Disentangling cues of different domains in transfer and development in L3 acquisition: An investigation of L2/L3 Mandarin *yes-no* questions

Yanyu Guo[1] and Boping Yuan[2]

[1]University College Cork, Cork, Ireland and [2]Shanghai Jiao Tong University, Shanghai, China / University of Cambridge, Cambridge, UK

![CAMBRIDGE UNIVERSITY PRESS]

Abstract

This empirical study aims to shed light on L3 initial-stage transfer and later development by investigating Q-operations in L1 English–L2 Cantonese and L1 Cantonese–L2 English bilinguals' L3 Mandarin and L1 English speakers' L2 Mandarin at low and high proficiency levels. Data from an online cross-modal priming task and an offline acceptability judgement task found that structural similarity determines transfer source selection. Adopting a de-compositional approach to cues of different domains, we have found both facilitative and detrimental transfer effects from Cantonese, with the latter triggered by orthographic and phonological cues. Our data also suggest that detrimental transfer effects can persist at an advanced stage and that L3 development and acquisition results can be affected by various factors such as word frequency and the nature of learning situations.

1. Introduction

The past decade has witnessed increasing attention to multilingualism in linguistic research, and transfer/cross-linguistic influence (CLI) in third language (L3) acquisition has been a hot topic of debate in the literature. Various models have been proposed to identify sources of transfer in L3 acquisition, i.e., which of the previously acquired languages is most likely to influence the L3 acquisition and whether the influence is wholesale or piecemeal (cf. Rothman et al., 2019; Schwartz & Sprouse, 2021; Slabakova, 2017; Westergaard et al., 2017). In this article, we report on an empirical study examining L3 acquisition of Q-operations in Mandarin *yes-no* questions by Cantonese–English bilinguals and English monolinguals. We will focus on the role of different types of cues (e.g., syntactic, semantic, phonological, orthographic cues) in triggering transfer at initial stages and key factors that influence L3 grammar.

2. Q-operations in *yes-no* questions

2.1 Formation of *yes-no* questions

Languages vary in the way of forming *yes-no* questions to request the addressee to indicate whether a given proposition is true. English employs subject-auxiliary inversion or *do*-support to form *yes-no* questions (Holmberg, 2015), whereas Mandarin and Cantonese use two phonetically realised forms to instantiate a [+Q] feature: sentence-final particles (SFPs) and an A-not-A structure (Law, 2002; Paul, 2015; Sybesma & Li, 2007).

(1) Ni   hui   shuo   yingyu   ma/a_1?   (Mandarin)
    you   know   speak   English   SFP
    the *ma* question: "Do you speak English?"
    the a_1 question: "You can speak English, don't you?" or
                        "(What?!) You can speak English?"

(2) Nei5   sik1   gong2   jing1-man2   maa3/aa4?   (Cantonese)
    you   know   speak   English   SFP
    the *maa3* question: "Do you speak English?"
    the *aa4* question: "You can speak English, don't you?" or
                        "(What?!) You can speak English?"

For example, in (1) and (2), the Mandarin SFPs *ma* and a_1 and the Cantonese SFPs *maa3* and *aa4* are of a [+Q] feature and can turn a sentence into a *yes-no* question. The difference

between *ma/maa3* questions and *a₁/aa4* questions is that the latter have an additional force of tone-softening or surprise (see the English translations of (3) and (4)).

To form a *yes-no* question, Mandarin and Cantonese can also use morphological reduplication, the so-called "A-not-A" structure, which involves reduplicating part or the entirety of the verb, adjective or proposition (represented by "A"), with a negator (*bu/mei* "not") in between, as illustrated in (3) and (4).

(3) Ni   hui-bu-hui   shuo   yingyu
    you   know-not-know   speak   English
    (a_2 [-Q]/*ma/*a_1)?   (Mandarin)
    SFP
    "Do you speak English?"

(4) Nei5   sik1-mh-sik1 gong2   jing1-man2
    you   know-not-know speak   English
    (*aa3*[-Q]/*maa3/*aa4)?   (Cantonese)
    SFP
    "Do you speak English?"

Moreover, in the three languages, only one [+Q] device can be employed to form a *yes-no* question due to economy (cf. Huang et al., 2009), and therefore, [+Q] SFPs such as the Mandarin *ma* and *a₁* and the Cantonese *maa3* and *aa4* are not compatible with A-not-A that is also of [+Q], as illustrated in (3) and (4). A-not-A questions can only be optionally followed by [-Q] SFPs, which are to express the speaker's attitudes and emotions. For instance, the Mandarin SFP *a₂* in (3) and the Cantonese SFP *aa3* in (4) do not violate the single Q-marking rule, and they act as a tone-softener, which makes the question more polite and softer.

It can be seen from the analysis above, Cantonese is typologically and structurally closer to Mandarin than English is in terms of the syntax of *yes-no* questions.

2.2 Mandarin SFPs ma and a and their Cantonese counterparts

Both Mandarin and Cantonese are well-documented as SFP languages, but they differ in the number of SFPs. Mandarin has around ten SFPs, most of which have multiple functions (cf. Li, 2006; Paul, 2015; Zhu, 1982), while the number of Cantonese SFPs ranges from some 30 (Kwok, 1984) to 95 (Leung, 1992), depending on how one counts them. Due to the difference in the number of SFPs, it is common that multiple functions of one Mandarin SFP are respectively expressed by several Cantonese SFPs. Moreover, for the written form of SFPs, Mandarin uses simplified Chinese characters but Cantonese adopts traditional ones, with substantial similarities shared by the two versions. Regarding the sound, Mandarin SFPs are normally pronounced in the neutral tone, but Cantonese employs tones to differentiate functions of the SFPs which share the same character.

The Mandarin SFPs involved in the present study, *a* (啊) and *ma* (吗), differ in their functions and distributions. It is documented that *a* (啊) is the most widely distributed SFP in Mandarin, as it can appear in exclamatives, imperatives, and (*wh-/ yes-no*) questions (Li & Thompson, 1989; Zhu, 1982). In (1), *a* is a [+Q] SFP, which we label as *a₁*, and it has a strong interrogative force to turn a sentence into an information-checking question with a soft tone or a tone of surprise (Paul, 2015). Moreover, *a* can also be a [-Q] SFP, which we label as *a₂*, and it only acts as a tone-softener after the A-not-A structure, as in (3). Cantonese uses the same characters but different tones to express the two functions. The Cantonese SFP *aa4* (written as 呀/啊) is the equivalent of the Mandarin [+Q] *a₁*, as illustrated in (2). The Cantonese *aa3* (also written as 呀/啊) corresponds to the Mandarin [-Q] SFP *a₂*, which is mainly used in interrogatives such as A-not-A questions, as in (4), to soften the force of a question (Matthews & Yip, 1994; Sybesma & Li, 2007; Tang, 2015).

Compared to the functions of *a*, the function of the Mandarin SFP *ma* is much more straightforward. It is used exclusively as a question marker, mainly in a neutral context where the speaker does not hold any pre-assumption of whether the proposition expressed by the question is true or false (Li & Thompson, 1989). It is documented that there is no neutral *yes-no* question marker – namely, no equivalent SFP of *ma*, in Cantonese (Matthews & Yip, 1994; Sybesma & Li, 2007; Tang, 2015). The question SFP *maa3* in (2) is normally used in a formal non-negation context[1] and considered a result of influence from Mandarin (Tang, 2015). Consequently, it is not included in most Cantonese dictionaries as an SFP. *Maa3* is written as 嘛 or 嗎[2] in Cantonese, with the former character shared with another SFP, [-Q] *maa5*.

The forms and meanings of the Mandarin *ma* and *a* and their potentially approximate Cantonese and English counterparts are summarised in Table 1. In Mandarin, the SFP *ma* (吗) can only be [+Q], whereas the form of *a*(啊) can be either [+Q] or [-Q], represented as *a₁* and *a₂* respectively. In Cantonese, the SFP *maa3* is an approximation of the Mandarin SFP *ma*. [+Q] and [-Q] SFPs in Cantonese are phonologically differentiated by tones (e.g., *aa4* and *aa3*). English does not have particles to indicate questions or the speaker's attitude but employs "subject-auxiliary inversion" or "*do*-support" to construct general purpose *yes-no* questions (Holmberg, 2015).

3. Issues on L3 initial-stage transfer and later development

Generative approaches to L3 acquisition have mainly discussed modelling morphosyntactic transfer at initial stages of the L3 and later development. Moreover, discussion about the implication of different methodologies (i.e., on-line versus offline measures) in L3 research remains insufficient. In this section, we will briefly review some influential L3 transfer models, key factors

Table 1. Mandarin ma and a and their approximate counterparts in Cantonese and English

Type	Function	Mandarin	Cantonese	English
+Q	neural	*ma* (吗)	*maa3* (嗎/嘛)	subject-auxiliary inversion/ *do*-support
	surprise/tone-softening	*a₁* (啊)	*aa4* (呀/啊)	∅
−Q	tone-softening	*a₂* (啊)	*aa3* (呀/啊)	∅

a. ∅ stands for no counterparts.

that moderate L3 acquisition development, and the use of different testing methods in L3 studies, which is yet to be well researched in the field.

3.1 What is transferred and what triggers transfer at L3 initial stages?

Different from L2 learners (L2ers), L3 learners (L3ers) have two already highly-activated languages. Hence, when modelling L3 acquisition, a key question to address is how transfer takes place among multiple sources available (the L1 and L2) in L3 initial stages. Some early work found that the L1 plays a privileged role in L3 initial-stage transfer (Hermas, 2010; Jin, 2009; Na Ranong & Leung, 2009), which is dubbed as "the L1 factor hypothesis" in Slabakova (2017) even though it is never formalised as a model. Some other studies, however, persist that the L2 can take on a stronger role than the L1 in the initial stage of L3 syntax, which is labelled as the L2 Status Factor Hypothesis (L2SF) (Bardel & Falk, 2007, 2012; Falk & Bardel, 2011). This model is built on Paradis' (2009) neurolinguistic framework which differentiates between implicit linguistic competence sustained by procedural memory and explicit metalinguistic knowledge by declarative memory. The L2SF contends that the L2 serves as the transfer source for the L3 when the L2 has been learned in a similar manner as the L3, because formally learned L2 and L3 are a function of explicit metalinguistic knowledge (Bardel & Falk, 2012).

In contrast, some other researchers argue that neither the L1 nor the L2 has a privileged status for the morphosyntactic transfer at the initial stages of L3 acquisition. This view is shared by four influential models: the Cumulative Enhancement Model (CEM) (Flynn et al., 2004), the Linguistic Proximity Model (LPM) (Westergaard, 2021a, 2021b; Westergaard et al., 2017), the Scalpel Model (SM) (Slabakova, 2017), and the Typological Primacy Model (TPM) (Rothman, 2010, 2011, 2015; Rothman et al., 2019). The CEM proposes that transfer to L3 occurs only when at least one of the previous two grammars instantiates the target property to avoid redundancy and that transfer can only be non-detrimental. The other three models take typological/structural similarity as the determinant in transfer source selection but are divergent on whether transfer is piecemeal or wholesale and on what triggers transfer.

As a wholesale transfer model, the TPM extends the Full Transfer/Full Access model for L2-interlanguage grammars (Schwartz & Sprouse, 1996) to L3 acquisition and argues that the linguistic parser selects a structurally similar language as the transfer source from the previously acquired languages, based on some hierarchical linguistic cues (Lexicon → Phonology → Morphology → Syntax). Among these cues, lexis is regarded as a primary source and given the biggest weight to trigger full transfer, and lexical similarities (especially semantic similarities) is prioritised at initial stages (Rothman et al., 2019). Cues of phonological, morphological and syntactic levels come into play in making comparative choices later, only when the motivation for selection cannot come from the lexicon. The LPM (Westergaard et al., 2017) and the SM (Slabakova, 2017), however, are proponents of piecemeal transfer and argue that cross-linguistic influence proceeds in a property-by-property fashion throughout development, given comparative structural similarities. Assuming a "learning by parsing" paradigm, the LPM proposes that variables such as saliency and early availability of cues in the L3 input may lead to the variation of CLI across time and linguistic domains (Westergaard et al., 2017).

3.2 Non-facilitation and L3 development

Based on the investigation of L3 initial stages, we can further discuss how L3 develops and whether L3 and L2 acquisition trajectories are different.

The modelling of non-facilitation is directly related to developmental predictions; if nonfacilitative transfer occurs, it will eventually have to be overcome at later stages. The current L3 models differ in whether non-facilitation would occur and what causes non-facilitation. The Cumulative Enhancement Model is the only model that advocates that the L3 initial transfer is always non-detrimental. However, many L3 studies (e.g., García Mayo & Slabakova, 2015; Guo & Yuan, 2020) have found that non-facilitation exists and remnant transfer from initial stages can linger in high proficiency levels of the L3.

The other L3 models acknowledge that transfer can be non-facilitative but diverge on how non-facilitation arises. The TPM, as a wholesale transfer model, persists that the initial state of L3 acquisition is the entirety of the typologically closer language that has been previously acquired, and therefore, non-facilitation brought by mismatches between the two grammar systems comes naturally. Although this model is not for L3 later development, learning difficulties in L3 acquisition under this framework can be predicted by those L2 models which also hold a full transfer position. Similar to L2ers, L3ers also need to re-configure features of the target L3 item by adding or discarding some features when there are some mismatches between the feature set of the target item and that of the corresponding item in the transfer source language, which may cause long term difficulty (cf. the Feature Reassembly Hypothesis in Lardiere, 2009).

On the other hand, the L3 models that advocate piecemeal transfer intrinsically concern developmental stages as transfer can take place throughout the L3 acquisition course. The factor of input plays a key role in the activation of transfer. The LPM advocates "learning by parsing" and predicts that cross-linguistic influence can be non-facilitative when learners misanalyse L3 input and/or have not had sufficient L3 input (Westergaard, 2021a). The Scalpel Model (Slabakova, 2017) considers effects of factors on the L3 acquisition process such as construction frequency, misleading input, necessity of negative evidence, which have been observed in some L3 studies (e.g., Guo & Yuan, 2021; Jensen et al., 2023; Slabakova & García Mayo, 2015). Although it is logical to assume that these factors can influence L3 acquisition as in the case of L2 acquisition, questions such as how the disparate factors interact in transfer and later development remain unclear. More importantly, although the current L3 models predict inhibition from a previous grammar transferred, they have little discussion about the process of L3ers overcoming detrimental transfer.

3.3 Online vs. offline tasks in L3 research

Generative L2/L3 studies inherently concern how mental representations of non-native grammars come to be in the mind of L2ers/L3ers. Tasks employed are likely to relate to different types of grammatical knowledge that they tap into (cf. Ellis, 2005; Marinis, 2010), and mismatches found in L3 research outcomes may be attributable to factors involved in methodology (Rothman et al., 2019).

Offline tasks such as untimed acceptability judgement tasks are typically considered conducive to testing and measuring explicit knowledge, while online tasks such as cross-modal priming

tasks (CMPTs), which detect participants' automatic response to stimuli, are considered appropriate for measuring implicit knowledge (cf. Godfroid et al., 2015; Marinis, 2010, 2018). The majority of L3 studies in the literature reported data elicited from offline tasks and not much work on L3 morphosyntax has employed online processing methods (see reviews in De Bot & Jaensch, 2015; Rothman et al., 2019). A few recent processing studies employing different online methods have revealed mixed results about the transfer effect. For example, Abbas et al. (2021) employed an online reading task with eye movement recording and an offline judgement task to investigate L1 Arabic–L2 Hebrew speakers' L3 English grammar and found that both L1 and L2 can be the source of cross-language influences in L3 processing and that the trilingual language system is fully interactive. Another pioneer study is Alonso et al. (2020), which used event related potentials (ERPs) on the acquisition of two artificial languages (Mini-Spanish and Mini-English) and found evidence in support of the similarity-based L3 models such as the TPM. Given the limited number of L3 studies with online methodology, more studies with online methodology are obviously necessary to provide clearer empirical evidence about the nature of initial-stage transfer and later development in the L3.

As shown above, efforts have been made to tackle questions concerning L3 initial stages, but empirical work is also needed to investigate L3 development by involving diverse methodologies. Moreover, there is an obvious lack of varieties in language pairings as most current L3 studies take Indo-European languages as the target language (with exceptions of Guo & Yuan, 2020, 2021; Leung, 2005; Na Ranong & Leung, 2009). In view of all this, we report on an empirical study, aiming to shed light on not only the transfer issue, but also L3 development and relevant influential factors involved, by comparing the L3 and L2 acquisition of Mandarin Q-operations by English–Cantonese bilinguals and English monolinguals.

4. The study

To investigate how transfer takes place at L3 initial stages and how it influences later development, we asked three main questions below, followed by relevant predictions.

Q1: Which of the previously acquired languages is the source of transfer at L3 initial stages?

Predictions: For the construction of *yes-no* questions, Cantonese is structurally closer to Mandarin than English, as the [+Q] feature in Cantonese can be realised in either SFPs or A-not-A, which is similar to the target L3. Hence, Cantonese is predicted to be the transfer source by the TPM, the CEM, the LPM and the SM, no matter whether Cantonese is the L1 or L2 of L3ers. However, the L2SF would predict that English is the transfer source when English is the L2 of L3ers.

Q2: Can transfer be non-facilitative? If yes, what causes non-facilitation?

Predictions: There are some mismatches between the Cantonese and Mandarin question SFPs. Cantonese has equivalent SFPs of Mandarin $a_{1/2}$, which will cause facilitative effects. However, Cantonese only has an approximation of [+Q] Mandarin *ma*, which is *maa3*. The Cantonese orthographic forms of *maa3* are 嘛 [±Q] and 嗎[+Q]. The orthographic cue of 嘛, which is shared with another SFP ([-Q] *maa5*), may lead to L3ers' incorrect use of Mandarin *ma* in *yes-no* questions when Cantonese grammar is transferred into their L3 Mandarin. This will reject the CEM and support the predictions

of the LPM and the SM regarding the effect of misleading input on non-facilitative transfer.

Q3: Do L2ers and L3ers of Mandarin have the same developmental pattern and similar acquisition results? If not, why not?

Predictions: We predict that L2ers and L3ers will behave differently as they are involved in different learning situations: L2ers are learning a new way to instantiate [+Q] (i.e., questions SFPs) and L3ers face feature re-configuration. L3ers will outperform their L2 counterparts on questions with $a_{1/2}$ due to a facilitative transfer from Cantonese, but will be less successful in the acquisition of questions with *ma* – as a feature discarding process (from [±Q] to [+Q]) is required at later stages, as predicted by the Feature Reassembly Hypothesis.

4.1 Participants

A total of 174 people participated in this study, including 28 Mandarin native speakers living in China as a control group. There were three types of learners: L1 English–L2 Cantonese–L3 Mandarin (E-C-M) learners; L1 Cantonese–L2 English–L3 Mandarin (C-E-M) learners; and L1 English–L2 Mandarin (E-M) learners. To a large extent, they were instructed learners of their non-native languages, particularly at early stages. Learners of each type were divided into two proficiency groups, a low proficiency group (beginners and pre-intermediate learners) and a high proficiency group (high-intermediate and advanced learners), based on participants' performance in a Mandarin proficiency test adapted from HSK[3] past papers.

The inclusion of E-C-M and C-E-M learners in the study is to achieve a mirror-image design. All participants in the E-C-M groups were English native speakers who were working or studying in Hong Kong for more than five years. In a background questionnaire, they indicated that their L2 Cantonese (including speaking, listening and reading) had reached an advanced or a near-native level and that they used Cantonese as a working, social and/or family language[4]. The C-E-M low proficiency group consisted of young immigrants who immigrated to the UK from Hong Kong before the age of 10 years old (17 participants) and British-born Chinese who used Cantonese as home language (3 participants) and were highly proficient in both Cantonese and English. All participants in the C-E-M high proficiency group were Hong Kong students who were studying in UK universities and scored 7.5 or above in IELTS when admitted to university, which indicates that they were advanced users of English.

As our test sentences in the experiment included A-not-A questions, we conducted a fill-in-the-blank task as a prerequisite and only selected participants who got 100% correct on the A-not-A questions for the study[5]. One-way ANOVAs and post-hoc tests show that there was a significant difference between the proficiency scores of the Mandarin native group and each of the learner groups ($p < .001$) and no significant difference was found between corresponding learner groups at the same proficiency levels (advanced groups: $F(2, 70) = 2.97$, $p = 0.06$; beginner groups: $F(2, 70) = 1.41$, $p = 0.25$). Table 2 provides more detailed information about the groups.

4.2 Materials and procedures

As discussed in Section 1, Mandarin allows a [+Q] SFP (*ma* or a_1) to follow a sentence to form a *yes-no* question, as illustrated in Speaker A's sentences in (5) and (6). However, *ma* and *a* differ

Table 2. Information of participants

	Group	Number	Mean age (in years)	Mean score in the Mandarin proficiency test (total = 40)
L2	E-M L	20	20.1 (1.6; 18-25)	9.1 (4.8; 2-16)
	E-M H	31	24.7 (3.0; 22-32)	31.7 (5.1; 20-36)
L3	E-C-M L	33	22.1 (6.1; 16-45)	8.8 (4.0; 1-16)
	E-C-M H	20	24.7 (8.4; 16-52)	30.2 (5.1; 21-38)
L3	C-E-M L	20	21.5 (2.4; 16-25)	9.7 (6.2; 1-19)
	C-E-M H	22	21.1 (1.9; 18-26)	32.1 (4.8; 24-39)
NS	Mandarin natives	28	24.6 (2.4; 20-30)	39.6 (1.1; 35.5-40)

a. M is for Mandarin, C for Cantonese, E for English, L for low proficiency, and H for high proficiency.
b. The figures in the brackets stand for (standard deviation; minimum-maximum).

in their interactions with A-not-A in Mandarin questions, which is constrained by the restriction on double Q-marking. The SFP *ma* cannot co-occur with A-not-A since this would involve double Q-marking, as in (7). However, a_2 in (8), which shares the same orthographic form with a_1, can follow A-not-A, because it is [-Q]. To test the two SFPs and their interactions with A-not-A in participants' L3 and L2 Mandarin, an online cross-modal priming task (CMPT) and an offline acceptability judgement task (AJT) were designed for the study.

(5) Type 1. Sentence + *ma*
    Speaker A: XiaoWang mingtian qu xuexiao ma?
                XiaoWang tomorrow go school SFP
                "Will XiaoWang go to school tomorrow?"
    Speaker B: Ta yao qu.
                he will go
                "Yes, he will."
(6) Type 2. Sentence + a_1
    Speaker A: XiaoWang mingtian qu xuexiao a?
                XiaoWang tomorrow go school SFP
                "XiaoWang will go to school tomorrow, won't he?"
    Speaker B: Ta bu qu.
                he not go
                "No, he won't."
(7) Type 3. *A-not-A + *ma*
    Speaker A: *XiaoWang wanshang kan-bu-kan dianying ma?
    XiaoWang evening watch-not-watch movie SFP
    Intended meaning: "Is XiaoWang going to watch movie tonight?"
    Speaker B: Ta bu kan.
    he not watch
    "No, he won't."
(8) Type 4. A-not-A + a_2
    Speaker A: XiaoWang wanshang   kan-bu-kan
                XiaoWang evening   watch-not-watch
                dianying   a?
                movie   SFP
    "Is XiaoWang going to watch movie tonight?"
    Speaker B: Ta yao kan.
                he will watch
                "Yes, he will."

The CMPT is a well-established method for detecting activation of lexical and syntactic information

during sentence comprehension, which integrates verbal and visual modalities (cf. Hu & Jiang, 2011; Marinis, 2018). This method was adapted in our study, in which all tokens were mini dialogues consisting of two sentences (uttered respectively by Speakers A and B), as illustrated in (5)–(8). Participants listened to a spoken prime (i.e., Speaker A's utterance without the final word) and then saw a visual display of the last word (a character) and punctuation on the computer screen. The character was either a Chinese, Japanese or Korean character. Participants were asked to make a judgement on whether the character they just saw on the computer screen is a Chinese character or not by pressing one of the designated response keys ("√" for "Yes", "✕" for "No" and "?" for "I don't know") on the keyboard. Participants were instructed to respond as quickly and accurately as possible. Participants' response times (RTs) were recorded in milliseconds (ms) as the key data, representing the effect of the syntactic structure in the auditory input on the recognition of the visual character. After pressing a designated response key on the keyboard, the participant would hear the second utterance in the dialogue (i.e., Speaker B's answer in (5) – (8)). A comprehension question based on the content of the whole mini dialogue was asked immediately after, to ensure meaningful comprehension and avoid mechanical answering. The rationale of the design is that the [+Q] or [-Q] feature attached to the utterance by Speaker A, i.e., whether the utterance contains A-not-A, will affect the time it takes to recognise the character of the SFP (*ma* or *a*) appearing on the computer screen. Twelve tokens were designed for each type and divided into four different lists based on a Latin square design. On each list, there were 12 critical items and 48 fillers/distractors half of which contained a Japanese or Korean character for the character recognition part.

In addition, a web-based acceptability judgement task was administered to all participants after the CMPT, which also included the four types, as illustrated in (5) – (8), with each type having 4 tokens corresponding to the CMPT tokens (only Speaker A's questions were included). The participant was asked to decide whether the sentence was "completely unacceptable", "probably unacceptable", "probably acceptable" or "completely acceptable". There was also an option of "I don't know".

5. Results

Data from the CMPT were composed of reaction times (RTs) and the accuracy of recognising the last character, with the former as the main data. Following Lo and Andrews (2015), our data trimming process consisted of three steps: 1) all incorrect responses to

character recognition were eliminated; 2) RTs faster than 200 ms or slower than 3,000 ms were removed; and 3) of the remaining trials, any latency that was 2.5 standard deviations away from the individual mean was also removed.[6] To compare RTs of the two SFPs (*ma* and *a*) across groups for each of the sentence types (particle questions and A-not-A questions) at the two proficiency levels (beginner and advanced), raw RTs were analysed with generalised linear mixed-effect models (*GLMM*) with Inverse Gaussian distributions (cf. Lo & Andrews, 2015). The modelling was conducted using the *glmer* function of the *lme4* package (Bates et al., 2015) in the R program 4.0.3 (R Core Team, 2021). In the *GLMM*s, Group and SFP were set as fixed effect factors and Subject and Item as random factors. The E-C-M group was chosen as the reference factor level in pair-wise comparisons of Group, since it was the mirror-image group of the C-E-M group and shared the same L1 English with the E-M group.

In the data analysis of the AJT, responses of "I don't know" were deleted and treated as missing values[7]. The four acceptability ratings were converted into numerical values of 1, 2, 3 and 4, respectively. Mean scores that reach 3 or above imply acceptance and those lower than 2 are interpreted as rejection. We checked the judgement patterns and conducted ordinal regression by following Veríssimo (2021). We generated cumulative link mixed effects models (*CLMM*) (Christensen, 2015) that included Group and SFP as fixed effects and Subject and Item as random factors, using the ordinal package *CLMM*s (Christensen, 2015) in R version 4.0.3 (R Core Team, 2021). The E-C-M group was also chosen as the reference factor level for the variable of Group.

5.1 Results of Type 1 (sentence + ma) and Type 2 (sentence + a)

Results of the CMPT

Figure 1 presents RT profiles of all groups for Types 1 and 2. To investigate whether L2/L3 learners process *a* and *ma* differently, we conducted *GLMM*s on a data set consisting of the three low proficiency groups and a set including the three high proficiency groups and the Mandarin native group, with the model outputs summarised in Table 3. At the low level, we did not observe a significant effect of SFP or an interaction effect between Group (E-C-M L vs. C-E-M L) and SFP but found a significant interaction effect between Group (E-C-M L vs. E-M L) and SFP. This suggests that neither of the L3 low proficiency groups

processed *ma* in Type 1 and *a* in Type 2 in significantly different ways, but the L2 group was divergent.

In light of the interactions, we further conducted *GLMM*s within each group, with SFP coded as the fixed effect and Subject and Item as random effects. The modelling returned a marginally significant effect of SFP in the L2 data (E-M L: β = -74.96, t = -2.05, p = .04), but not in the L3 data (C-E-M L: β = -11.79, t = -.17, p = .86; E-C-M L: β = -52.1, t = -.67, p = .50), which indicates that only the L2 group spent significantly longer RTs on *a* than on *ma*. At the high level, as shown in Table 3, there was no main effect of SFP or interaction effects found in the *GLMM*, which suggests that the L2/L3 high proficiency learners and Mandarin natives did not process *ma* differently from *a*.

Results of the AJT

As presented in Figure 2, mean scores of the four L3 groups for Types 1 and 2 are above 3, whereas the scores of the L2 groups for Type 2 are between 2.5 and 3. *CLMM* outputs from L2/L3 low proficiency learners' data are presented in Table 4. We found a robust effect of SFP, meaning that the judgements for Type 1 were significantly different from those for Type 2 among the L2/L3 low proficiency learners. Moreover, the *CLMM* returned a significant effect of Group (E-C-M L vs. E-M L) but no effect of Group (E-C-M L vs. C-E-M L) or any interaction effects, which suggests that the two L3 low proficiency groups' means were significantly different from those of the L2 group for both types. *CLMM*s were further run on the data set of Type 1 and that of Type 2 separately to compare group means. For Type 1, we found no significant Group effects (E-C-M L vs. C-E-M L: β = .31, z = 1.25, p = .21; E-C-M L vs. E-M L: β = -.24, z = -1.08, p = .27). For Type 2, we found a simple effect of Group (E-C-M L vs. E-M L) (β = -.74, z = -2.66, p = .007) but no significant effect of Group (E-C-M L vs. C-E-M L) (β = -.06, z = -.21, p = .82). This suggests that the L2 low proficiency learners' acceptance rates of Type 2 sentences were significantly lower than those of their L3 counterparts.

For the data set of the high proficiency learner groups and the Mandarin native group, as shown in Table 4, the *CLMM* returned a significant effect of SFP and a simple effect of Group (E-C-M H vs. NS) but no interaction effects. This implies that the L3 high proficiency learners shared a similar pattern with their L2 counterparts: they judged Type 1 *ma* sentences significantly
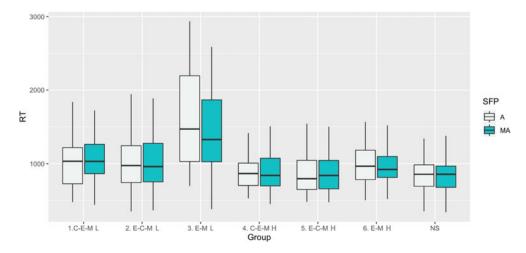

Figure 1. Mean reaction times (in millisecond) of Type 1 (sentence + ma) and Type 2 (sentence + a) in the CMPT

Table 3. GLMM outputs for Type 1 (sentence + ma) and Type 2 (sentence + a) in the CMPT

		β	SE	t	p
Beginner	(Intercept)	1235.1	165.34	7.47	< .001 ***
	Group (E-C-M L vs. C-E-M L)	35.47	132.95	.26	.789
	Group (E-C-M L vs. E-M L)	485.05	214.78	22.45	.023 *
	SFP (A vs. MA)	−18.36	92.74	−.19	.843
	Group (E-C-M L vs. C-E-M L) : SFP	32.40	50.00	.64	.517
	Group (E-C-M L vs. E-M L) : SFP	−130.17	66.75	−2.03	.048 *
Advanced	(Intercept)	930.95	48.41	19.23	< .001 ***
	Group (E-C-M H vs. C-E-M H)	19.73	48.62	.41	.685
	Group (E-C-M H vs. E-M H)	126.77	56.35	2.24	.024 *
	Group (E-C-M H vs. NS)	15.37	59.86	.26	.797
	SFP (A vs. MA)	7.31	35.49	.21	.836
	Group (E-C-M H vs. C-E-M H): SFP	2.39	35.33	.07	.946
	Group (E-C-M H vs. E-M H) : SFP	−31.82	32.49	.98	.327
	Group (E-C-M H vs. NS) : SFP	−62.65	32.38	−1.93	.053

Significance codes: *** $p < 0.001$; ** $p < 0.01$; * $p < 0.05$

more acceptable than Type 2 *a* sentences, although their mean scores were not nativelike.

5.2 Results of Type 3 (*A-not-A + ma) and Type 4 (A-not-A + a)

Results of the CMPT

Mean RTs of all groups are presented in Figure 3. As shown in Table 5, at the low proficiency level, the *GLMM* found no significant effect of SFP but a simple effect of Group (E-C-M L vs. E-M L) and a significant interaction effect between Group (E-C-M L vs. E-M L) and SFP. To further investigate the RT difference between the two types across groups, we conducted GLMMs within each group and found a main effect of SFP in the L2 data (E-M L: β = -119.11, t = -8.75, p = .003), but not in the L3 data (C-E-M L: β = -61.95, t = -.60, p = .54; E-C-M L: β = -16.73,

t = -.25 p = .79), which suggests that only the L2 low proficiency group processed the two SFPs differently.

At the high proficiency level, as shown in Table 5, we found a significant effect of SFP, a simple effect of Group (E-C-M H vs. E-M H) and a significant interaction between Group (E-C-M H vs. E-M H) and SFP but no other interactions. This suggests that the L3 advanced groups and the Mandarin native group spent significantly longer RTs on *ma* than on *a*, indicating sensitivity to the illicit double Q-marking in Type 3, whereas the L2 advanced group showed an opposite pattern.

Results of the AJT

As presented in Figure 4, the mean scores of all the groups for Type 4 are above 3, which indicates acceptance of this type of sentences. However, their mean scores for Type 3 sentences vary. As
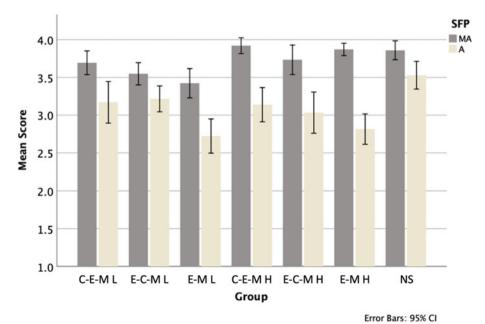

Error Bars: 95% CI

Figure 2. Mean acceptability judgment scores for Type 1 (sentence + ma) and Type 2 (sentence + a) in the AJT

Table 4. CLMM outputs for Type 1 (sentence + ma) and Type 2 (sentence + a) in the AJT

		β	SE	z	p
Beginner	Group (E-C-M L vs. C-E-M L)	−.052	.230	−.228	.819
	Group (E-C-M L vs. E-M L)	−.653	.224	−2.910	.003**
	SFP (A vs. MA)	.559	.177	3.155	.001**
	Group (E-C-M L vs. C-E-M L) : SFP	.371	.309	1.199	.230
	Group (E-C-M L vs. E-M L) : SFP	.402	.287	1.401	.161
Advanced	Group (E-C-M H vs. C-E-M H)	.074	.399	.187	.851
	Group (E-C-M H vs. E-M H)	−.408	.369	−1.106	.268
	Group (E-C-M H vs. NS)	.956	400	2.389	.016 *
	SFP (A vs. MA)	1.772	.333	5.317	< .001 ***
	Group (E-C-M H vs. C-E-M H): SFP	.593	.439	1.352	.176
	Group (E-C-M H vs. E-M H) : SFP	.635	.365	1.740	.081
	Group (E-C-M H vs. NS) : SFP	−.035	.443	−.080	.936

Significance codes: *** $p < 0.001$; ** $p < 0.01$; * $p < 0.05$

shown in Table 6, the *CLMM* conducted on the low proficiency groups' data set returned no effect of Group or SFP but a significant interaction between Group (E-C-M L vs. E-M L) and SFP. To compare the group mean scores for each type, *CLMM*s were further run in the data sets of Types 3 and 4. For Type 3, we found a significant Group effect between the E-C-M L group and the E-M L group ($\beta = -1.31$, $z = -3.01$, $p = .002$), but not between the two L3 groups ($\beta = -.04$, $z = -.09$, $p = .92$). This indicates that the L2 low proficiency judged the illicit Type 3 as significantly less acceptable than their L3 counterparts. For Type 4, the *CLMM* found no Group effects (E-C-M L vs. C-E-M L: $\beta = .54$, $z = 1.89$, $p = .06$; E-C-M L vs. E-M L: $\beta = .11$, $z = .41$, $p = .68$).

When it comes to the high proficiency level, as shown in Table 6, the *CLMM* found a significant effect of SFP and significant interactions between different group pairs and SFP, which suggests that high proficiency L2/L3 learners start to show some sensitivity to the illicit Q-marking. To compare group means for each SFP, *CLMM*s were run on the data set of Type 3 and that of the Type 4, separately. For Type 3, we found that the two L3 high proficiency groups (E-C-M H vs. C-E-M H: $\beta = -.76$, $z = -1.37$, $p = .17$) behaved significantly differently from their L2 counterparts (E-C-M H vs.

E-M H: $\beta = -2.37$, $z = -4.01$, $p < .001$) and from the Mandarin natives (E-C-M H vs. NS: $\beta = -2.25$, $z = -3.91$, $p < .001$), which shows that the L3 high proficiency groups were not as sensitive as the L2 group and Mandarin natives to the illicit Q marking. In the *CLMM* outputs of Type 4, we did not find main effects of Group between the learner groups (E-C-M H vs. C-E-M H: $\beta = 72$, $z = 1.15$, $p = .24$; E-C-M H vs. E-M H: $\beta = -.61$, $z = -1.11$, $p = .26$) but between the E-C-M H group and the native group ($\beta = 2.31$, $z = 3.85$, $p < .001$), which suggests that the L2/L3 learners accepted Type 4 sentences to a significantly lesser extent than Mandarin natives even though they all judged Type 4 as acceptable in general (Means > 3).

6. Discussion

6.1 Transfer source at L3 initial stages

A key aim of this study (research question 1) is to ascertain the transfer source at L3 initial stages. The mirror image design of L3 groups (i.e., C-E-M vs. E-C-M) and the comparison between the L2 and L3 low proficiency groups provide us with answers
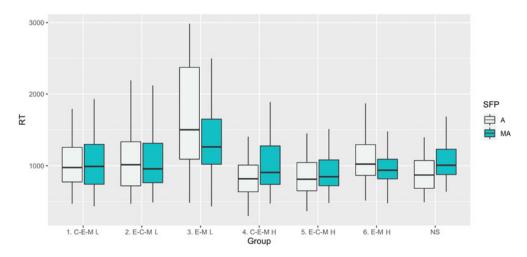

Figure 3. Mean reaction times (in millisecond) of Type 3 (*A-not-A + ma) and Type 4 (A-not-A + a) in the CMPT

Table 5. GLMM outputs for Type 3 (*A-not-A + ma) and Type 4 (A-not-A + a) in the CMPT

		β	SE	t	p
Beginner	(Intercept)	1299.50	68.97	18.84	< .001 ***
	Group (E-C-M L vs. C-E-M L)	17.57	98.48	.17	.858
	Group (E-C-M L vs. E-M L)	428.18	73.75	5.81	< .001 ***
	SFP (A vs. MA)	6.98	59.62	.19	.846
	Group (E-C-M L vs. C-E-M L) : SFP	−23.45	51.94	−.45	.651
	Group (E-C-M L vs. E-M L) : SFP	−172.06	63.17	−2.72	.006 **
Advanced	(Intercept)	914.83	69.51	13.16	< .001 ***
	Group (E-C-M H vs. C-E-M H)	7.71	60.41	.128	.898
	Group (E-C-M H vs. E-M H)	169.87	55.70	3.04	.002 **
	Group (E-C-M H vs. NS)	15.58	62.19	.25	.802
	SFP (A vs. MA)	125.24	57.18	2.19	.028 *
	Group (E-C-M H vs. C-E-M H): SFP	11.55	44.32	.26	.794
	Group (E-C-M H vs. E-M H) : SFP	−154.85	39.98	−3.87	< .001 ***
	Group (E-C-M H vs. NS) : SFP	58.13	38.83	1.49	.134

Significance codes: *** $p < 0.001$; ** $p < 0.01$; * $p < 0.05$

to the first research question about transfer source selection. As discussed in Section 1, Cantonese is similar to Mandarin regarding the Q-operation, as [+Q] in both languages is overtly realised as an SFP or as A-not-A rather than auxiliary-verb inversion or *do*-support, and the two Q devices cannot co-occur.

In the CMPT, neither of the L3 low proficiency groups (E-C-M L and C-E-M L) processed *ma* differently from *a*, while the E-M group always spent longer times on *ma* than on *a*. In the AJT, the L3 low proficiency groups always patterned together on all types and behaved differently from the L2 group on Types 2 and 3. Recall that the difference between L3 and L2 groups is that the former, but not the latter, had the knowledge of Cantonese. Our results suggest that the L3 beginners are not influenced by their English. Cantonese, the structurally more similar language, is the main source of transfer. The findings provide strong evidence against the predictions of the L1 factor hypothesis that L1 plays a

privileged role in initial stage transfer. Our data can also falsify the L2 Status Factor Hypothesis as the C-E-M learners' L2 English does not influence their L3 Mandarin in either of the tasks, even though their L2 English and their L3 Mandarin were mostly learned via formal instructions. These results confirm findings from the majority of previous L3 studies reviewed in Puig-Mayenco et al. (2020) that the L1/L2 is not always the main (only) source of transfer. In general, our data support the Typological Primacy Model, the Cumulative Enhancement Model, the Linguistic Proximity Model and the Scalpel Model, in which structural similarity plays a key role in transfer source selection.

6.2 Causes of facilitative and detrimental effects

Research question 2 investigates detailed transfer effects and relevant causes. Our data show that Cantonese does not always
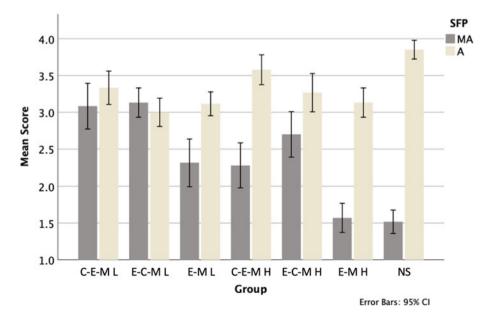

Figure 4. Mean acceptability judgment scores for Type 3 (*A-not-A + *ma*) and Type 4 (A-not-A + *a*) in the AJT

Table 6. CLMM outputs for Type 3 (*A-not-A + ma) and Type 4 (A-not-A + a) in the AJT

		β	SE	z	p
Beginner	Group (E-C-M L vs. C-E-M L)	.422	.283	1.491	.135
	Group (E-C-M L vs. E-M L)	.082	.266	.309	.757
	SFP (A vs. MA)	.173	.179	.967	.333
	Group (E-C-M L vs. C-E-M L) : SFP	−.493	.295	−1.668	.095
	Group (E-C-M L vs. E-M L) : SFP	−1.167	.278	−4.192	< .001 ***
Advanced	Group (E-C-M H vs. C-E-M H)	.545	.327	1.667	.095
	Group (E-C-M H vs. E-M H)	−.219	.288	−.761	.446
	Group (E-C-M H vs. NS)	1.416	.355	4.114	< .001 ***
	SFP (A vs. MA)	−.722	.221	−3.266	.001 **
	Group (E-C-M H vs. C-E-M H): SFP	−1.041	.318	−3.271	.001 **
	Group (E-C-M H vs. E-M H) : SFP	−1.238	.288	−4.294	< .001 ***
	Group (E-C-M H vs. NS) : SFP	−2.951	.367	−8.039	< .001 ***

Significance codes: *** $p < 0.001$; ** $p < 0.01$; * $p < 0.05$

assist L3ers' acquisition of Mandarin: both facilitative and detrimental transfer effects are observed at initial stages, which rejects the Cumulative Enhancement Model that explicitly excludes the possibility of detrimental transfer. This finding echoes many L3 studies that also show evidence of non-facilitative transfer (cf. Puig-Mayenco et al., 2020). Recall that Cantonese has equivalent SFPs of Mandarin a_1 and a_2, but only has an approximate counterpart of the Mandarin *ma*. A deeper dive into features attached to the two SFPs in L3 initial-stage grammars can help us further understand details of transfer and non-facilitation.

Facilitative effects on the L3 acquisition of questions with $a_{1/2}$
The Mandarin SFPs a_1 and a_2 correspond to the [+Q] SFP *aa4* and the [-Q] *aa3* in Cantonese, respectively. As these Mandarin and Cantonese SFPs all share the same character 啊 and sound similar to each other phonetically, it is very natural for L3ers to transfer the features of *aa4* and *aa3* to the Mandarin a_1 and a_2 at initial stages and accept both the [+Q] a_1 and the [-Q] a_2. The L2 beginners, however, have no knowledge of Cantonese, and thus have to entirely rely on the Mandarin input to acquire the SFPs. The character 啊 is shared by multiple SFPs in Mandarin: it can be used to express an exclamative, imperative or interrogative meaning, with the first two as its main functions. In the corpus of the Centre for Chinese Linguistics at Peking University[8], there are around 58,000 occurrences of 啊 in non-interrogative sentences and around 7000 in questions (around half of which are for the [+Q] a_1 and the other half for the [-Q] a_2, e.g., in A-not-A or *wh*-questions). Compared to the [+Q] of *ma* (108,000 occurrences), the [+Q] feature of a_1 is more difficult to acquire as the (phonological and orthographic) form of a_1 (i.e., *a* 啊) mainly appears in non-interrogative contexts. This can account for the difficulty that the L2 beginners had in handling *a*. Without any knowledge of Cantonese, the L2 beginners were able to accept the [-Q] function of *a* (Type 4) but were uncertain about whether *a* has a [+Q] feature. It is obviously challenging for L2 beginners, who have no experience in dealing with SFPs, to make distinctions between the multi-functions of *a*.

Detrimental effects on the L3 acquisition of questions with ma
As discussed above, the L3 low proficiency groups outperformed their L2 counterpart on Type 2 questions with *a* in the CMPT and the AJT, but they were less native-like than their L2 counterparts on the ungrammatical A-not-A questions in the AJT. The L3 low proficiency learners accepted all types tested in the AJT, including the ungrammatical Type 3. An important question here is why they erroneously accepted these double [+Q] marking sentences.

We can rule out the possibility that the [+Q] feature is not attached to A-not-A in their L3 Mandarin, because our prerequisite test has ensured that participants selected all acquired that A-not-A is interrogative. The SFP *ma* is the most frequently used *yes-no* question particle in Mandarin (Li & Thompson, 1989; Zhu, 1982). Although Cantonese has only an approximate counterpart but not an equivalent of *ma* and English has no question particles, the high frequency and the transparent semantics of *ma* make its [+Q] feature rather salient in the input and, therefore, easy to acquire for both L2 and L3 beginners. This is supported by the L2 and L3 data of Type 1 questions. Moreover, given that none of the three languages involved allows double Q-marking, it is very unlikely that the L3ers start with a grammar that allows the non-economical way of double Q-marking. Therefore, to account for why the L3 low proficiency groups accepted ungrammatical A-not-A questions with the SFP *ma* in Type 3, the only possibility left is that a [-Q] feature is also attached to *ma* in their L3 initial Mandarin grammars.

In Cantonese, the SFP *maa3* approximately corresponding to the Mandarin *ma* is written either as 嗎 or 嘛: the former is used in formal texts and can only be used as a [+Q] SFP, whereas the latter is a character shared by both the [+Q] SFP *maa3* and the [-Q] SFP *maa5* indicating that something is obvious. At initial stages, it is natural for L3ers to associate the Mandarin *ma* with the [+Q] Cantonese *maa3* 嗎/嘛 based on the similarities of sounds and meanings. In addition, the [+Q] *maa3* 嘛 and the [-Q] *maa5* 嘛 share the same orthographic form in Cantonese. The similarities in sounds and the identical orthographic form shared are believed to serve as detrimental cues, leading the L3 beginners to erroneously map features of both Cantonese [+Q] *maa3* and [-Q] *maa5* onto the Mandarin SFP *ma*.

Consequently, the Mandarin *ma* at L3 initial stages can function as either a [+Q] or a [-Q] SFP. This kind of two-to-one mapping paradigm between Cantonese and Mandarin SFPs results from three main factors: the significantly bigger number of SFPs in Cantonese than that in Mandarin, similarities between the two languages in phonological and logographic forms, and the lack of standardisation in the Cantonese writing system. Consequently, "overgeneralisation" or "overuse" of SFPs is commonly observed in Cantonese speakers' L2/L3 Mandarin.

If the analysis above stands, the learning situation of the L3ers is different from that of the L2ers: the L3ers face a feature unlearning process, whereas the L2ers need to acquire a new way to instantiate [+Q] – namely, the Mandarin SFP *ma*. The L2 low proficiency learners judged the ungrammatical Type 3 as significantly less acceptable than Type 4. This is because they had no knowledge of Cantonese, and the Mandarin input that they were exposed to is the only source of information concerning the use of Mandarin SFPs. The SFP *ma* is always used in interrogative sentences in Mandarin, which helps L2ers acquire the [+Q] feature of *ma* in a rather straightforward manner. This provides an account for the E-M L group's sensitivity to ungrammatical sentences in Type 3.

Disentangling cues of different domains

Questions about what triggers transfer and what kind of input is misleading can further be answered with our data. The findings clearly show that transfer can take place at the morpho-lexical level – namely, the unit of transfer can be functional items (e.g., individual SFPs). Features bundled on a certain functional item can be transferred into L3 initial-stage grammars. At initial stages, L3ers will first search in their previously acquired languages for items perceived to be identical or similar to those in the L3 and then map them onto the perceived corresponding items in the L3. In this attempt, cues in different domains (syntax, semantics, morphology, phonology, orthography, etc.) can trigger transfer from previously acquired languages to L3. In other words, L3ers do not just check cues in one specific domain, but rather scan different domains for potential cues when they are initially exposed to L3 input.

It is noteworthy that cues from different domains may play various roles in the L3 initial-stage grammar. Syntactic cues concerning macro-level grammars, such as the formation of *yes-no* questions (subject-auxiliary inversion, A-not-A or SFPs), serve for the measurement of structural similarity and play a crucial role in the transfer source selection. Phonological/phonetic and orthographic cues of individual SFPs are about the external realisation of functional items, and they are more language specific. Although they are not directly related to the core grammar, they can trigger (non-)facilitative transfer in L3 acquisition, as observed in the detrimental transfer from Cantonese in our L3 Mandarin study.

6.3 Development of L2 and L3 Mandarin grammars

The last research question compares the development and attainment of L3 and L2 acquisition. The L3 low proficiency groups correctly allowed the use of *ma* (Type 1) and a_1 (Type 2) in the AJT and processed *ma* and *a* similarly in the CMPT, whereas their L2 counterparts had difficulty with the [+Q] SFP a_1. When it comes to the high proficiency level, the L2/L3 discrepancy disappears in the AJT but remains in the CMPT; the L2 and L3 advanced learners accepted the use of *ma* and *a* in particle questions but the L2ers still found it more difficult to process *a* than *ma*.

These findings suggest that the facilitative transfer from Cantonese pertains at L3 later stages and helps the L3 advanced learners acquire both implicit and explicit knowledge of the two types of particle questions. However, it seems arduous for L2 learners to fully acquire the [+Q] of a_1 and integrate this knowledge into automatic online processing, which is presumably attributable to the multiple functions of the Mandarin *a* 啊. Moreover, word frequency also seems to affect the acquisition result. All groups judged *ma* questions significantly more acceptable than *a* questions, although the scores of the two types are within the threshold of acceptance across groups. The successful L2/L3 acquisition of *ma* is believed to be largely influenced by the high frequency of *ma* questions in Mandarin. In contrast, the [+Q] a_1 is of low frequency and non-salient in comparison with the other SFPs sharing the same character and sound, which results in a relatively lower degree of acceptance across groups.

The L3 groups' online processing patterns were consistent with their offline judgement at both proficiency levels whereas the L2 groups' online and offline behaviours were different. As discussed in the previous section, the L3 beginners benefit from the facilitative transfer from Cantonese cues in the case of $a_{1/2}$, but due to the phonologically and orthographically detrimental cues of the [-Q] SFP *maa5* in Cantonese, they erroneously transfer the [-Q] feature of the non-interrogative Cantonese SFP into the Mandarin *ma* at initial stages. The task facing L3ers at later stages is to discard the [-Q] feature from the feature set of *ma* in their L3. In the AJT, we can see that, although the L3 high proficiency learners could differentiate between Type 3 and Type 4, they could not reject the former type as firmly as the L2ers did, which suggests that unlearning a certain feature transferred can be more arduous than acquiring a new way of instantiation.

Since the L3ers showed some sensitivity to the illicit double Q-marking in both the AJT and the CMPT at advanced stages, another interesting question to ask is how they retreat from the overgeneralisation. We argue that the orthographic form (Chinese characters) is a helpful cue for them to differentiate between the functions of SFPs. Unlike Cantonese, Mandarin has a standardised writing system, in which the character of 吗 is for the [+Q] SFP *ma* only, and the character of 嘛 is always for a non-interrogative SFP *ma*, which denotes a meaning of "obviously" or an imperative function. With more exposure to Mandarin, and particularly with the increased ability to distinguish Mandarin characters orthographically, the distinction between 吗 and 嘛 can serve as a useful cue triggering the attachment of [+Q] onto 吗 *ma* as well as the dissociation of [-Q] from 吗*ma* in their L3 Mandarin. This is confirmed by the L3 data of the CMPT, which presented SFPs in Chinese characters. The advanced L3 learners showed sensitivity to the interaction between A-not-A and the SFP吗, which indicates that they gradually associate the orthographic form of 吗only with the [+Q] feature in their L3 grammars.

In general, the L3 groups had some difficulty with *ma* even at advance stages, which suggests that influences from Cantonese remain in advanced L3 grammars. However, this difficulty can be overcome as the L3ers become able to distinguish SFPs orthographically, resulting in the proper attachment of [+Q] to and dissociation of [-Q] from 吗 *ma* in their L3 Mandarin development.

7. Conclusion

The present study is devoted to L3 acquisition of Mandarin *yes-no* questions by English–Cantonese bilinguals at low and high-

proficiency levels. By comparing L3 and L2 Mandarin grammars, it is found that structural similarity plays a deterministic role in transfer source selection at L3 initial stages and syntactic cues (the use of SPFs) serve as the measurement of structural similarity. Both facilitative and detrimental transfer effects are observed. Our findings reject the L3 models that advocate a privileged role of L1/L2 in L3 initial stage transfer and the Cumulative Enhancement Model that excludes non-facilitative transfer in L3. When a Mandarin SFP has an equivalent in Cantonese (as in the case of questions with $a_{1/2}$), learners' L1/L2 Cantonese assists their L3 Mandarin acquisition process. Orthographic and phonological cues of individual SFPs are found to be a main cause of non-facilitation in our study. In the case of *ma*, orthographic similarities between Mandarin and Cantonese trigger a mis-mapping between relevant SFPs and result in an overgeneralisation requiring a feature discarding process (from [±Q] to [+Q]). The detrimental influence from Cantonese remains in later development, i.e., in advanced L3 Mandarin grammars, which suggests that feature unlearning can be more arduous than learning a new instantiation. Unlike L2ers, who performed differently in the online and offline tasks, L3ers' online processing is consistent with their offline judgement at both proficiency levels, which indicates that facilitative and detrimental transfer from Cantonese exists in both their implicit and explicit knowledge.

This study has adopted a de-compositional approach to cues, which we believe is useful in accounting for some sporadic and irregular patterns observed in initial and later stages of L3 grammars. Further examinations of a wider range of linguistic properties and new language combinations are in demand to systematically investigate the exact effect of cues of different domains (syntax, semantics, morphology, phonology, orthography, register, etc.) on L3 acquisition.

Notes

[1] Some scholars argue that the form *maa3* consists of the negation marker *m* "not" and the tone-softening SFP *aa3* and thus it cannot co-occur with negators (Leung, 2005; Sybesma & Li, 2007).

[2] Cantonese is mainly used as a colloquial language and lacks standardisation of the writing script. This results in a hybrid writing system of Cantonese, a mixture of colloquial written forms and standard Chinese forms, in which a certain lexical item might be associated with different characters.

[3] The HSK is an official Mandarin language proficiency test established by the Chinese government.

[4] The Hong Kong government has implemented a 'biliterate and trilingual' language policy since 1997, aiming at developing biliterate (written Chinese and English) and trilingual (Cantonese, Mandarin and English) citizens. Cantonese and Mandarin differ in the spoken language but their written forms enjoy a high degree of similarity, although Cantonese uses the traditional Chinese characters and Mandarin adopts the simplified version.

Consequently, Cantonese proficiency cannot be assessed with written tests. To the best of our knowledge, no official language test is available assessing non-native speakers' Cantonese proficiency and therefore we had to rely on the participants' self-evaluations in eight questions in the questionnaire concerning their Cantonese proficiency, the frequency of their use of the language and the quantity of Cantonese input they were exposed to, etc.

[5] The prerequisite test (3 tokens) conducted before the main experiment was to ensure that participants are able to form A-not-A questions, as illustrated in the sample item below. Two E-M learners and one E-C-M learner did not pass the screening test.

小王 XiǎoWáng __ 哥哥 gēge 今天 jīntiān __ __ 去 qù __？(学校 xuéxiào，不 bú，的 de，去 qù)

□ Impossible to use all the words provided to make a correct Mandarin sentence.

[6] The trimming procedure affected 8.5% of the learners' data and 5.7% of the Mandarin native speakers'.

[7] This procedure affected 2.7% of the learners' data.

[8] The PKU-CCL Corpus contains 581 million characters of text written in modern Chinese and covers a wide range of genres (e.g., spoken, fictions, newspapers, movies and academic books/articles).

References

Abbas, N., Degani, T., & Prior, A. (2021). Equal Opportunity Interference: Both L1 and L2 Influence L3 Morpho-Syntactic Processing. *Frontiers in Psychology*, 12, 673535. https://doi.org/10.3389/fpsyg.2021.673535

Alonso, J. G., Banón, J. A., DeLuca, V., Miller, D., Soares, S. M. P., Puig-Mayenco, E., Slaats, S., & Rothman, J. (2020). Event related potentials at initial exposure in third language acquisition: Implications from an artificial mini-grammar study. *Journal of Neurolinguistics*, 56, 100939. https://doi.org/10.1016/j.jneuroling.2020.100939

Bardel, C., & Falk, Y. (2007). The role of the second language in third language acquisition: The case of Germanic syntax. *Second Language Research*, 23(4), 459–484. https://doi.org/10.1177/0267658307080557

Bardel, C., & Falk, Y. (2012). Behind the L2 Status Factor: A neurolinguistic framework for L3 research. In J. Cabrelli Amaro, S. Flynn, & J. Rothman (Eds.), *Third language acquisition in adulthood* (pp. 61–78). John Benjamins.

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting Linear Mixed-Effects Models Using lme4. *Journal of Statistical Software*, 67(1), 1–48. doi:10.18637/jss.v067.i01.

Christensen, R. H. B. (2015). *Ordinal: Regression models for ordinal data.* R package version, 28, 2015.

De Bot, K., & Jaensch, C. (2015). What is special about L3 processing? *Bilingualism: Language and Cognition*, 18(2), 130–144. https://doi.org/10.1017/S1366728913000448

Ellis, R. (2005). Measuring implicit and explicit knowledge of a second language: A psychometric study. *Studies in Second Language Acquisition*, 27 (2), 141–172. https://doi.org/10.1017/S0272263105050096

Falk, Y., & Bardel, C. (2011). Object pronouns in German L3 syntax: Evidence for the L2 status factor. *Second Language Research*, 27(1), 59–82. https://doi.org/10.1177/0267658310386647

Flynn, S., Foley, C., & Vinnitskaya, I. (2004). The cumulative-enhancement model for language acquisition: Comparing adults' and children's patterns of development in first, second and third language acquisition of relative clauses. *The International Journal of Multilingualism*, 1(1), 3–16. https://doi.org/10.1080/14790710408668175

García Mayo, M. D. P., & Slabakova, R. (2015). Object drop in L3 acquisition. *International Journal of Bilingualism*, 19(5), 483–498. https://doi.org/10.1177/1367006914524643

Godfroid, A., Loewen, S., Jung, S., Park, J. H., Gass, S., & Ellis, R. (2015). Timed and untimed grammaticality judgments measure distinct types of knowledge: Evidence from eye-movement patterns. *Studies in Second Language Acquisition*, 37(2), 269–297. https://doi.org/10.1017/S0272263114000850

Guo, Y., & Yuan, B. (2020). Facilitative transfer only? L3 acquisition of Mandarin sentence-final particle clusters by English-Cantonese bilinguals. *Journal of Second Language Studies*, 3(2), 261-289. https://doi.org/10.1075/jsls.00013.guo

Guo, Y., & Yuan, B. (2021). Asymmetric transfer and development of temporal-aspectual sentence-final particles in English-Cantonese bilinguals' L3 Mandarin grammars. *Linguistic Approaches to Bilingualism.* https://doi.org/10.1075/lab.20116.guo. Published online by John Benjamins on November 26, 2021.

Hermas, A. (2010). Language acquisition as computational resetting: verb movement in L3 initial state. *International Journal of Multilingualism, 7* (4), 202–224. https://doi.org/10.1080/14790718.2010.487941

Holmberg, A. (2015). *The syntax of yes and no.* Oxford University Press. https://doi.org/10.1093/acprof:oso/9780198701859.001.0001

Hu, G. G., & Jiang, N. (2011). Semantic integration in listening comprehension in a second language. *Applying priming methods to L2 learning, teaching and research*, 199–218. http://digital.casalini.it/9789027286925

Huang, C. T. J., Li, T. H. A., & Li, Y. (2009). *The Syntax of Chinese.* Cambridge University Press. https://doi.org/10.1017/CBO9781139166935

Jensen, I. N., Mitrofanova, N., Anderssen, M., Rodina, Y., Slabakova, R., & Westergaard, M. (2023). Crosslinguistic influence in L3 acquisition across linguistic modules. *International Journal of Multilingualism, 20*(3), 717-734. https://doi.org/10.1080/14790718.2021.1985127.

Jin, F. (2009). Third Language Acquisition of Norwegian Objects: Interlanguage Transfer or L1 Influence? In Y. K. I. Leung (Ed.) *Third Language Acquisition and Universal Grammar* (pp.144–161). Multilingual Matters.

Kwok, H. (1984). *Sentence Particles in Cantonese, Vol. 56.* University of Hong Kong.

Lardiere, D. (2009). Some thoughts on the contrastive analysis of features in second language acquisition. *Second language research, 25*(2), 173-227. https://doi.org/10.1177/0267658308100283

Law, A. (2002). Cantonese sentence-final particles and the CP domain. *UCL working papers in linguistics, 14*, 375–398.

Leung, C. S. (1992). *A Study of the Utterance Particles in Cantonese as Spoken in Hong Kong* [Unpublished M.Phil. thesis], Hong Kong Polytechnic.

Leung, Y. K. I. (2005). L2 vs. L3 initial state: A comparative study of the acquisition of French DPs by Vietnamese monolinguals and Cantonese–English bilinguals. *Bilingualism: Language and cognition, 8*(1), 39–61. https://doi.org/10.1017/S1366728904002044

Li, B. Y. (2006). *Chinese final particles and the syntax of the periphery* [Unpublished doctoral dissertation]. University of Leiden.

Li, C., & Thompson, S. (1989). *Mandarin Chinese: A Functional Reference Grammar.* University of California Press.

Lo, S., & Andrews, S. (2015). To transform or not to transform: Using generalized linear mixed models to analyse reaction time data. *Frontiers in psychology, 6*, 1171. https://doi.org/10.3389/fpsyg.2015.01171

Marinis, T. (2010). Using on-line processing methods in language acquisition research. In E. Blom, & S. Unsworth (Eds.), *Experimental Methods in Language Acquisition Research* (pp. 139–162). John Benjamins. http://digital.casalini.it/9789027287953

Marinis, T. (2018). Cross-modal priming in bilingual sentence processing. *Bilingualism: Language and Cognition, 21*(3), 456–461. https://doi.org/10.1017/S1366728917000761

Matthews, S., & Yip, V. (1994). *Cantonese: A Comprehensive Grammar.* Routledge.

Na Ranong, S., & Leung, Y. K. I. (2009). Null objects in L1 Thai-L2 English-L3 Chinese: An empirical take on a theoretical problem. In Y. K. I. Leung (Ed.), *Third Language Acquisition and Universal Grammar* (pp. 162–191). Multilingual Matters.

Paradis, M. (2009). *Declarative and procedural determinants of second languages.* John Benjamins. https://doi.org/10.1075/sibil.40

Paul, W. (2015). *New Perspectives on Chinese Syntax.* De Gruyter Mouton. https://doi.org/10.1515/9783110338775

Puig-Mayenco, E., González Alonso, J., & Rothman, J. (2020). A systematic review of transfer studies in third language acquisition. *Second Language Research, 36*(1), 31-64. https://doi.org/10.1177/0267658318809147

R Core Team. (2021). *R: A language and environment for statistical computing.* https://www.R-project.org/

Rothman, J. (2010). On the typological economy of syntactic transfer: Word order and relative clause high/low attachment preference in L3 Brazilian Portuguese. *International Review of Applied Linguistics in Language Teaching, 48*, 245–273. https://doi.org/10.1515/iral.2010.011

Rothman, J. (2011). L3 syntactic transfer selectivity and typological determinacy: The typological primacy model. *Second Language Research, 27*(1), 107–127. https://doi.org/10.1177/0267658310386439

Rothman, J. (2015). Linguistic and cognitive motivation for the typological primacy model of third language (L3) transfer: Considering the role of timing of acquisition and proficiency in the previous languages. *Bilingualism: Language and Cognition 18*(2), 179–190. https://doi.org/10.1017/S136672891300059X

Rothman, J., González Alonso, J., & Puig-Mayenco, E. (2019). *Third language acquisition and linguistic transfer.* Cambridge University Press. https://doi.org/10.1017/9781316014660

Schwartz, B. D., & Sprouse, R. A. (1996). L2 cognitive states and the full transfer/full access hypothesis. *Second Language Research, 12*(1), 40–72. https://doi.org/10.1177/026765839601200103

Schwartz, B. D., & Sprouse, R. A. (2021). The full transfer/full access model and L3 cognitive states. *Linguistic Approaches to Bilingualism, 11*(1), 1–29. https://doi.org/10.1075/lab.20055.sch

Slabakova, R. (2017). The scalpel model of third language acquisition. *International Journal of Bilingualism, 21*(6), 651–665. https://doi.org/10.1177/1367006916655413

Slabakova, R., & García Mayo, M. P. (2015). The L3 syntax-discourse interface. *Bilingualism: Language and Cognition, 18*(2), 208–226. https://doi.org/10.1177/S1366728913000369

Sybesma, R., & Li, B. (2007). The dissection and structural mapping of Cantonese sentence final particles. *Lingua, 117*(10), 1739–1783. https://doi.org/10.1016/j.lingua.2006.10.003

Tang, S. W. (2015). *Lectures on Cantonese grammar.* The Commercial Press.

Veríssimo, J. (2021). Analysis of rating scales: A pervasive problem in bilingualism research and a solution with Bayesian ordinal models. *Bilingualism: Language and Cognition, 24*(5), 842–848. https://doi.org/10.1017/S1366728921000316

Westergaard, M. (2021a). Microvariation in multilingual situations: The importance of property-by-property acquisition. *Second Language Research, 37*(3), 379–407. https://doi.org/10.1177/0267658319884116

Westergaard, M. (2021b). L3 acquisition and crosslinguistic influence as co-activation: Response to commentaries on the keynote 'Microvariation in multilingual situations: The importance of property-by-property acquisition'. *Second Language Research, 37*(3), 501–518. https://doi.org/10.1177/02676583211007897

Westergaard, M., Mitrofanova, N., Mykhaylyk, R., & Rodina, Y. (2017). Crosslinguistic influence in the acquisition of a third language: The Linguistic Proximity Model. *International Journal of Bilingualism, 21*(6), 666–682. https://doi.org/10.1177/1367006916648859

Zhu, D. X. (1982). *On grammar.* The Commercial Press.