Japanese L1 learners, who lack functional equivalents to *will* in their native language, exhibit underuse as well as omission in obligatory contexts. The study therefore confirms the hypothesis that at lower proficiency levels, the presence or absence in L1 of partial functional equivalents to a target form can affect the latter's frequency of use in L2. However, these trends are restricted to one of two essay tasks, suggesting task-related factors.

1. Introduction

It has long been recognized by researchers that acquisition of a foreign language can be influenced by learners' native languages as well as any other languages previously acquired (Jarvis & Pavlenko, 2008; Luk & Shirai, 2009). This paper uses corpus data to conduct Contrastive Interlanguage Analysis of writing by Chinese L1 and Japanese L1 learners of English. In particular, it attempts to ascertain the presence or absence of crosslinguistic influence in learners' use of the modal auxiliary verb *will*.

Modal verbs have frequently attracted attention from researchers due to the difficulties learners face in reaching nativelike use, both quantitatively and qualitatively. The current paper selects *will* as the focus of analysis because the acquisition of this modal auxiliary is predicted to present different sets of difficulties for the two groups of learners.

2. Literature Review

This section reviews previous studies relevant to the focus of the present study. Section 2.1 considers corpus studies of L2 modal verb use. Section 2.2 examines modal auxiliaries in Chinese and their similarities to the English modal *will*. Section 2.3 combines the conclusions of the two preceding sections to explain the rationale behind the current study and its hypotheses.

2.1 Corpus Studies of L2 English Modal Verb Use

This section provides an overview of corpus-based studies examining the use of modal verbs by Japanese learners of English (JLE) and Chinese learners of English (CLE).

Nakayama (2020) compares JLE and two groups of native speakers (students and teachers) using ICNALE's written component. JLE are found to overuse *can*, *should* and *must*, but underuse *will* and *would*. Nakayama suggests this reflects the greater difficulty of epistemic modality markers, but does not consider learners' proficiency levels. Nakayama (2021), using the spoken module of ICNALE, finds that JLE at A2 and B1 levels underuse *could*, *might*, *would* and *will*, and use modal verbs to express deontic modality more frequently than epistemic modality, contrasting with native speakers. While Nakayama provides analysis for a selection of individual verbs, there is no specific explanation for the underuse of *will*.

Xiao (2017) compares data from learner and native corpora and reports that CLE overuse *must*, *should*, *will* and *can*, but underuse *would*, *might* and *could* in their writing. Likewise, in spoken language, CLE overuse *must*, *should*, *will* and *can*, but underuse *would* and *might*. Xiao adopts an analytical framework from functional grammar, which groups *will* with other "middle-value" modals, *would* and *shall*. As a result, the analysis cannot adequately explain the high frequency of use of *will*.

Yang (2018) reports that modal verbs appear more frequently in learners' academic writing than in published academic papers, and that learners overuse *can*, *will*, *could* and *would*. Yang suggests that one-to-one translations of modal verbs in course books may cause pragmatically inappropriate uses of *should* by CLE (p. 127).

Taken together, the above studies appear to show trends towards underuse and overuse of *will* by JLE and CLE respectively. This paper aims to directly compare the two learner groups using a unified data set and offer explanations for any differences observed. While a principled analysis of course books is beyond the scope of the current paper, the characteristics of learners' native languages will be analyzed as one cause of differing patterns of use of modal verbs in L2 English. In particular, real and perceived similarities between modal verbs in Chinese and English will be shown to influence the use of *will* by CLE. The following section briefly compares and contrasts *will* and corresponding modal verbs in Chinese.

2.2 Modal Auxiliaries in Mandarin Chinese

Mandarin Chinese possesses a wide inventory of modal verbs, but this section will focus on the modal verb *hui* due to its similarities with *will*.

Tsai (2015) distinguishes 5 uses of *hui* as a modal verb. While the future and epistemic uses in (2) and (5) correspond to *will*, what Tsai terms “dispositional” (3) and “generic” (4) modals are not typically expressed using *will* (Tsai's idiomatic English translations of the Chinese sentences have been slightly adjusted). Dispositional and generic modals are referred to below as “non-future” uses. Table 1 summarizes the partial correspondence between *hui* and *will*.

- (1) Yiqian waijiaoguan dou hui shuo fayu. [dynamic modal]
Before diplomat all can speak French
'In the past, all diplomats could speak French.'
- (2) Waijiaoguan hui changchang lai zheli. [future modal]
Diplomat will often come here
'Diplomats will come here often.'
- (3) Waijiaoguan changchang hui lai zheli. [dispositional modal]
diplomat often tend.to come here

'Diplomats often tend to come here.'

- (4) Shui hui wang dichu liu. [generic modal]

water HUI towards low.land flow

'Water flows to lower places.'

- (5) Waijiaoguan dagai hui lai zheli. [epistemic modal]

diplomat probably Irr come here

'Diplomats will probably come here.'

(Tsai, 2015, p. 278)

Table 1. Correspondence between *hui* and *will*

Uses of <i>hui</i> (Tsai 2015)	Correspondence with <i>will</i>
1. Ability	no
2. Future	yes
3. Dispositional	limited
4. Generic	limited
5. Epistemic	yes

Examples of uses of *will* resembling the generic and dispositional uses of *hui* are shown in (6) and (7) below. These are examples of corrected learner production displayed on the English Grammar Profile Online and are described there as “habitual and typical” (6) and “willfulness or disapproval” (7) uses of *will*. Such uses are deemed to be limited for the following reasons. First, they are categorized at CEFR C1 and C2 levels respectively, so are typically acquired only at the highest proficiency levels. This is likely related to their low frequency of use by native speakers. Given the high CEFR ratings, it is unlikely that learners at the proficiencies focused on in this study will have received sufficient input to use them in their own writing.

- (6) “habitual and typical” (C1)

Can use 'will' to talk about something which is typical or habitual.

Example: She will often knock on the door to see you.

(Japan; C1 EFFECTIVE OPERATIONAL PROFICIENCY; 1993; Japanese; Fail)

(7) “willfulness or disapproval” (C2)

Can use 'will' to talk about general behaviour, often disapprovingly.

Example: Indeed no one can imagine what children will do!

(France; C2 MASTERY; 1993; French; Pass) (English Grammar Profile Online)

Secondly, as alluded to in the description in (7), *will* used to talk about general or typical states of affairs often expresses an air of disapproval which is absent from dispositional and generic uses of *hùi*. Carlson (2012, p. 834) discusses a further restriction, namely that habitual *will* cannot appear with individual-level states. Even where a habitual reading is plausible, as in (8c), this reading is excluded in favor of an epistemic reading in which the speaker is making a prediction about future conditions.

(8) a. Bob will be an attorney.

b. The girl will like ice cream.

c. The weather will be very mild here.

(Carlson, 2012, p. 834)

2.3 Functional Similarities Between L1 and L2 and the Potential for Crosslinguistic Influence

Section 2.2 demonstrated that *will* and *hùi* have limited functional similarities, namely their future and epistemic uses. In contrast, dispositional and generic uses of *will* are infrequent, marked and unlikely to be encountered in input learners receive. Nonetheless, there is potential for crosslinguistic influence in all five uses shown in Table 1 above. According to Jarvis & Pavlenko (2008, pp. 178–180), crosslinguistic influence typically occurs where there are subjective similarities between L1 and L2. A pertinent example is reported in Odlin (2008, pp. 317–318), who refers to a study by Sastry-Kuppa (1995). This study showed that native speakers of Tamil used *will* as a marker of habitual aspect, not only in the present tense but also in the past tense, where *would* or *used to* would be appropriate. Sastry-Kuppa concludes that this reflects overgeneralization of the similarities between *will* and the future tense marker in Tamil.

If CLE overgeneralize the similarities in Table 1 and assume functional equivalence in categories 3 and 4, they are expected to overuse *will* to express dispositional and generic meaning. This in turn may lead to overuse of *will* overall. This indeed appears to be the case, as the corpus studies in section 2.1 have revealed. What

the present study seeks to demonstrate is that CLE are indeed using *will* to express dispositional and generic meaning where native speakers do not (or do so at a significantly lower frequency).

In contrast to CLE, previous studies found that JLE underuse *will* in comparison to native speakers. This can be explained by considering the means of expressing modality in Japanese. Modal verbs such as *can*, *should* and *must*, which JLE were found to overuse, are typically expressed in Japanese using sentence final expressions or verbal inflections (9). In contrast, future and epistemic meanings are not expressed by a dedicated, obligatory morpheme, although *-daro* or *-ka mo shirenai*, expressing a subjective judgement of probability, are optionally attached to the non-past form of the verb (10). Furthermore, dispositional and generic meaning can also be expressed using an unmarked verbal form. This means that Japanese lacks formal equivalents to *will* and therefore L1 forms are not predicted to aid the acquisition of L2 forms. This lack of morphological salience in L1 is expected to manifest itself in underuse of *will* compared to both native speakers and CLE.

- (9) Gakusei wa apuri de benkyō {suru koto ga dekiru / suru beki da / shinakereba naranai}.ⁱ

Student TOP app INS study {do NMLZ NOM can / do ought.to COP / do-NEG-COND become-NEG}.

‘The students {can / should / have to} study by accessing the online resources.’

- (10) Gakusei wa apuri de benkyō suru (darō / ka mo shirenai).

Student TOP app INS study do (COP-CONJEC / Q also know-POTEN-NEG)

‘The students (will probably/might) study by accessing the online resources.’

Newbery-Payton and Mochizuki (2020) analyzed L1 to English translations by CLE and JLE in order to explore the effect that the absence or presence of comparable L1 forms has on the production of L2 forms. Errors of omission of *will* appeared exclusively in JLE data, while translations by CLE were characterized by inappropriate use of *will* in habitual senses. Newbery-Payton and Mochizuki explained these contrastive error trends through reference to the kinds of characteristics of Chinese and Japanese discussed above.

3. Research Design

The current paper seeks to verify the findings of Newbery-Payton & Mochizuki (2020) using different methodology. Specifically, it adopts a larger data set, examines a different task format, and compares both native speakers and learners at different proficiency levels using statistical testing.

3.1 Aim and Research Questions

This paper considers the research questions listed below. Research Questions 1 and 2 concern the overall frequency of use of *will*. CLE and JLE are expected to differ in their use of *will*. CLE are expected to exhibit a higher frequency of use than NS, while JLE are expected to exhibit a lower frequency of use. In addition, both groups of learners are expected to become more native-like in terms of frequency of use at higher proficiency levels.

RQ1: To what extent do CLE and JLE differ in their use of the modal auxiliary verb *will*?

RQ2: To what extent does the use of *will* by CLE and JLE become more native-like with increasing proficiency?

Research Question 3 concerns the effect of L1 forms on the use of *will* in L2 English. CLE are expected to overuse *will* in dispositional and generic senses, as a result of overgeneralization from L1. A similar phenomenon is not expected in the JLE data due to the lack of functional equivalents in L1.

RQ3: To what extent can the use of *will* by CLE be explained by reference to L1 forms?

3.2 Data and Method

Data is sourced from the Written Essays module of the International Corpus Network of Asian Learners of English (ICNALE; Ishikawa, 2013). Use of this data set provides the following advantages. First, two essay topics are specified for participants to write about, allowing both topic control and comparison of topics. The prompts for

each essay are shown in (11) below. In the remainder of this paper, topics A and B are abbreviated as “PTJ” (part-time job) and “SMK” (smoking) respectively.

(11) Do you agree or disagree with the following statements? Use reasons and specific details to support your opinion.

(Topic A) It is important for college students to have a part-time job.

(Topic B) Smoking should be completely banned at all the restaurants in the country. (Ishikawa, 2013, p. 97)

Although neither topic uses *will*, the prompt for the SMK task includes the modal verb *should*, which potentially affects the use of other modal verbs. Nevertheless, use of this data set is still preferable to the translation task used by Newbery-Payton & Mochizuki (2020), as the latter task type may enhance the potential for L1-related effects to occur. This is because learners may be directly influenced by features of the L1 text they are required to translate. Further discussion of task-related effects is provided in section 5.

The second advantage of using ICNALE is that it includes data from learners judged to be at A2, B1 and B2 CEFR levels. This allows pseudo-longitudinal analysis in order to examine the effect of proficiency. As alluded to in the previous section, crosslinguistic influence is predicted to be mediated by increasing proficiency.

48 essays on each topic were randomly selected from the data sets for JLE and CLE at A2, B1-1 and B1-2 levels. This reflects the size of the smallest of the subcorpora under consideration (JLE B1-2, N=49). B2 level learners were excluded from the analysis due to data size limitations. ICNALE contains three groups of NS; the student group was selected for analysis as this was considered to best match social characteristics of the learner groups. A total of 672 files totaling 154,088 words were selected for analysis. Summaries of the data size and learner attributes are provided in Tables 2 and 3.

Table 2. Data Summary

	PTJ					SMK			
	A2	B1-1	B1-2	Total		A2	B1-1	B1-2	Total
CLE	10923	11972	12287	35182	CLE	10713	11203	11438	33354
JLE	10933	10540	11057	32530	JLE	10423	10349	10850	31622
Total	21856	22512	23344	67712	Total	21136	21552	22288	64976
NS		10774		78486	NS		10626		75602

Total Files: 672 / Total Words: 154,088

Table 3. Summary of Learner Attributes

	CLE				JLE		
	A2	B1-1	B1-2		A2	B1-1	B1-2
F	25	24	21	F	22	17	21
M	23	24	27	M	26	31	27
Average age	19,2	19,5	19	Average age	18,5	18,6	18,8

Data was tagged using TagAnt and relevant examples were then extracted using AntConc. Each instance of *will* was examined within the wider context of the essay to determine the most likely intended meaning. In particular, “non-future” uses of *will* were identified and extracted for further analysis (see section 4.2).

4. Results

Quantitative analysis, relating to RQ1 and RQ2, is presented in section 4.1. Qualitative analysis, relating to RQ3, is presented in section 4.2.

4.1 Quantitative Analysis

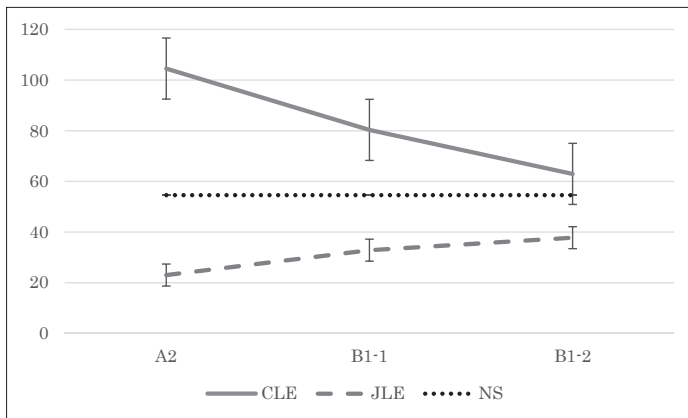
Table 4 shows the adjusted frequency of use of *will* by each group of learners in the two tasks. The data from the SMK and PTJ tasks are also shown in Figures 1 and 2 respectively. Black dotted lines in the figures show the performance of NS on each task.

Table 4. Adjusted Frequency (per 10,000 words) of *will*

	Part Time Job				Smoking		
	A2	B1-1	B1-2		A2	B1-1	B1-2
CLE	67.75	78.5	74.9	CLE	104.5	80.3	62.9
JLE	47.6	58.8	44.3	JLE	23.0	32.9	37.8
NS		72.4		NS		54.6	

On both tasks and at all proficiency levels, adjusted frequency is highest for CLE and lowest for JLE, with frequencies for NS appearing between the two groups of learners (the exception is the PTJ task, where adjusted frequency is slightly higher for NS than for CLE at A2 level). However, Figures 1 and 2 reveal different trends beyond these general similarities.

On the SMK task, frequency is particularly high for A2 CLE and particularly low for A2 JLE, resulting in a high degree of disparity between the two groups at A2 level. With increasing proficiency, however, CLE frequency of use falls and JLE frequency of use rises. In this way, proficiency effects are visible, with both groups of learners approaching native-like frequencies of use at higher proficiency levels.

Figure 1. Adjusted Frequency (per 10,000 words) of *will* in "Smoking" Task

The PTJ task displays less variation, both between groups and over different proficiency levels. In both groups of learners, there are marginal increases in frequency at B1-1 level, followed by marginal decreases in frequency at B1-2 level. Furthermore, A2 level learners' frequency of use is already relatively close to that of NS. As a result, there are no obvious proficiency effects comparable to those in the SMK task.

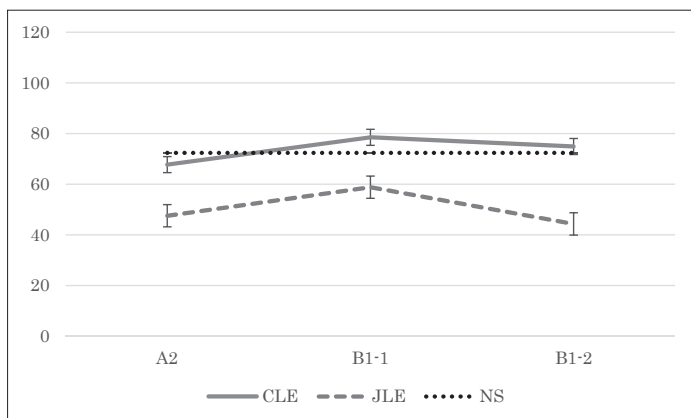


Figure 2. Adjusted Frequency (per 10,000 words) of *will* in “Part-Time Job” Task

The data for each task was next tested for statistical significance. A Kruskal-Wallis test ($df=6$, $\chi^2=57.334230$, $p=1.563306e-10$) revealed that the groups were not homogenous on the SMK task. Results of post-hoc tests (Dunn method, adjusted with Holm FWER for multiple comparisons) are reported in Table 5. A Kruskal-Wallis test on the PTJ data found no significant difference between groups.

Table 5. Dunn Adjusted p-values for Pairwise Comparisons (SMK)

	CLE_A2	CLE_B1-1	CLE_B1-2	JLE_A2	JLE_B1-1	JLE_B1-2
CLE_B1-1	.289					
CLE_B1-2	.035	1				
JLE_A2	<.001	<.001	.014			
JLE_B1-1	<.001	.012	.121	1		
JLE_B1-2	<.001	.052	.364	1	1	
NS	.019	1	1	.027	.207	.537

The significant differences in Table 5 (shown in bold) can be summarized as follows. CLE at A2 level use *will* significantly more frequently than almost all other groups, the only exception being B1-1 level CLE. CLE at B1-1 level also show significantly higher frequency of use than A2 and B1-1 level JLE. No significant differences were found between learners at B1-2 level. Taken together, the results corroborate the trends observed in Figure 1.

Comparing learners to NS, while A2 level learners show significantly higher (CLE) or lower (JLE) frequency than NS, at B1-1 and B1-2 these differences are no longer significant. In other words, non-nativelike frequency of use, both overuse and underuse, is limited to A2 level. This provides answers to RQ1 and RQ2: CLE and JLE differ significantly at A2 and partially at B1-1 level; from B1-1 level onwards, learners' frequency of use converges and becomes more native-like.

These results are not simply a reflection of idiosyncratic modal auxiliary use by a minority of learners. As summarized in Table 6, the percentages of learners in each group using *will* on at least one occasion in their writing show largely similar trends to the adjusted frequencies shown in Table 4. In other words, the high but falling frequency of use by CLE and the low but rising frequency of use by JLE appear to be characteristics of each group as a whole.

Table 6. Percentage of Learners Using *will* in the SMK Task

	A2	B1-1	B1-2
CLE	83	79	65
JLE	33	40	44

The next section considers RQ3, namely whether the significantly higher frequency of use of *will* produced by CLE can be explained, at least in part, by crosslinguistic influence in the form of overgeneralization of L1 forms.

4.2 Qualitative Analysis

RQ3 concerns the extent to which the “non-future” uses of the Chinese modal auxiliary *hui* might influence CLE use of *will*. In order to answer this question, uses of *will* were categorized and the proportion of “non-future” uses calculated. Analysis was

conducted by the author using the following heuristics (judgement by multiple annotators was not feasible due to practical constraints). Instances of *will* were considered “non-future” if they expressed states of affairs current to the reference time and could be replaced with present tense forms with minimal change of meaning. This amounts to a division between the future and epistemic uses of *will* on the one hand, and generic and dispositional uses on the other. This level of granularity was judged to be sufficient for the purposes of the current research question as it reflects the key parallels between English and Chinese. Conditional sentences containing *if* such as (12) allow an epistemic interpretation and so were not considered “non-future” uses. Sentences including *when* were judged on the content of the sentence and the wider context of the essay. For instance, (13) was regarded as expressing a future state of affairs whereas (14) was regarded as expressing a generic state of affairs (note the use of *sometimes*); only the latter was considered a “non-future” use.

- (12) If the law of banning is through, the atmosphere of restaurants **will** be more perfect. CHN_SMK_039_A2
- (13) When smokers cut down the number of cigarettes, the good dining atmosphere **will** be built easily. CHN_SMK_269_B1_1
- (14) They said sometimes inspire **will** come across in their mind when they smoked. CHN_SMK_310_A2

Contrary to expectations, A2 level JLE also showed some non-future uses of *will* (15). However, as proficiency rises, non-future uses of *will* largely disappear from the JLE data, while continuing to account for 15-20% of the overall use of *will* by CLE (Table 7). Examples of “non-future” uses of *will* by B1-level CLE are given in (17) and (18). The persistence of such examples suggests that CLE continue to use *will* in a manner analogous to L1, providing partial confirmation of the prediction for RQ 3.

- (15) Especially, in the restaurant, many people **will** enjoy eating and talking with friends or families. JPN_SMK_344_A2

- (16) There are many people who like smoking, even in the public places they **will** take a cigarette in hand. CHN_SMK_289_A2
- (17) Finally, it is a good idea to ban smoking in any restaurants because When someone smoke cigarettes, harmful gases **will** arise and fulfill the room. CHN_SMK_104_B1-1
- (18) France has forbidden people smoking in cafes a few years ago. The citizens who defy it **will** be punished and feed for a lot. CHN_SMK_363_B1-2

Table 7. Frequency and proportion of non-future uses of *will*

	A2	A2 (%)	B1-1	B1-1 (%)	B1-2	B1-2 (%)
CLE	20	18	18	20	11	15
JLE	5	21	3	9	1	2

It must be recognized that “non-future” use of *will* alone cannot explain the significant differences in frequency at A2 and B1-1 levels. The phenomenon might best be understood as one expression of L1 transfer occurring more generally in the writing of CLE. The remainder of this section will consider one further aspect of the low frequency of use by JLE in the SMK task, namely errors of omission.

JLE are expected to omit *will* in obligatory contexts more frequently than CLE due to the absence of functional equivalents to *will* in Japanese. One such situation is in conditional clauses, as verbs in consequent clauses are frequently marked with *hurai* in Chinese but receive no dedicated morphological marking in Japanese. Conditional clauses were extracted from the data set by searching for sentences including *if* then filtering manually. Examples are provided below, with relevant errors in bold. As predicted, JLE show a greater raw frequency and proportion of errors of omission in obligatory contexts (Table 8). Extraction of all errors of omission was beyond the scope of the current paper, but it seems reasonable to expect errors of omission to occur more frequently throughout JLE’s writing, not only in conditional clauses.

- (19) However, if smoking is banned at all the restaurants, the smoker **is** uncomfortable

and dissatisfied.

W_JPN_SMK0_312_A2

(20) If smoking is forbidden at all restaurants, I think only nonsmokers and light smokers **enjoy** meals.

W_JPN_SMK0_269_B1_1

(21) If smokers stop smoking at the public places in order not to be fined, the public places **become** more comfortable and cleaner.

W_JPN_SMK0_179_B1_2

Table 8. Omission of *will* in Obligatory Contexts (Conditional Clauses)

	A2	A2 (%)	B1-1	B1-1 (%)	B1-2	B1-2 (%)
CLE	3	4	3	6	2	5
JLE	10	14	12	17	13	17

In summary, it appears that crosslinguistic influence may have affected the overall frequency of use by the two groups of learners, as well as influencing the proportion of uses of *will* expressing “non-future” meaning and the proportion of errors of omission in obligatory contexts. The implications of these findings are discussed in the following sections.

5. Discussion

This section considers a number of issues raised by the present study and their implications for future research. Section 5.1 reviews the study’s main findings and their relation to theoretical distinctions in the field of second language acquisition. Section 5.2 considers task- and proficiency-related effects in relation to previous studies, while Section 5.3 considers task-related and other effects within the current data set.

5.1 Crosslinguistic Influence, Suppletion and Addition in Second Language Acquisition

While L1-related effects on acquisition (measured in terms of accuracy of use) have been demonstrated for a range of grammatical categories (Luk & Shirai, 2009; Murakami & Alexopoulou, 2016), few such effects have been demonstrated for *will* or

for modal verbs more generally. Recent explanations offered for the overuse or underuse of individual modal verbs (see section 2.1) rest on categorizations of modal verbs as a group, without considering L1-related factors. This study's findings suggest that specific consideration of relevant L1-related factors can provide nuance that such approaches miss.

Murakami & Alexopoulou (2016, p. 368) hypothesize that “lack of the equivalent feature in the L1 leads to low accuracy”. The present study confirms this for JLE, while also showing that the presence of an (apparently) equivalent feature leads to lower accuracy, due to overuse of the target form. Gabriele (2009) argues that addition (the acquisition of new interpretations of a given linguistic form) should be distinguished from preemption (the ruling out of interpretations present in L1 but not in L2). Gabriele examines the acquisition of different interpretations of progressive forms in English and Japanese and concludes that preemption is more difficult than addition, especially in the absence of explicit input showing otherwise.

In the context of the present study, CLE must preempt the “non-future” uses of *will*, while JLE must acquire the semantics of *will* due to a lack of a functional equivalent in Japanese. If acquisition is incomplete, CLE are expected to overuse non-future senses of *will* and JLE are expected to underuse and/or omit *will* in obligatory contexts. While this is what the results have indicated, the current study does not offer evidence either for or against the assertion that preemption is more problematic than addition.

5.2 Task- and Proficiency-Related Effects in Relation to Previous Studies

Newbery-Payton & Mochizuki (2020) analyzed L1 to English translations by high proficiency learners. The present study, however, found that CLE and JLE had converged by B1-2 level, suggesting that task type influences the extent to which L1-influence occurs. Translation tasks, which provide an L1 text to translate into L2, may induce even higher proficiency learners to emulate certain features of L1 in their L2 writing, whereas free-writing tasks appear to show L1-related effects only at lower proficiency levels. This underlines the importance of confirming research findings using different data sets, task types and groups of learners. In the present study, there was no L1 source text that might induce L1-like norms in L2 writing, such as the inclusion of *will* wherever *hui* appeared in the source text. Furthermore, the essay

prompts did not contain *will*, so direct linguistic influence from the prompts cannot be assumed.

In addition, the studies referred to in Section 2 examining L2 English modal use cannot be said to have fully considered proficiency effects. It is possible that the modal verbs reported to be underused or overused were in fact used at native-like frequencies by higher proficiency learners. Clarification of such issues could be beneficial when considering, for example, which aspects of modal verb pedagogy would most benefit from reconsideration and at which stages of EFL study.

5.3 Task-Related and Other Effects in the Present Study

As stated in Section 4.1, differences between the three groups in the PTJ task were not statistically significant. This may reflect lesser (perceived) “opportunity of use”, a term used to refer to “the opportunity the learner is afforded to use a linguistic feature”, which can be affected by factors including task type, task topic and document length (Buttery & Caines, 2018, p. 6). Of these three factors, task topic is the most relevant for the ICNALE data.

While neither essay topic discourages use of *will*, the SMK task is arguably more conducive to writing about hypothetical future events, as learners are encouraged to write about the implications of a possible future change in the law. This provides two contexts – sentences including future time reference and consequent clauses in conditional sentences – where functional similarities between *hui* and *will* encourage the use of the latter by CLE. However, while CLE at A2 and B1-1 level did indeed use *will* more frequently on the SMK task, the opposite is true for the B1-2 group. NS and JLE at all proficiencies similarly displayed a higher adjusted frequency on the PTJ task than they did on the SMK task (Table 4).

It is difficult to provide conclusive answers to this puzzling phenomenon, but one explanation may lie in the use of other modal verbs, which are in syntactic competition to appear before the main verb in a given sentence. As stated in section 2, Nakayama (2020) reported overuse of *can*, *should* and *must* by JLE in the ICNALE data. This is likely related to the fact that *should* appears in the essay prompt for the SMK task (11). JLE may have selected *should* more frequently in the PTJ task, leaving fewer opportunities for the use of *will*. If the conclusions of this paper are valid, then CLE are already primed to use *will* due to L1-related factors, causing the prominent differences

between the two groups on the SMK task. NS may be less likely to be influenced by the prompt, given the wider range of linguistic devices available to them.

The PTJ essay prompt does not include any modal verbs so learners are not explicitly induced to select one modal over another and the topic may be more conducive to a mix of temporal references and real and hypothetical situations. This may be why the PTJ task exhibited more homogeneous use of *will* by the three groups.

A reviewer suggests that cultural differences may influence the trends reported in this study. While it is possible that one group of learners is more likely to hedge statements using other modal verbs, thus avoiding *will*, in the view of the author this is more likely to occur with auxiliaries used primarily as deontic modals, which more directly reflect the writer's stance. Chen and Zhang (2017, pp.19–21), in their study of hedging by Chinese and Anglophone writers, report that the only modal verb with significant differences in frequency was *should*; Chinese writers were found to overuse *should* as a deontic modal but underuse it as an epistemic modal. In regard to writing by Japanese native speakers, Takimoto (2015, pp. 95–96) states that although Japanese speakers may express themselves indirectly in their native language, such L1 norms are not necessarily replicated in L2 English writing. Takimoto reports that JLE use boosters (including *will*) as frequently as NS, and hedges significantly less frequently than NS.

High frequency hedges in Takimoto's study include the modal verbs *could* and *may*. A comprehensive study of learners' selection of modal verbs is beyond the scope of the present paper, but the occurrence of these two modal verbs in the current data set can be summarized as follows. First, their frequency generally rises with increasing proficiency for CLE. This might be expected on the SMK task, where the frequency of *will* falls significantly for CLE at B1-2 level (Figure 1). It cannot, however, account for the PTJ task, where the frequency of *will* displays minimal change despite the increase in frequency of these hedging modals.

Second, the frequency of hedging modals is typically lower for JLE than it is for CLE. Furthermore, as learners' proficiency rises, frequency of use either decreases or returns to its original level after an initial rise. The exception is *could* on the PTJ task, where JLE also exhibit a large increase, exceeding the frequency for NS. The general tendency may be another instantiation of underuse of modal verbs by JLE due to the absence of equivalent obligatory morphemes in L1 (see Section 2.3).

It should be noted that the combined frequency of *could* and *may* across tasks and

groups (N=459) is less than half of that of *will* (N=932). In short, it seems unlikely that trends in the overuse and underuse of *will* can be reduced to an epiphenomenon caused by selection trends among other modal auxiliaries.

Use of other modal verbs does not appear to explain the differing frequency of use of *will* by NS either. For instance, adjusted frequencies of *could* and *may* are comparable on both tasks, and the adjusted frequency of *can* is higher on the PTJ task (100) than on the SMK task (59). The PTJ task therefore appears to be generally more conducive to the use of modal verbs – at least for NS. The author hopes to address this issue more fully in future studies.

Finally, another reviewer asks whether the frequency of *will* is related to the frequency of *going to*, particularly at lower proficiency levels. While the highest frequency of use was indeed observed in the data for the PTJ task by A2 level JLE, *going to* appeared only 13 times in the whole data set. Given this low frequency, preference for one future expression over another appears to have a relatively small effect on trends of use, at least for the current topics and task types.

6. Conclusion

This study examined the extent to which CLE and JLE differ in their use of *will* (RQ1), the effect of proficiency on frequency of use (RQ2), and the extent to which overuse or underuse of *will* can be explained with reference to L1 forms (RQ3). Analysis revealed significant differences in the use of *will* by CLE and JLE at lower proficiency levels, whereas learners at higher proficiency levels did not differ significantly from each other or from native speakers. Qualitative analysis showed that non-future uses of *will* were significantly higher among CLE, suggesting learners use the form in an analogous manner to the Chinese modal auxiliary *hui*. JLE do not display this characteristic and also show a tendency to omit *will* in obligatory contexts, suggesting that the absence of a comparable L1 form is one factor in the underuse of *will*.

The current study was limited to written language, so spoken data from ICNALE could also be analyzed in future. Online processing demands during speech may cause higher rates of omission of target forms, particularly at lower proficiency levels. It is unclear, however, whether or not this will significantly affect the differences between

JLE and CLE in terms of frequency of use of *will* or other modal verbs.

While this study has focused on one particular linguistic form, similar methodology could be used to investigate the frequency of other forms. Principled selection of these forms, and of the L1 groups to include in analyses, can be aided by careful consideration of L1 characteristics.

Finally, more attention has been paid in recent years to the interface between the fields of corpus linguistics and second language acquisition (Le Bruyn & Paquot, Eds., 2021). It is hoped that corpus analyses like the present study can complement existing SLA research or provide the impetus for new studies. For example, researchers could examine whether overuse and underuse of modal verbs by different groups of learners are also observable in cloze tasks or other experimental designs.

References

- Anthony, L. (2015). TagAnt (Version 1.2.0) [Computer Software]. Tokyo, Japan: Waseda University. Available from <https://www.laurenceanthony.net/software>
- Anthony, L. (2020). AntConc (Version 3.5.9) [Computer Software]. Tokyo, Japan: Waseda University. Available from <https://www.laurenceanthony.net/software>
- Binnick, R. (2005). The markers of habitual aspect in English. *Journal of English Linguistics*, 33(4), 339–369.
- Caines, A., & Buttery, P. (2018). The effect of task and topic on opportunity of use in learner corpora. In Brezina, V., & Flowerdew, L. (Eds.). *Learner corpus research: New perspectives and applications* (pp. 5–27). Bloomsbury.
- Carlson, G. (2012). Habitual and generic aspect. In Binnick, R. (Ed.). *The Oxford handbook of tense and aspect* (pp. 828–851). Oxford University Press.
- Chen, C., & Zhang, L. (2017). An intercultural analysis of the use of hedging by Chinese and Anglophone academic English writers. *Applied Linguistics Review*, 8(1), 1–34.
- Gabriele, A. (2009). Transfer and transition in the SLA of aspect A bidirectional study of learners of English and Japanese. *Studies in Second Language Acquisition*, 31, 371–402.
- Ishikawa, S. (2013). The ICNALE and sophisticated contrastive interlanguage analysis of Asian learners of English. School of Languages & Communication, Kobe University. *Learner Corpus Studies in Asia and the World*, 1, 91–118.
- Jarvis, S., & Pavlenko, A. (2008). *Crosslinguistic influence in language and cognition*. Routledge.
- Luk, Z., & Shirai, Y. (2009). Is the acquisition order of grammatical morphemes impervious to L1 knowledge? Evidence from the acquisition of plural -s, articles, and possessive 's. *Language Learning*, 59, 721–754.
- Le Bruyn, B., & Paquot, M. (2021). *Learner corpus research meets second language*

- acquisition*. Cambridge University Press.
- Murakami, A., & Alexopoulou, T. (2016). L1 influence on the acquisition order of English grammatical morphemes: A learner corpus study. *Studies in Second Language Acquisition*, 38(3), 365–401.
- Nakayama, S. (2020). Contrastive interlanguage analysis of modal auxiliary verb usage by Japanese learners of English in argumentative essays. *The IAFOR International Conference on Education – Hawaii 2020 Official Conference Proceedings*.
- Nakayama, S. (2021). Modal auxiliary verbs in Japanese EFL learners’ conversation: A corpus-based study. *Asia Pacific Journal of Corpus Research*, 2(1), 23–34.
- Newbery-Payton, L., & Mochizuki, K. (2020). L1 influence on use of tense/aspect by Chinese and Japanese learners of English. School of Languages & Communication, Kobe University. *Learner Corpus Studies in Asia and the World*, 4, 67–93.
- Odlin, T. (2008). Conceptual transfer and meaning extensions. In Robinson, P., & Ellis, N. (Eds.). *Handbook of Cognitive Linguistics and Second Language Acquisition* (pp.306-340).
- Sastry-Kuppa, S. (1995, March 3). *That’s why he will talking for English: The expression of habitual aspect in the English of untutored and low-level tutored Indian speakers*. Presentation at the Ninth International Conference on Pragmatics and Language Learning, University of Illinois.
- Takimoto, M. (2015). Assertions and lexical invisibility in EFL learners’ academic essays. *Journal of Pragmatics*, 89, 85–99.
- Tsai, W. (2015). On the topography of Chinese modals. In Shlonsky, U. (Ed.). *Beyond functional sequence* (pp.275–294). Oxford University Press.
- Xiao, Y. (2017). Chinese EFL learners’ acquisition of modal verbs: A corpus-based study. *International Journal of English Linguistics*, 7(6), 164–170.
- Yang, X. (2018). A corpus-based study of modal verbs in Chinese learners’ academic writing. *English Language Teaching*, 11(2), 122–130.

This research has made use of the English Grammar Profile. This resource is based on extensive research using the Cambridge Learner Corpus and is part of the English Profile programme, which aims to provide evidence about language use that helps to produce better language teaching materials. See <http://www.englishprofile.org> for more information.

¹Notation follows conventions in the Handbooks of Japanese Language and Linguistics published by the National Institute for Japanese Language and Linguistics.

COND	conditional	INS	instrumental	POTEN	potential
CONJEC	conjunctural	NEG	negative	Q	question marker
COP	copula	NMLZ	nominalizer	TOP	topic

