

Li, Y. & H. Goad (2024) Naïve English-speaking learners' use of indirect positive evidence: The case of Mandarin plural marking. *Linguistic Approaches to Bilingualism* 14: 115-146. [online first 29 July 2022]
<https://doi.org/10.1075/lab.21024.li>

Naïve English-speaking Learners' Use of Indirect Positive Evidence: The Case of Mandarin Plural Marking

Ying Li & Heather Goad

liying_22@163.com

Qinghai University

heather.goad@mcgill.ca

McGill University

June 2022

Abstract

When second language learners are faced with acquiring a grammar that is a subset of their native language grammar, direct positive evidence is unavailable. We question whether learners can instead use indirect positive evidence: evidence drawn from errors in the learner's L1 made by native speakers of the learner's L2. We examine if naïve English-speaking learners of Mandarin can determine from plural omission errors in Mandarin speakers' English productions that Mandarin marks plural in a subset of conditions under which English does. Participants were exposed to indirect positive evidence via an English-medium dialogue where a native Mandarin-speaking interlocutor produced all contextually plural nouns as singulars. Subsequently, participants learnt 12 Mandarin-like nouns in singular contexts, after which their word learning was tested using both singular and plural pictures as prompts. 40% of participants correctly deduced that strings to which they had assigned singular interpretations were also appropriate in plural contexts. Follow-up questions revealed that they noticed the errors in the dialogue and used these to inform their understanding of plural marking in Mandarin. This result suggests that indirect positive evidence may be an effective tool for real language learners to acquire a grammar that is a subset of their native grammar.

Keywords

indirect positive evidence, subset grammar, inflectional morphology, plurals, Mandarin

1. Introduction

In second language acquisition, when the learner's native language (L1) is a subset of the language being acquired (L2), acquisition can be successful based on direct positive evidence alone: evidence in the primary linguistic data that signals to the learner that a particular construction is well-formed in the language being acquired. For example, the initial clusters in words like *store* and *smile*, if appropriately perceived, indicate to Brazilian Portuguese-speaking learners that *sC* clusters are well-formed in English, pointing to the need to modify the transferred L1 grammar, which lacks this type of complexity. When the learner's L1 is a superset of the L2, acquisition is not always successful, as learners must make use of negative evidence—either direct negative evidence, where learners' attention is explicitly drawn to the ill-formedness of a particular construction in the language being acquired; or indirect negative evidence, where learners infer that a construction is ungrammatical from its absence in the available data.

Schwartz and Goad (2017) explore the possibility that L2 learners of subset grammars may have access to another type of evidence, what they call *indirect* positive evidence. They define this as evidence drawn from errors in the learner's L1 made by native speakers of the learner's L2. To illustrate, if a native speaker of English hears a native speaker of Brazilian Portuguese (BP) repair *sC*-initial words in English via [i] epenthesis (e.g., *I went to a really cool [i]store*), the English speaker may correctly surmise that BP does not permit *sC* clusters in initial position. This type of evidence is indirect because the constraints that are operative in one language are detected through errors made in the other language.

Schwartz and Goad (2017) experimentally examined the efficacy of indirect positive evidence for learning a subset grammar in the syllable structure domain. We build on this work by probing the potential role that this type of evidence may play in another domain: inflectional morphology. Constraints on the marking of inflection are, of course, completely different from those that regulate syllable structure well-formedness, yet the repairs that learners apply when confronted with an L2 that differs from the L1 in either of these domains can overlap, namely, omission of the offending material. This calls for experimental examination of learners' ability to use indirect positive evidence for inflection. Specifically, we test whether naïve English-speaking learners of Mandarin notice plural omission errors in the English productions of Mandarin speakers; and, if so, whether they can appropriately conclude that omission indicates that plural is not (obligatorily) marked in Mandarin nominal constructions.

2. Types of evidence for learning a subset grammar

As mentioned, when the learner's L1 is a superset of the language being learnt, direct positive evidence that a particular structure is ill-formed in the target grammar is not available. Successful acquisition would appear then to rely on some type of negative evidence. If the negative evidence is indirect, learners must infer that a construction is ungrammatical from its absence in the available data. Accordingly, we might expect successful acquisition under this scenario to be uneven. A glance at the L2 literature is consistent with this.

Prawatmuang and Yuan (2020) investigate learners' understanding of constraints on the use of collective/plural markers in Thai and Mandarin. The grammar of Mandarin is a subset of that of Thai: Mandarin *-men* cannot follow animate non-human nouns nor indefinite nouns (see further section 3), unlike Thai *phûak-*. Prawatmuang and Yuan observe that Mandarin-speaking learners of Thai at intermediate levels of proficiency successfully acquire the constraints on Thai *phûak-*, given positive evidence in the input to which they are exposed. Thai-speaking learners of Mandarin, however, must rely on indirect negative evidence and Prawatmuang and Yuan find that,

even at advanced levels of proficiency, learners still struggle with the more restricted conditions that hold of Mandarin *-men*.

The results from Trapman and Kager's (2009) study suggest that indirect negative evidence can sometimes be effective. They examine Russian- and Spanish-speaking learners' acquisition of Dutch phonotactic constraints on onset and coda clusters. Russian is a superset of Dutch, while Spanish is a subset. As expected, the Spanish speakers, for whom positive evidence is available, show sensitivity to the Dutch constraints. Notably, the Russian speakers, for whom there is no positive evidence, also show sensitivity to the Dutch constraints and, in some conditions, outperform the Spanish speakers.

Cross-study differences in the efficacy of indirect negative evidence are no doubt due, in part, to the type of construction under consideration (cf. Saleemi, 1992). For the comparison between Trapman and Kager (2009) and Prawatmuang and Yuan (2020), it is likely that the range of possible consonant clusters in a language is both more frequently present in the input and easier to draw generalisations from than are the contexts for use of collective/plural markers. In the former case, every syllable has the potential to contain a cluster, independent of the category and interpretation of the word that contains it; in the latter, collective/plural markers are expressed on nominals only, and their use is subject to both syntactic/semantic and pragmatic constraints. These differences may lead learners to conclude that data they have not encountered reflect ungrammaticality more readily or perhaps only in the former type of case.

Turning to direct negative evidence, the definition provided above, that learners' attention is explicitly drawn to the ill-formedness of a particular construction in the language being acquired, suggests that learners' awareness of ill-formedness can arise under different conditions: they can be taught that a construction is ill-formed or they can observe this via native speakers' repairs to ill-formed constructions. A body of literature has examined the former situation, with mixed results.

Izumi and Lakshmanan (1998) tested whether targeted instruction can positively impact Japanese learners' understanding of passive constructions in English. English is a subset of Japanese, as it only permits direct passives, while Japanese permits both direct and indirect passives. After being provided with instruction on the ill-formedness of indirect passives in English, Japanese speakers showed considerable improvement in their understanding of the constraints holding of the smaller English grammar, thereby demonstrating that direct negative evidence can be effective for acquisition.

The results from White (1991) are less encouraging. White examined whether French-speaking adolescent learners of English can determine that verb raising does not apply in English, in contrast to in French, through probing adverb placement. Classroom instruction employing negative data resulted in successful learning. However, a follow-up study conducted one year later revealed that the children did not retain what they had learnt, leading White to question whether direct negative evidence impacts competence (see also Schwartz, 1993).

As to whether learners can successfully identify a particular construction as ill-formed by observing the repairs that apply to such constructions, we are not aware of studies that have probed this type of direct negative evidence, but we sketch a scenario here. Consider learners from an L1 like English that permits *sC* clusters faced with acquiring a language like Spanish that does not. For learners of Spanish, direct negative evidence that this type of complexity is ill-formed is available through morphophonemic alternations. Learners can compare verbs like *subscribir* 'to sign, endorse' and *subtitular* 'to subtitle' with their nonprefixed counterparts, *escribir* 'to write' and *titular* 'to title', to arrive at the conclusion that *sub-* is a prefix. Further pairwise comparison should lead learners to conclude that the initial vowel in *escribir* has no morphological function (it is not a prefix), given its corresponding absence from *titular* and *subscribir* and must have been epenthesised due to a ban on *sC* clusters in initial position.

We conjecture that acquisition through exposure to this type of direct negative evidence would reflect a change in competence because the context is more organic: learners arrive at generalisations through their own analysis of the data. Success, though, is not guaranteed. Indeed, for the case at hand, understanding morphophonemic alternation likely requires a high degree of metalinguistic awareness.

From the scenarios we have discussed, it would appear that learning a subset grammar often poses challenges that are not present when learning a superset grammar, regardless of the type of negative evidence employed. We mentioned earlier, however, that Schwartz and Goad (2017) propose that another type of evidence may be available for learning a subset grammar, namely indirect positive evidence. For the construction that they examined, *sC* clusters in Brazilian Portuguese, the evidence should be relatively accessible to learners: English-speaking learners of BP should correctly surmise from BP speakers' productions like [i]*store* for English *store* that BP does not permit *sC* clusters in initial position. We illustrate next that the situation with inflectional morphology is less straightforward.

3. Mandarin plural marking

As mentioned, our goal is to explore the potential value of indirect positive evidence in the domain of inflectional morphology, by focusing on naïve English-speaking learners' understanding of plural marking in Mandarin nominal constructions. In English, plural is obligatorily marked on count nouns in plural contexts, but this is not the case in Mandarin. Focusing on constructions with common nouns, the data in (1a-c) show that *-men* optionally appears with human nouns only; (1c-e) show that there are further conditions on its use: *-men* restricts interpretation to definite reference, (1c), and it cannot cooccur with numerals or measure word expressions, (1d)-(1e) (examples adapted from Po-Ching & Rimmington, 2016).

- | | | | |
|--------|---------------------------------|--|----------------------------------|
| (1) a. | xīn shū
new book | | 'a new book'/'new books' |
| b. | háizi
child | | 'child'/'children' |
| c. | háizi-men
child-PL | | 'the children', *'some children' |
| d. | sān ge háizi
three CL child | *sān ge háizi-men
*three CL child-PL | 'three children' |
| e. | jǐ ge háizi
several CL child | *jǐ ge háizi-men
*several CL child-PL | 'several children' |

There is disagreement in the literature as to whether *-men* is truly a plural marker (Li, 1999) or instead a collective marker (Cheng & Sybesma, 1999; Iljic, 1994), and whether plural morphology is present in classifier languages more generally (Chierchia, 1998). Because *-men* serves a similar function to English *-s*, we follow Lardiere (2009) in assuming that English-speaking learners will initially equate Mandarin *-men* with English *-s*, expressed via transfer of L1 constraints on plural marking. If *-men* is indeed a plural marker, to acquire the target Mandarin grammar, English learners must, at a minimum, come to understand (i) that if a nominal construction is overtly marked with *-men*, it must also be [+human] and [+definite]; and (ii) that

even under the conditions in (i), *-men* is optionally expressed, determined by pragmatic considerations.

3.1 *The learning challenge*

Given the constraints identified in (i)-(ii) above, we first consider what English-speaking learners can conclude about plural marking in Mandarin when exposed to data like the well-formed constructions in (1). We focus on (1a-c). The presence of *-men* in [+human, +definite] contexts like (1c) provides direct positive evidence to learners that plural can be overtly expressed in Mandarin nominal constructions, consistent with the transferred grammar. There is also direct positive evidence that nouns in plural contexts can appear without *-men*, in non-human or indefinite constructions (1a-b), inconsistent with the transferred grammar. The challenge for learners is that they cannot use direct positive evidence to draw conclusions about the ill-formedness of *-men* in these contexts, especially given that the use of *-men* in [+human, +definite] contexts is optional. In the absence of direct negative evidence—corrections to learners' oversuppliance of *-men* in non-human or indefinite contexts—learners might reasonably conclude that they simply have not yet heard nouns with *-men* in plural contexts like those in (1a-b). Clearly, direct positive evidence is not available for English-speaking learners to determine all constraints on plural marking in Mandarin.

3.2 *Indirect positive evidence?*

In view of the limitations of the direct positive evidence available for English-speaking learners of Mandarin plural marking, we question whether learners could employ indirect positive evidence. Specifically, if English-speaking learners converse in English with Mandarin speakers who fail to reliably mark plural in their English productions, could the English speakers use these errors to conclude that Mandarin marks plural in a subset of the conditions under which it is marked in English?

Whether the English speakers can truly use this evidence requires, of course, that the omission errors be noticed. This would seem to be more likely if Mandarin speakers' failure to mark plural in English led to misinterpretation on the part of their English-speaking interlocutors. Misinterpretation, though, is not inevitable, if the pragmatic context is sufficiently rich or if syntactic cues, such as a disambiguating determiner (*two book, some book*), are present. The second question we therefore ask is: if there is a low probability of widespread misinterpretation, does this mean that errors in the Mandarin-speaking learners' productions of English will go undetected by English-speaking interlocutors? We might be tempted to conclude that the answer is no. After all, in English-type languages, where uninflected stems can surface as words, every time inflection is omitted (*books* → *book*), the learner's production will inadvertently collide with the word's uninflected counterpart (*book*). At the same time, though, since the core meaning and lexical category of the word have not changed, errors may be noticed by some, but perhaps not by all, English speakers.

Assuming that at least some English speakers do notice plural omission errors made by Mandarin speakers, the third question we ask is: can these English speakers (who, recall, are striving to learn Mandarin) use the errors to inform their understanding of plural marking in Mandarin? English speakers may fail to connect the behaviour observed with the Mandarin grammar it reflects. Alternatively, they may be unsure how to interpret the behaviour observed, that is, how to determine which part of the Mandarin grammar is implicated (what Pinker (1989: 14) labels as 'blame assignment').

If the environments for plural suppliance and omission in Mandarin speakers' English productions aligned with the constraints in (1) (which we return to in section 7.4), the indirect positive evidence available to English-speaking learners of Mandarin could be particularly instructive in helping learners narrow down the hypothesis space. Lardiere (2007, 2009), however, points out that even when Mandarin learners of English have overcome the restrictions identified in (1), they can still fail to mark plural morphology in all obligatory contexts in their L2 productions. This is because factors other than syntax/semantics and pragmatics can impact suppliance, for example (i) phonological mismatches between the L1 and L2: the absence of necessary segmental or syllable structure complexity at the right edge of Mandarin words (Lardiere, 2009; Young, 1991); and/or (ii) morpho-phonological mismatches between the L1 and L2: differences in the organisation of inflection into higher prosodic structure in the two languages (Goad & White, 2019; Goad, White & Steele, 2003). The observation that omission of plural morphology can arise from mismatches between the L1 (Mandarin) and L2 (English) grammars on multiple levels affirms Slabakova's (2019) assertion that functional morphology is the bottleneck in L2 acquisition.

Clearly, the real world value of indirect positive evidence for the problem at hand—understanding (the absence of) plural marking in Mandarin nominal constructions—is difficult to determine, because constraints on the grammar of Mandarin conspire such that there are many potential sources for plural omission that English-speaking learners of Mandarin must consider in interpreting Mandarin speakers' English productions. In view of this, as our goal in this paper is to examine the potential efficacy of indirect positive evidence for the learning of inflectional morphology, we limit the number of factors that could be at play in the errored data that naïve learners of Mandarin are exposed to, as we discuss more concretely in section 5.2. We first present the research questions that underlie our study.

4. Research questions

The research questions (RQs) that we experimentally examine stem directly from the questions we raised in the preceding section concerning how acquisition of constraints on the marking of Mandarin *-men* might proceed in a real language learning context:

- (2) RQ1: Can indirect positive evidence be effectively used in the domain of inflectional morphology?
- RQ2: Are omission errors made by non-native speakers in the inflectional domain noticed by native speakers?
- RQ3: Can native speakers connect omission errors observed in non-native speakers' productions with the grammar of inflection that the errors reflect?

The overarching question we pose, RQ1, is whether indirect positive evidence can be used when acquiring a subset grammar in the domain of inflectional morphology. We explore this question by comparing the behaviour of two groups of naïve English-speaking learners of Mandarin, one of which, the exposure group, was presented with indirect positive evidence suggesting that Mandarin does not mark plural. RQ2 and RQ3 are examined via answers to follow-up questions addressed to individual participants in the exposure group. If we are able to answer RQ1-RQ3 in the affirmative, this would extend the potential benefit of indirect positive evidence beyond the realm of syllable structure (examined in Schwartz & Goad, 2017).

5. Methods

To probe RQ1-RQ3, participants first listened to a dialogue in English providing indirect positive evidence for the absence of plural marking on Mandarin nouns. The dialogue was designed to mimic a real-world interaction between a Mandarin-speaking learner of English and a native speaker of English. In this way, the methodology builds on that employed by Schwartz and Goad (2017). Participants in their study were also exposed to indirect positive evidence in the form of a dialogue; however, the two interlocutors were both BP-speaking learners of English. In a naturalistic setting, interlocutors who share the same native language are unlikely to converse in their L2, which is why our dialogue involved interlocutors who do not share the same L1. The type of interaction we employed is thus likely to more closely reflect the indirect positive evidence a real language learner would be exposed to.

Subsequent to hearing the dialogue, participants were required to learn the names for several Mandarin-like nouns in singular contexts. They were then tested on their word learning success, using both singular and plural pictures as prompts. If learners are sensitive to the indirect positive evidence they were exposed to in the dialogue, it was expected that they would correctly deduce that the same string to which they had assigned some singular interpretation would also be appropriate in plural contexts, that is, that Mandarin (generally) does not mark plural on nouns.

5.1 Participants

Participants were 43 native speakers of North American varieties of English, 18-35 years in age. All but one reported that they had completed or were currently pursuing post-secondary education. Thirty-nine participants were students; four were professionals. None had any knowledge of or previous exposure to any Chinese language (aside from hearing Chinese spoken now and again in their environment). Other exclusionary criteria were as follows: participants could not have attained higher than an intermediate level of proficiency in any other language; they could not have a hearing deficit; they could not have had chronic ear infections as children; and they could not have taken any courses in Linguistics. As for their language profile, most participants had attained some level of proficiency in French (30 out of 39 participants) and most had studied at least one other language, including French (36 participants). Further details are provided in Appendix A. Participants were recruited via online advertisements. They were tested in person, and were compensated for their time.

Participants were randomly assigned to two experimental groups: an exposure group ($n = 21$; 9 males, 12 females) and a control group ($n = 22$; 10 males, 12 females) (see further section 5.3). One participant from the exposure group was excluded due to the level of proficiency he had attained in Japanese; Japanese is similar to Mandarin in the conditions under which plural morphology, which is optional, is marked (Martin, 1975).

5.2 Materials

The method used to expose participants to indirect positive evidence was a 4-minute dialogue (736 words in length) between a native speaker of Mandarin and a native speaker of English (see Appendix B). Productions were relatively balanced across interlocutors (Mandarin speaker: 359 words; English speaker: 378 words; 19 turns each). In the Mandarin speaker's utterances, there were 24 nominals in plural contexts, all of which were incorrectly produced as their singular counterparts. In contrast, to limit the number of factors that could be responsible for plural omission, the speaker's productions of agreement morphology, *s*-shaped clitics, and [s~z] at the right edge of monomorphemic forms were error free. Participants listened to the dialogue twice

(see section 5.3), for a total of 48 errored productions over 8 minutes, which is similar to the amount of exposure participants were reported to have had in Schwartz and Goad (2017) (48 errored productions over 7 minutes).

The participants were also required to learn 12 pseudo-Mandarin words: existing Mandarin monosyllabic strings, each of which was paired with a singular concrete count noun referent to ensure picturability and facilitate learning.¹ All words were CVV or CVN in shape (C = consonant, V = vowel, N = nasal). Segments were selected to ensure that the words would be distinct enough from each other to minimise perceptual difficulties for the English-speaking participants who, recall, had had no previous exposure to Mandarin. All words were produced with tone 4 (falling tone) so that they would sound natural to anglophones when produced in isolation, a context where falling intonation is observed in English. The words and associated meanings are provided in (3) (transcribed in IPA; a grave accent indicates tone 4):

(3)	[pèi]	‘cup’	[bìn]	‘fan’	[màu]	‘cat’	[sùn]	‘cake’
	[tòu]	‘plum’	[dàn]	‘star’	[màn]	‘hat’	[lín]	‘tree’
	[kài]	‘car’	[gòu]	‘book’	[nèi]	‘bag’	[lùn]	‘mop’

The English counterparts of the nouns in (3) did not appear in the productions of the Mandarin speaker in the dialogue, in neither singular nor plural contexts. Thus, participants should not have made a connection between the dialogue and the word learning task at the point when they were learning the sound–meaning pairings. Any connection would be made later, if at all, when participants were tested on their ability to use the indirect positive evidence that they were exposed to in the dialogue in order to generalise the names for the singular items in (3) to plural contexts. We provide further details on this in the Procedure section.

The dialogue and pseudo-Mandarin words were recorded in a sound-attenuated booth, using a Roland-05 recorder (settings: 16-bit mono channel, 44.1 KHz). The voices for the dialogue were a male native speaker of English and a female native speaker of Mandarin. The pseudo-Mandarin words were recorded by the same female speaker whose voice was used in the dialogue. All recordings were saved in wav format and later transferred to computer.

5.3 Procedure

The experiment involved four activities: listening to the dialogue; training on Mandarin singular nouns; testing on Mandarin singular nouns and their plural counterparts; and answering follow-up questions. Participants in the exposure group were involved in all four activities, whereas participants in the control group were involved in three: only the exposure group listened to the dialogue and, thus, only this group received indirect positive evidence about the absence of plural marking in Mandarin. Participants in both groups received identical training on the 12 pseudo-Mandarin words and were involved in the same testing sessions on these words. All participants answered questions about their responses to the last testing session, although the questions differed somewhat for the two groups (see below).

The experiment was conducted on a MacBook Pro laptop and took place in a sound-attenuated booth at McGill University. All participants wore headphones (Beats ML992LL/A). The training and testing activities were presented using PsychoPy v3.0 (Peirce et al., 2019). After giving informed consent, participants filled out a questionnaire, which solicited basic demographic and language background information. They then proceeded to the experiment.

¹ Participants were told that the words were actual words of Mandarin. At the end of the experiment, they were informed that this was not in fact the case, that the sound-meaning pairings had been fabricated.

The experiment took place over two consecutive days, for approximately one hour per day (see Table 1). The two days were structured similarly in terms of the activities that participants were engaged in and the order in which tasks were conducted within these activities. The difference between the two days concerned the nature of the final task: maxi test 1 vs. maxi test 2, detailed below.

Table 1. Schedule for experiment: exposure and control groups.

	Exposure group	Control group
Day 1	Step 1: listen to dialogue	Step 1: mini training 1
	Step 2: mini training 1	Step 2: mini test 1
	Step 3: mini test 1	Step 3: mini training 2
	Step 4: mini training 2	Step 4: mini test 2
	Step 5: mini test 2	Step 5: mini training 3
	Step 6: mini training 3	Step 6: mini test 1
	Step 7: mini test 1	Step 7: maxi test 1
	Step 8: maxi test 1	
Day 2	Repeat steps 1-7 from Day 1	Repeat steps 1-6 from Day 1
	Step 8: maxi test 2	Step 7: maxi test 2
	Step 9: follow-up questions	Step 8: follow-up questions

Participants in the exposure group first listened to the dialogue. Training and testing on the 12 pseudo-Mandarin words followed for both groups. To optimise learning, the 12 words were divided into three training sessions (mini training 1-3), each of which focused on four words. In a given mini training session, each auditory stimulus and its corresponding picture were presented three times, yielding a total of 12 trials per session. The stimuli were presented in random order using PsychoPy. Each stimulus was separated by an interval of 350ms.

Each mini training was immediately followed by a mini test on the four words previously taught. Each word occurred twice, in two different trials, yielding a total of eight trials per mini test. Words were presented in random order. In each trial, a single auditory stimulus was heard (using the same audio recording from the training session). Its corresponding picture and another picture randomly selected from the other three in the same session occurred synchronously on the laptop screen, as exemplified in Figure 1. Participants were asked to listen to the word and select its corresponding picture by pressing LEFT (referring to the left-hand picture) or RIGHT (referring to the right-hand picture), as per labels attached to keys on the computer keyboard. Their response triggered a 350ms interval of silence, after which the next trial began. Feedback on participants' selections was provided immediately after a response was made: a high-pitched *ding* indicated a correct response, while a low-pitched *buzz* referred to an incorrect response. The correct answers (LEFT or RIGHT) were counterbalanced.

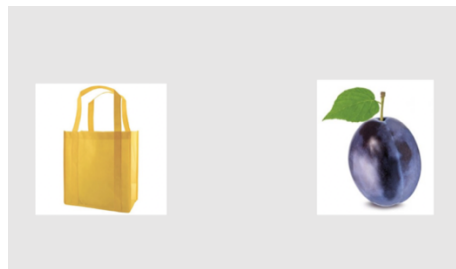


Figure 1. Sample trial in mini test, for target word [nèi] ‘bag’.

In maxi test 1, at the end of Day 1, all 12 words were tested with the same method used in the mini tests. Each word occurred twice in two different trials, yielding 24 trials presented in random order. Feedback was provided, as in the mini tests.

Day 2 repeated the same steps from Day 1 to ensure that participants remembered the words they had previously been trained on. For the exposure group, this was preceded by participants once again listening to the dialogue. At the end of Day 2, participants completed maxi test 2, which was designed to probe whether participants in the exposure group could use the indirect positive evidence they were presented with in the dialogue to generalise the names for the singular items on which they were trained to plural contexts. Importantly, participants were not given feedback in maxi test 2; thus, we also used this test to determine how well they had learnt the names for the pictures, that is, in singular contexts.

In maxi test 2, participants were asked to listen to a word and select a corresponding answer among three choices: two pictures showing a single object or multiple objects displayed on the left and middle of the screen, and NEITHER displayed on the right side of the screen to be selected if neither the left nor middle object was the correct answer; see Figure 2. The plural pictures did not involve two identical copies of the corresponding singular picture (as is common in wug tests designed for children), because pilot testing suggested that this increased participants' decision-making based on analogy.

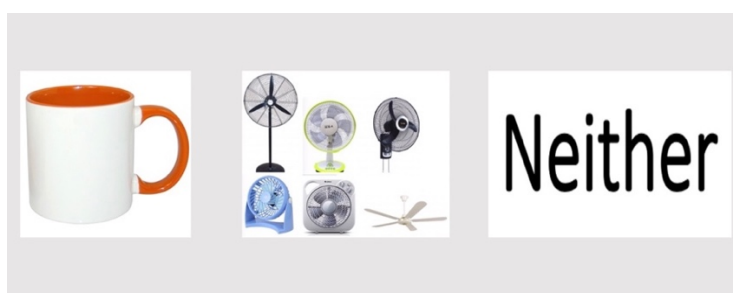


Figure 2. Sample trial in maxi test 2 for target word [bɪn] ‘fan’.

In addition to the 12 words that participants had been trained on, pictures of six new objects that had not appeared in the mini training sessions, mini tests, or maxi test 1 were added to maxi test 2. The pictures were of objects that are semantically similar to six of the words that participants were trained on: *bowl* (similar to *cup*), *moon* (similar to *star*), *broom* (similar to *mop*), *air conditioner* (similar to *fan*), *grape* (similar to *plum*), and *truck* (similar to *car*). These pictures were introduced to try to disguise the purpose of the task: we hoped that their inclusion would detract away from participants noticing that all training words were paired with their plural counterparts in testing. Each of the new objects occurred twice; one picture displayed a single object and one displayed multiple objects.

Maxi test 2 included a total of 96 trials: 24 trials where the correct answer was a singular picture from the learnt words; 24 trials where the correct answer was a plural picture from the learnt words; and 48 distractor trials, where the correct answer was NEITHER. For 24 of the distractors, the pictures displayed corresponded to the learnt words; for the remaining 24, one picture was of a learnt word and the other was of a semantically similar word (see preceding paragraph). To illustrate, consider the word [pɛi] ‘cup’ in Figure 3. Trials 1 and 5 served to test whether participants had learnt the Mandarin words from the mini training sessions; trials 2 and 6 examined whether indirect positive evidence in the form of plural errors in the dialogue was used to make correct responses (for the exposure group); and trials 3, 4, 7 and 8 were distractors, where

for trials 4 and 7, singular and plural pictures of *bowl* substituted for singular and plural pictures of *cup*.² Stimuli were presented in random order.

















Trial	Word	Choices			Correct answer	Testing target
		LEFT	MIDDLE	NEITHER		
1	[pèi] ‘cup’			neither	LEFT	singular
2	[pèi] ‘cup’			neither	LEFT	plural
3	[bìn] ‘fan’			neither	NEITHER	distractor
4	[pèi] ‘cup’			neither	NEITHER	distractor
5	[pèi] ‘cup’			neither	MIDDLE	singular
6	[pèi] ‘cup’			neither	MIDDLE	plural
7	[màn] ‘hat’			neither	NEITHER	distractor
8	[màn] ‘hat’			neither	NEITHER	distractor

Figure 3. Test trials for the target word [pèi] ‘cup’ in maxi test 2.

After maxi test 2 was complete, participants were asked some follow-up questions, the answers to which were used to help interpret their results. Participants in the exposure group were asked questions 1-3 below; participants in the control group were asked question 3 only.

² No participant incorrectly selected pictures of the semantically similar words as their response. Thus, we do not discuss these words further.

1. Did you notice anything unusual or anything wrong with the dialogue; if yes, what did you notice?
2. Did you see any connection between the dialogue and the training/testing sessions; if yes, what was the connection?
3. In the final testing session, were there any cases when a word was heard, such as [pèi] meaning ‘cup’, where you selected the plural picture (*cups*) as the correct answer; if yes, why?

6. Results

In this section, we report on the data collected from maxi test 2. Individual participants’ accuracy rates in responding to singular and plural test items were first calculated. To ensure that poor performance on the plural stimuli was not due to deficiencies in learning the names for the singular objects on which participants were trained, minimum performance on the singulars was set at 21/24 (87.5%) correct. The probability that a participant could randomly select the correct answer among three equally likely options 21 times out of 24 is less than 0.0001%. Two participants from the control group and one from the exposure group did not meet this criterion and so were excluded from further analysis.

6.1 Results by group

The results by group for the remaining 39 participants (control group: $n = 19$; exposure group: $n = 20$) are provided in Table 2.

Table 2. Performance of the control and exposure groups on maxi test 2.

Test items	Group	Minimum score (%)	Maximum score (%)	Mean (%)	S.E.
singulars	control	87.5	100	96.3	0.95
	exposure	87.5	100	95.4	0.94
plurals	control	0	91.7	20	7.16
	exposure	0	95.8	55.5	8.54

As can be inferred from the table, most participants in both groups achieved high accuracy rates on the singulars: both groups had a minimum score of 87.5% with similar means: 96.3% for the control group and 95.4% for the exposure group. The situation is markedly different for the plurals: performance was overall much lower, and for some participants in both groups, it was as low as 0%. Although the minimum and maximum scores provided may suggest that the two groups performed similarly, the means are considerably different: 20% and 55.5%, respectively.

A two-way ANOVA was conducted with individual participants’ accuracies on the singular and plural items as dependent variables, while group (control vs. exposure) and test item (singular vs. plural) were coded as fixed factors. There were significant effects of group ($F(1,2) = 8.31$, $p = 0.005$, $\eta_p^2 = 0.1$) and test item ($F(1,39) = 108.55$, $p < 0.001$, $\eta_p^2 = 0.60$), as well as a significant interaction between group and test item ($F(1,2) = 9.66$, $p = 0.003$, $\eta_p^2 = 0.12$) on individual participants’ accuracy rates. Two follow-up t-tests were carried out to compare the control and exposure groups’ performance on the singular and plural test items in maxi test 2. It was found

that the exposure group's accuracy rate differed significantly from that of the control group on the plural items ($t(58) = 6.36, p = 0.003$), but not on the singular items ($t(58) = 2.03, p = 0.14$).

6.2 Results by individual

Figures 4 and 5 provide individual results on the singular and plural stimuli for the control and exposure groups, respectively. Participants in each group are ordered according to their performance on the plurals. For participants in the control group, who were not exposed to the dialogue, we would expect to find near 0% correct on the plurals. Figure 4 shows that this holds for 16 of the 19 participants (C10 to C12, who got 0%-16.7% correct). There are, though, three participants (C5, C4, C15) whose performance is 87.5% or higher. Participants thus follow two markedly different patterns of behaviour.

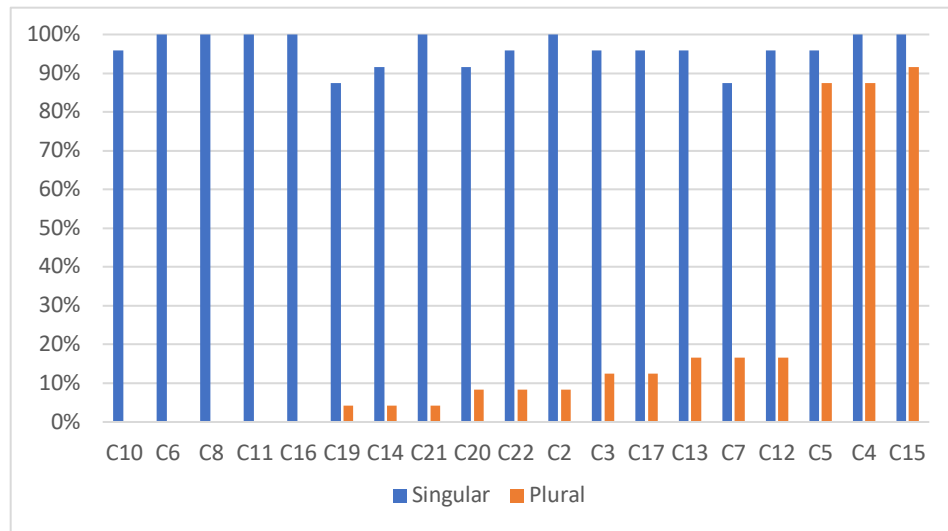


Figure 4. Results by individual for the control group.

Recall that participants in the control group were asked the following question after they completed maxi test 2: In the final testing session, were there any cases when a word was heard, such as [pèi] meaning ‘cup’, where you selected the plural picture (*cups*) as the correct answer; if yes, why? The 11 participants in the middle of Figure 4 (C19 to C12), who got 4.2%-16.7% of plurals correct, responded that their plural responses were random guesses. In contrast, participants C5, C4 and C15, who got a minimum of 87.5% of plurals correct, responded that their plural responses were based on analogy with the singulars they had learnt.

Figure 5 provides the results for the exposure group. Recall that participants in this group listened to the dialogue and thus they could, in principle, take advantage of the indirect positive evidence it provided and generalise the names for the singular items that they were trained on to plural contexts. It is conceivable, then, that these participants would be near ceiling on the plurals, parallel to their performance on the singulars. Like the control group, however, participants in the exposure group showed two distinct patterns of behaviour: ten participants were correct on the plurals 0%-25% of the time, while the other ten were correct 75-95.8% of the time.

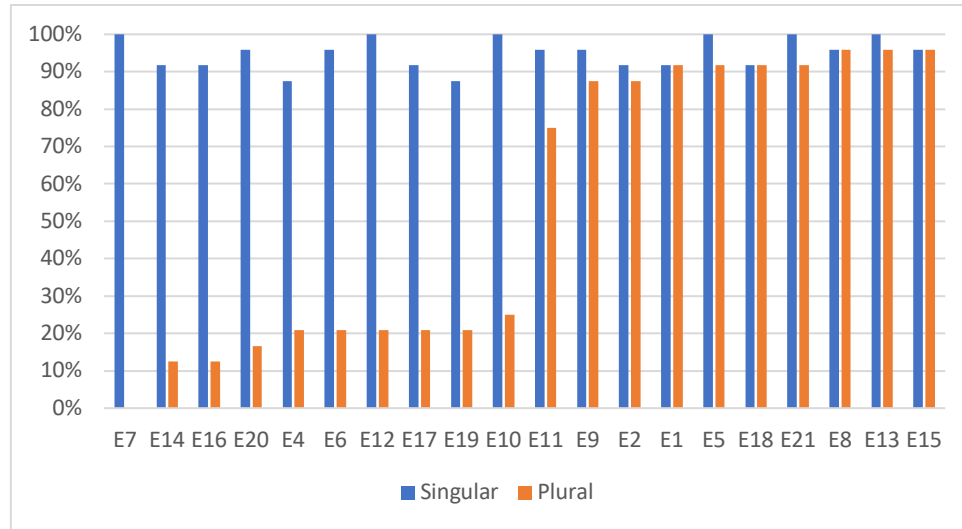


Figure 5. Results by individual for the exposure group.

Since the three learners in the control group who performed well on the plurals reported using analogy, we must assess whether the ten high performing participants in the exposure group similarly used analogy or whether they instead used the indirect positive evidence available in the dialogue. To tease these options apart, we examine participants' responses to the first two questions they were asked after they completed maxi test 2: Did you notice anything unusual or anything wrong with the dialogue; if yes, what did you notice? Did you see any connection between the dialogue and the training/testing sessions; if yes, what was the connection?

We begin with the ten participants in the exposure group whose performance ranged from 0% to 25%. Five of these participants (E7, E14, E4, E12, E17) reported that they failed to notice the plural mistakes in the dialogue and therefore did not see a connection between the dialogue and the training/testing sessions. The other five (E16, E20, E6, E19, E10) said they noticed several plural mistakes in the dialogue but did not see any connection between these mistakes and the training/testing sessions. Clearly, these ten participants did not use the indirect positive evidence available. Not surprisingly, then, in response to the final question (In the final testing session, were there any cases when a word was heard, such as [pèi] meaning 'cup', where you selected the plural picture (*cups*) as the correct answer; if yes, why?), these participants, like the 11 controls in the middle of Figure 4, responded that their plural responses were random guesses.

Let us now turn to the ten participants whose performance on the plurals was at 75% or higher. In response to the first question, all participants reported that they noticed the plural errors in the dialogue. Two participants (E11, E9), though, indicated that they failed to see a connection between these mistakes and the training/testing sessions. In response to the third question, they reported that their correct responses were based on analogy. The eight remaining participants (E2, E1, E5, E18, E21, E8, E13, E15) mentioned that they saw the connection between the dialogue and the training/testing sessions. In response to the third question, they reported that they concluded from the errors they observed in the dialogue that Mandarin does not mark plural and they used this to generalise the names for the singular items they were trained on to plural contexts, thereby using the indirect positive evidence available.

7. Discussion

7.1 Research questions revisited

Our first research question, RQ1, asked whether indirect positive evidence can be effectively used in the domain of inflectional morphology. The results of the statistical tests indicate that we can answer this question in the affirmative: the ANOVA revealed a significant interaction between group and test item and follow-up t-tests revealed that the exposure group outperformed the control group on the plural items only.

However, when we examine individuals' performance in Figure 5 and consider their answers to the follow-up questions, it is evident that the participants in the exposure group are not behaving uniformly. We further interpret their behaviour by returning to our other research questions. RQ2 asked whether omission errors made by non-native speakers in the inflectional domain are noticed by native speakers. We can answer this question positively for 15 of the 20 participants in the exposure group. These participants noticed plural mistakes in the dialogue. However, noticing the mistakes is only the first step toward being able to use indirect positive evidence, which leads us to RQ3.

RQ3 asked whether native speakers can connect omission errors observed in non-native speakers' productions with the grammar of inflection that the errors reflect. We can answer this question affirmatively for eight out of the 15 participants who noticed the plural mistakes in the dialogue. These participants successfully used the indirect positive evidence available to deduce that plural is not (obligatorily) marked in Mandarin nominal constructions. Specifically, these learners correctly concluded that the same string to which they had assigned a singular interpretation in the word learning task was appropriate for use in plural contexts, evidenced through their performance on maxi test 2 and their answers to the follow-up questions we posed.

The seven remaining participants who noticed plural errors in the dialogue did not connect these errors with the grammar of Mandarin. Five of the participants scored quite poorly on the plural items and reported that their answers were random guesses. Clearly, they did not generalise a regular pattern of plural omission; indeed, they may have failed to connect all instances of omission specifically to plurality. If this is the case, they would not have been able to connect the word learning task to the dialogue via their answers to maxi test 2. As for the two remaining participants, both scored well on the plural items and reported that their answers were due to analogy. It is highly unlikely that these participants failed to connect all instances of omission to plurality. In fact, it is possible that these learners used the indirect positive evidence available, but are not aware of this.

7.2 Individual variation

We have seen that participants in the exposure group arrive at different conclusions about aspects of the data they were presented with, but we have yet to consider why some learners in this group were successful whereas others fell short; that is, what could potentially underlie the variation observed across individuals. Recall from Figure 5 that the variation is considerable: ten participants in the exposure group scored from 0% to 25% on the plurals while ten others scored between 75% and 95.83%. The former group either provided no plural responses ($n = 1$) or reported that their plural responses were random guesses ($n = 9$), while the latter group indicated that they used analogy ($n = 2$) or the indirect positive evidence available ($n = 8$). For the discussion that follows, we will compare these two groups with each other, referring to the former as the low performing group and the latter as the high performing group. We have included the two individuals who claimed that they used analogy in the high performing group based on their scores,

even though they either did not use, or are not aware that they used, the indirect positive evidence available in the dialogue.

Two possible explanations for the individual differences observed that can likely be dispensed with are poor learner attention and cognitive fatigue. Concerning learner attention, recall that all participants successfully learnt the names for the singular items, indicating that they were on task. Concerning cognitive fatigue, participants were exposed to the dialogue as the first activity on each testing day when they were fresh (see Table 1 above). Further, test items that solicited a plural response or a singular response were integrated into the same task, so fatigue could not have selectively impacted performance on the plurals.

A more plausible explanation of individual variation lies in differences in participants' metalinguistic awareness. Previous experimental studies have shown that the best predictor of successful language learning in adults is metalinguistic awareness (e.g., Brooks & Kempe, 2013). We did not include a measure of metalinguistic awareness in our study, but a proxy for this could perhaps be found in differences in participants' language backgrounds: a richer language experience could augment metalinguistic awareness or learners with high metalinguistic awareness could be drawn to language learning.

Details on the language profiles of individual participants are provided in Appendix A. In the exposure group, all participants aside from two had studied one or more other languages so this alone cannot be a determining factor. We consider instead (i) exposure to and proficiency attained in French, and (ii) the number of languages studied. Concerning (i), we point out that although plural marking is orthographically expressed in French, it is rarely overtly produced on nouns (see further section 7.4), which could give learners who have studied French an advantage over those who have not when faced with plural omission errors in the productions of Mandarin speakers. This, however, is unlikely to be a relevant consideration, as the two groups look similar: of the ten low performing participants in the exposure group, eight have studied French, four of whom have reached an intermediate level of proficiency (the highest level of proficiency possible for inclusion in the study); of the ten high performing participants, seven have studied French, five of whom have achieved an intermediate level of proficiency.

Concerning (ii), the number of languages that participants in the exposure group have studied, the findings are more promising. Only two of the ten low performing participants have studied more than one additional language (and only to lower intermediate levels of proficiency), while seven of the ten high performing participants have. We do not know whether experience with multiple languages sensitises participants to learners' errors and helps them draw inferences based on this. It is clear, though, that future experimental work exploring the efficacy of indirect positive evidence should include a measure of metalinguistic awareness.

7.3 Probing participants' high rate of success

As we have discussed, we cannot determine with certainty what is responsible for individual participants' success in using the indirect positive evidence available. Nevertheless, we have seen that 40% of participants (8/20) in the exposure group appropriately connected plural omission errors in the dialogue to the absence of plural marking in Mandarin, which stands in contrast to the 32% of participants (11/32) in Schwartz and Goad (2017) who used indirect positive evidence to infer that *sC* clusters are ill-formed in BP. Although the two studies cannot be directly compared, the higher success rate of participants in the current study appears surprising, given the type of error involved and the position of the error in the word, which we detail as follows.

Because the type of error observed in the productions of Mandarin-speaking learners of English plural is different from that observed in the productions of BP-speaking learners of English *sC* clusters (omission vs. epenthesis), the type of indirect positive evidence available to native

English-speaking interlocutors is different. In the latter case, where Schwartz and Goad (2017) found that indirect positive evidence was utilised by some English speakers, the evidence available came in the form of insertion of segmental material in relation to native speaker productions, namely, vowel epenthesis (e.g., *store* → [i]store). Epenthesis typically signals a problem at the level of syllable structure and thus the hypothesis space of the learner who is trying to understand the source of the error is likely narrowed.

In the case under present investigation, the indirect positive evidence available to English speakers comes in the form of absence of morphological material relative to native speaker productions, namely, -s omission (*books* → *book*). Indeed, difficulties that L2 learners have in the inflectional domain rarely involve morphological insertion: inflection is seldom oversupplied; rather, it is undersupplied in obligatory contexts (Lardiere, 1998; Prévost & White, 2000), even when the L2 is more restrictive than the L1 (see, e.g., Hwang & Lardiere, 2013 on English-speaking learners' production of Korean intrinsic plural -*tul*).³ The challenge when the error involves omission is that there are several components of the grammar that could be implicated, as discussed in section 3.2, making such errors more difficult for learners to interpret. In our study, however, omission of [s~z] was limited to contextually plural contexts: the Mandarin speaker's productions of agreement morphology, s-shaped clitics, and [s~z] at the end of monomorphemic words were produced as error free (section 5.2). We have thereby reduced the hypothesis space for the naïve learner considerably, in contrast to the challenge that no doubt holds for real language learners faced with interpreting omission errors. We return to this issue in section 7.4.

Another difference between Schwartz and Goad (2017) and the current study concerns the edge of the word where errors are present. For BP-speaking learners of English, initial sC clusters are repaired via prothesis: epenthesis before the offending cluster (*store* → [i]store). This kind of repair can impede word recognition (Marslen-Wilson, 1984; Marslen-Wilson and Zwitserlood, 1989) and, thus, errors of this sort should be more noticable to English-speaking interlocutors than errors present at the right edge of words. Indeed, material at the right edge tends to erode across languages, and lenition or deletion can target segments in monomorphemic words as well as segments that have a morphological function. For example, in Andalusian, Caribbean and coastal varieties of Latin American Spanish, the trajectory *s* → *h* → ∅ is observed in final position (Alcina Franch and Blecua, 1975; Lipski, 1994), in both monomorphemic (e.g., /brindis/ → [brindis] ~ [brindih] ~ [brindi] 'toast, party') and plural contexts (e.g., /baka-s/ → [bakas] ~ [bakah] ~ [baka] 'cows'; examples from Miller & Schmitt, 2012).

An explanation for the success of our participants, though, may lie in the prosodic profile of many of our stimuli. Omission of plural morphology occurred from nouns with final stressed syllables in 15/24 cases (63%), and from nouns at the right edges of phonological or intonational phrases in 11/24 cases (46%). Both of these contexts are prominent: Cole and Jakimik (1976) observe that pronunciation errors are detected more readily in stressed syllables than in unstressed syllables; and lengthening in phrase-final position is observed in many languages, including English (Wightman, Shattuck-Hufnagel, Ostendorf & Price, 1992) and Mandarin (Yang & Wang, 2002). In short, these prosodic factors could have facilitated listeners' noticing plural omission errors, more so perhaps than in naturalistic language learning settings.

³ Jia (2003) reports that child L2 Mandarin-speaking learners of English oversupply plural on count nouns in singular contexts. From the way the data are reported, it is not possible to determine exactly how many errors were of this type, but Jia notes (p.c., 13 June 2020) that they were comparatively rare. McIlwain and Peterson (2005) report a handful or such errors in two French-speaking 'post-basic' learners of English (1% and 7% of target regular plurals). As mentioned above (see also section 7.4), orthographically-expressed plural marking on French nominals is rarely overtly produced.

7.4 Indirect positive evidence and real language learning

A number of questions arise from our study about the efficacy of indirect positive evidence for real language learning. In our experiment, omission of plural morphology was categorical: as mentioned previously, participants were exposed to errored data in 100% of contexts where plural would be obligatorily marked in English. In real language learning contexts, however, the productions of L2 speakers are variable. For the case at hand, the challenge for the English-speaking learner of Mandarin is to try to identify the syntactic/semantic conditions under which plural is and is not marked in Mandarin using Mandarin speakers' variable productions of English plural. If the contexts of plural suppliance and omission in speakers' productions broadly aligned with the constraints in (1), the indirect positive evidence available to learners of Mandarin could be particularly informative. Results available from the literature, however, suggest that this is not necessarily the case. Young (1991) observes that definite DPs do not favour plural marking, even though *-men* restricts interpretation to definite reference; animacy inhibits plural marking for low proficiency learners, which is unexpected from the marking of *-men* on human nouns only; and the presence of redundant plural marking, as in plural constructions containing numerals or measure word expressions, which disallow *-men*, promotes plural marking, in contrast to DPs where there is no other indication of plurality.⁴ Further, the English speaker must be able to isolate other sources of variation in plural suppliance in the English productions of Mandarin speakers that could obscure the conditions in (1), phonological factors as well as non-linguistic factors (e.g., cognitive load and speaker-specific attributes like proficiency level).

As we have mentioned, the challenge with Mandarin plural marking is very different from that of *sC* cluster well-formedness discussed earlier: only a mismatch between the L1 and L2 at the level of syllable structure could be responsible for the epenthesis pattern observed (*store* → [i]store), although variable epenthesis tied to related constraints (e.g., what type of consonant follows *s*), as well as to non-linguistic and speaker-specific factors could, of course, still play an obfuscating role.

The comparison we have made between plural marking in Mandarin and *sC* clusters in BP might imply that we assume that indirect positive evidence is likely to be more effective in the phonological domain than in the inflectional domain. This might often be the case, because phonological constraints always have the potential to impact the realisation of morphology (see Cabrelli Amaro, Campos-Dintrans & Rothman, 2018). Accordingly, the value of indirect positive evidence for the acquisition of inflection may depend on how similar the segmental and syllable structure constraints are for the two languages under consideration.

To illustrate, we can compare English-speaking learners of plural marking in Mandarin with English-speaking learners of plural marking in Hexagonal French. French plural is orthographically suffixed (*-s*) to nouns, as well as to elements in the DP that agree with the noun. However, it is not produced at the right edge of the words that morphologically host it (e.g., *les tricots bleus* [le triko blø], *[les trikos bløs] 'the blue sweaters'). Plural marking from a determiner is obligatorily realised (as [z]) on a following vowel-initial noun or adjective, that is, where it can be syllabified in onset position (*les amis* [le zami] 'the friends', *les anciens professeurs* [le zãsje pɔfɛsœʁ] 'the former teachers'). Plural marking from a noun is optionally realised in the same segmental context, more commonly in formal speech (e.g., *les amis intelligents* [le zami (z)êteližã] 'the intelligent friends'). One conclusion we can draw from these examples is that plural *-s* is never pronounced at the right edge of a DP in French. This is surprising, as both [s] and [z] are well-

⁴ The latter has also been found for Korean- and Japanese-speaking learners of English (Austin, 2014; Mellow & Cumming, 1994), two languages with similar constraints on plural marking as hold in Mandarin (Song, 2005; Martin, 1975).

formed at the right edge of words in non-plural contexts (e.g., *lisse* [lis] ‘smooth’, *lisent* [liz] ‘read’-3PL, *Lise* [liz] (name)).

Turning to acquisition, plural marking is often omitted in the English of French-speaking learners (e.g., McIlwain & Peterson, 2005), due to negative transfer: plural in English must be realised on the noun, which appears in DP-final position (e.g., *the blue sweaters*). French speakers, though, face no challenges with [s] and [z] at the right edge of monomorphemic words (e.g., *loose*, *lose*). Thus, after some exposure to French, English-speaking learners who try to use indirect positive evidence to tap into constraints on the distribution of French plural are unlikely to consider that plural omission errors in French speakers’ English productions are due to segmental or syllable structure constraints, unlike in the case of English-speaking learners of Mandarin. In short, since English and French are similar in the phonological profiles they permit at the right edge of words, the hypothesis space is considerably narrowed for the English-speaking learner who is trying to determine what underlies -s omission in French speakers’ English.

8. Conclusion

In this paper, we have questioned whether second language learners can employ indirect positive evidence to acquire a subset grammar in the domain of inflectional morphology: evidence drawn from errors in the learner’s L1 made by native speakers of the learner’s L2. We experimentally probed this question by testing whether naïve English-speaking learners of Mandarin can deduce that plural is not obligatorily marked in Mandarin using plural omission errors in the English productions of Mandarin speakers. The results revealed a significant difference between the two groups of participants: those in the exposure group, who were presented with plural omission errors via a dialogue, outperformed those in the control group on the plural items. Examination of the exposure group’s responses to the follow-up questions indicated that 40% of participants in this group used the indirect positive evidence available. Specifically, these individuals both noticed the plural omission errors in the dialogue and appropriately used this information to generalise the names for the singular items they were trained on to plural contexts.

Based on this finding, we conjecture that indirect positive evidence may be an effective tool in real language learning contexts. Indeed, including accented input in the classroom could be pedagogically advantageous. For uninstructed learners, though, how effective this type of evidence is may depend on whether the hypothesis space can be sufficiently narrowed such that the interpretation of errors can be confined to the appropriate area of the grammar. English-speaking learners of BP may be able to attribute epenthesis errors in BP speakers’ English productions of sC clusters to an L1–L2 mismatch at the level of syllable structure. English-speaking learners of French may be able to set aside the possibility of L1–L2 mismatches at the level of segment and syllable structure to interpret plural omission errors in French speakers’ English productions. English-speaking learners of Mandarin, however, will likely face more of a challenge interpreting plural omission errors in Mandarin speakers’ English productions, as these errors can arise from multiple types of mismatches between the L1 and L2 grammars, as we have discussed. Nevertheless, if learners have acquired sufficient knowledge of the grammar of Mandarin through direct positive evidence, they may be able to appropriately attribute omission errors in Mandarin speakers’ English productions to their various sources. We leave exploration of the efficacy of indirect positive evidence for real language learning to future research.

Acknowledgements

We would like to thank the reviewers, Associate Editor Ian Cunnings, and the audience at BUCLD 46 for helpful comments, as well as Jeff Lamontagne for recording the English stimuli. We would also like to thank the following funding sources: China Scholarship Council, MITACS Globalink Early Career Fellowship, Social Sciences and Humanities Research Council of Canada, and Fonds de recherche du Québec - Société et culture.

References

- Alcina Franch, J., & Blecua, J. M. (1975). *Gramática española*. Barcelona: Ariel.
- Austin, G. (2014). *Asymmetries in L2 functional morphology: The Korean speaker of English*. PhD dissertation, University of New England, Australia.
- Brooks, P. J., & Kempe, V. (2013). Individual differences in adult foreign language learning: The mediating effect of metalinguistic awareness. *Memory & Cognition*, 41(2), 281–296. <https://doi.org/10.3758/s13421-012-0262-9>
- Cabrelli Amaro, J., Campos-Dintrans, G., & Rothman, J. (2018). The role of L1 phonology in L2 morphological production: L2 English past tense production by L1 Spanish, Mandarin, and Japanese speakers. *Studies in Second Language Acquisition*, 40, 503–527. <https://doi.org/10.1017/S0272263117000122>
- Cheng, L. L.-S., & Sybesma, R. (1999). Bare and not-so-bare nouns and the structure of NP. *Linguistic Inquiry* 30, 509–542. <https://doi-org.proxy3.library.mcgill.ca/10.1162/002438999554192>
- Chierchia, G. (1998). Reference to kinds across languages. *Natural Language Semantics*, 6, 339–405. <https://doi-org.proxy3.library.mcgill.ca/10.1023/A:1008324218506>
- Cole, R. A., & Jakimik, J. (1979). How are syllables used to recognize words? *Journal of the Acoustical Society of America*, 67(3), 965–970. <https://doi.org/10.1121/1.383939>
- Goad, H. & White, L. (2019). Prosodic effects on L2 grammars. *Linguistic Approaches to Bilingualism*, 9, 769–808. <https://doi.org/10.1075/lab.19043.goa>
- Goad, H., White, L., & Steele, J. (2003). Missing inflection in L2 acquisition: Defective syntax or L1-constrained prosodic representations? *Canadian Journal of Linguistics*, 48, 243–263. <https://doi.org/10.1017/S0008413100000669>
- Hancin-Bhatt, B. (2000). Optimality in second language phonology: Coda in Thai ESL. *Second Language Research*, 16(3), 201–232. <https://doi.org/10.1191/026765800671362605>
- Hwang, S. H., & Lardiere, D. (2013). Plural-marking in L2 Korean: A feature-based approach. *Second Language Research*, 29(1), 57–86. <https://doi.org/10.1177/0267658312461496>
- Iljic, R. (1994). Quantification in Mandarin Chinese: Two markers of plurality. *Linguistics*, 32, 91–116. <https://doi-org.proxy3.library.mcgill.ca/10.1515/ling.1994.32.1.91>
- Izumi, S., & Lakshmanan, U. (1998). Learnability, negative evidence and the L2 acquisition of the English passive. *Second Language Research*, 14, 62–101. <https://doi.org/10.1191/026765898675700455>
- Jia, G. (2003). The acquisition of the English plural morpheme by native Mandarin Chinese-speaking children. *Journal of Speech, Language, and Hearing Research*, 46(6), 1297–1311. [https://doi.org/10.1044/1092-4388\(2003/101\)](https://doi.org/10.1044/1092-4388(2003/101))
- Lardiere, D. (1998). Case and Tense in the ‘fossilized’ steady state. *Second Language Research*, 14(1), 1–26. <https://doi.org/10.1191/026765898674105303>
- Lardiere, D. (2007). *Ultimate attainment in second language acquisition: A case study*. Mahwah, NJ: Lawrence Erlbaum Associates.

- 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>
- Li, Y.-H. A. (1999). Plurality in a classifier language. *Journal of East Asian Linguistics*, 8, 75–99.
<https://doi.org/10.1023/A:1008306431442>
- Lipski, J. (1994). *Latin American Spanish*. London: Longman.
- Marslen-Wilson, W. D. (1984). Function and process in spoken word recognition. In H. Bouma & D. G. Bouwhuis (Eds.), *Attention and performance X: Control of language processes* (pp. 125–150). Hillsdale, NJ: Erlbaum.
- Marslen-Wilson, W. D., & Welsh, A. (1978). Processing interactions and lexical access during word recognition in continuous speech. *Cognitive Psychology*, 10, 29–63.
[https://doi.org/10.1016/0010-0285\(78\)90018-X](https://doi.org/10.1016/0010-0285(78)90018-X)
- Marslen-Wilson, W. D., & Zwitserlood, P. (1989). Accessing spoken words: The importance of word onsets. *Human Perception and Performance*, 15(3), 576–585.
<https://doi.org/10.1037/0096-1523.15.3.576>
- Martin, S. (1975). *A reference grammar of Japanese*. New Haven, CT: Yale University Press.
- McIlwain, J., & Peterson, P. (2005). Signaling plurality in learner English. In I. Mushin (Ed.), *Proceedings of the 2004 Conference of the Australian Linguistic Society*. Sydney: University of Sydney. <https://ses.library.usyd.edu.au/handle/2123/93>
- Mellow, J. D., & Cumming, A. (1994). Concord in interlanguage: Efficiency or priming? *Applied Linguistics*, 15, 442–473.
- Miller, K. L., & Schmitt, C. (2012). Variable input and the acquisition of plural morphology. *Language Acquisition*, 19(3), 223–261. <https://doi.org/10.1080/10489223.2012.685026>
- Peirce, J., Gray, J. R., Simpson, S., MacAskill, M., Höchenberger, R., Sogo, H., Kastman, E., & Lindeløv, J. K. (2019). PsychoPy2: Experiments in behaviour made easy. *Behavior Research Methods*, 51(1), 195–203. <https://doi.org/10.3758/s13428-018-01193-y>
- Pinker, S. (1989). *Learnability and cognition*. Cambridge, MA: MIT Press.
- Po-Ching, Y., & Rimmington, D. (2016). *Chinese: A comprehensive grammar*. 2nd edition. London and New York: Routledge.
- Prawatmuang, W., & Yuan, B. (2020). Roles of positive and indirect negative evidence in L2 feature reassembly: An empirical study of L2 acquisition of Chinese and Thai collective markers. *Journal of Second Language Studies*, 3(2), 205–232.
<https://doi.org/10.1075/jsls.00011.pra>
- Prévost, P., & White, L. (2000). Missing surface inflection or impairment in second language acquisition? Evidence from tense and agreement. *Second Language Research*, 16, 103–133.
<https://doi.org/10.1191/026765800677556046>
- Saleemi, A. (1992). *Universal Grammar and language learnability*. Cambridge: Cambridge University Press.
- Sato, C. J. (1984). Phonological processes in second language acquisition: Another look at interlanguage syllable structure. *Language Learning*, 34(4), 43–57.
<https://doi.org/10.1111/j.1467-1770.1984.tb00351.x>
- Schwartz, B. D. (1993). On explicit and negative data effecting and affecting competence and linguistic behavior. *Studies in Second Language Acquisition*, 15, 147–163.
<http://doi.org/10.1017/S0272263100011931>
- Schwartz, M., & Goad, H. (2017). Indirect positive evidence in the acquisition of a subset grammar. *Language Acquisition*, 24(3), 234–264.
<https://doi.org/10.1080/10489223.2016.1187616>
- Slabakova, R. (2019). The Bottleneck Hypothesis updated. In T. Ionin & M. Rispoli (Eds.), *Three streams of generative language acquisition research: Selected papers from the 7th meeting of*

- Generative Approaches to Language Acquisition – North America, University of Illinois at Urbana-Champaign* (pp. 319–345). Amsterdam: John Benjamins.
- Song, J.-J. (2005). *The Korean language: Structure, use and context*. London and New York: Routledge.
- Trapman, M., & Kager, R. (2009). The acquisition of subset and superset phonotactic knowledge in a second language. *Language Acquisition*, 16(3), 178–221.
<https://doi.org/10.1080/10489220903011636>
- White, L. (1991). Adverb placement in second language acquisition: Some effects of positive and negative evidence in the classroom. *Second Language Research*, 7(2), 133–161.
<https://doi.org/10.1177/026765839100700205>
- White, L. (2003). *Second language acquisition and Universal Grammar*. Cambridge: Cambridge University Press.
- Wightman, C. W., Shattuck-Hufnagel, S., Ostendorf, M., & Price, P. J. (1992). Segmental durations in the vicinity of prosodic phrase boundaries. *The Journal of the Acoustic Society of America*, 91(3), 1707–1717. <https://doi.org/10.1121/1.402450>
- Yang, Y., & Wang, B. (2002). Acoustic correlates of hierarchical prosodic boundary in Mandarin. *Proceedings of Speech Prosody, 2002*, 707–710.
https://www.isca-speech.org/archive_open/sp2002/sp02_707.html
- Young, R. (1991). *Variation in interlanguage morphology*. New York: Peter Lang.

Appendix A. Participant information.

Participant ID	Age	Gender	Other languages and level of proficiency ⁱ	Performance on plurals (%)	Noticed errors in dialogue?	Connected errors to word learning?	Correct plurals due to what? ⁱⁱ
C2	35	F	French: I	8.33			guesses
C3	21	F	French: LI	12.5			guesses
C4	20	F	French: B	87.5			analogy
C5	23	M	French: B; Italian: LI	87.5			analogy
C6	19	F	French: B; German: LI	0			–
C7	24	F	Spanish: LI	16.67			guesses
C8	18	M	French: B; Spanish: B	0			–
C10	19	M	French: LI	0			–
C11	27	F	French: I; Spanish: LI; Italian: I	0			–
C12	33	M	French: LI	16.67			guesses
C13	21	F	French: LI	16.67			guesses
C14	23	F	Spanish: LI	4.17			guesses
C15	21	F	French: LI	91.67			analogy
C16	21	M	French: LI; Vietnamese: B	0			–
C17	25	F	Spanish: B; Tamil: B	12.5			guesses
C19	22	F	–	4.17			guesses
C20	20	F	French: LI	8.33			guesses
C21	19	M	French: LI; Spanish: LI	4.17			guesses
C22	24	M	Spanish: B	8.33			guesses
E1	21	M	French: I	91.67	yes	yes	IPE
E2	20	F	French: LI; Italian: LI	87.5	yes	yes	IPE
E4	24	M	French: LI; Arabic: B	20.83	no		guesses
E5	19	F	French: I; Turkish: B	91.67	yes	yes	IPE
E6	34	F	French: LI	20.83	yes	no	guesses
E7	20	M	French: LI; Spanish: LI	0	no		–
E8	21	F	French: LI; Spanish: LI	95.83	yes	yes	IPE
E9	19	F	French: LI; Spanish: I	87.5	yes	no	analogy

E10	26	F	French: I	25	yes	no	guesses
E11	21	F	French: I; Hebrew: B	75	yes	no	analogy
E12	21	M	French: LI	20.83	no		guesses
E13	26	F	–	95.83	yes	yes	IPE
E14	20	M	French: I	12.5	no		guesses
E15	19	F	French: I; Spanish: I	95.83	yes	yes	IPE
E16	23	F	Spanish: B	12.5	yes	no	guesses
E17	20	M	French: I	20.83	no		guesses
E18	24	M	Spanish: I; Thai: LI	91.67	yes	yes	IPE
E19	20	F	French: I	20.83	yes	no	guesses
E20	23	M	–	16.67	yes	no	guesses
E21	19	F	French: I	91.67	yes	yes	IPE

ⁱ B = beginner; LI = lower intermediate; I = intermediate.

ⁱⁱ IPE = indirect positive evidence.

Appendix B. Dialogue (contextually plural nouns produced without -s are underlined).

1. Liu Hua: Excuse me, I'm looking for the Student Service Centre. Do you know where it is?
2. Jerry: Yes. I'm heading there now. We can go together if you like.
3. Liu Hua: That would be great! Thank you!
4. Jerry: So, are you a new student at McGill?
5. Liu Hua: Yes, I am. I just arrived in Montreal. What about you?
6. Jerry: I'm a 3rd year student. Where are you from?
7. Liu Hua: I'm from China. This is my first time abroad. What about you? Are you a Canadian? I mean, you have blue eye, which is just like what I saw in western movie.
8. Jerry: Haha, you're right! I grew up in Montreal, so I know every corner of this city. I can show you around if you like.
9. Liu Hua: Cool! Then I assume you know where I can buy some food, like real coffee bean and some fresh fruit. Oh, and some snack too. I love potato chip.
10. Jerry: You can find all that stuff in local supermarkets, like Metro and IGA. Where do you live?
11. Liu Hua: I live in LaCité. Not far from campus. I haven't had time to explore nearby place yet. Are there many shop near where I live?
12. Jerry: Lucky you – your place is within walking distance of the campus. I know that Metro and Provigo are both close to where you live.
13. Liu Hua: What about kitchen item, like plate?
14. Jerry: I usually go to Dollorama. Things are very cheap there. And you can find almost everything you need for your kitchen in that one store. Do you live alone or do you share an apartment with others?
15. Liu Hua: I live alone. Though I'd like to have some dog, but the building manager doesn't allow me to have any pet.
16. Jerry: Yeah, that's pretty common in Montreal. It's for sanitary and safety reasons. Is your apartment a 3 and a ½?
17. Liu Hua: What do you mean by a 3 and a ½? I have four room: one bedroom, a living room, a kitchen, and a bathroom. Is that a 3 and a ½?
18. Jerry: Yes, that's a 3 and a ½. The ½ is the bathroom. It must be spacious for you since you live alone.
19. Liu Hua: Definitely! And it is totally empty now. I mean, literally. So I'm thinking about buying some flower to brighten it up.
20. Jerry: Good idea! You can find flowers at the supermarket or in a flower shop.
21. Liu Hua: Great! How about cinema and bookstore or maybe a library? I love watching film and reading. They are definitely two of my favourite hobby.
22. Jerry: There are lots of cinemas and bookstores downtown. The ones I usually go to are on Sainte-Catherine Street, very close to the campus. And there's a library on campus, right next to the Student Service Centre.
23. Liu Hua: Sounds good! I'm planning on visiting some nearby city during the summer break. Do you have any suggestion where I should go?
24. Jerry: Well, you could begin with Ottawa. It's the capital of Canada.
25. Liu Hua: Is it far from Montreal?
26. Jerry: Not at all. It takes about 3 hours to get there by train or bus. There are many interesting places to visit in Montreal too.
27. Liu Hua: I heard that Mont-Royal is famous.

28. Jerry: Yeah, it's nice. And there are interesting places that are close to Montreal too, like the Granby Zoo in the Eastern Townships. My family and I used to go there a lot when I was a kid.
29. Liu Hua: A zoo? I used to visit the zoo in my hometown with other kid from my school. It was fun. I especially love the elephant, but usually you can't find them in an urban zoo. It's amazing that their trunk are so flexible and they can pick up so many different object. How about the Granby Zoo? Are there any exotic animal there?
30. Jerry: Of course! It's a big zoo. There are elephants there too. But it's a bit far from Montreal. You'll need a car to get there. Actually, my friends and I are planning on going there next weekend. You can join us if you like.
31. Liu Hua: That's wonderful! I'd love to, thank you!
32. Jerry: You're welcome. Okay, here we are. You see the building on the left-hand side? That's the Student Service Centre.
33. Liu Hua: Is it the one behind the wall?
34. Jerry: Yup. And that's the library beside it.
35. Liu Hua: Thank you so much! By the way, my name is Liu Hua.
36. Jerry: I'm Jerry. Nice to meet you Liu Hua.
37. Liu Hua: It's nice to meet you Jerry. This is my e-mail. I'll see you next weekend!
38. Jerry: Great, see you then!