"She has many… *cat*?": On-line Processing of L2 Morphophonology by Mandarin Learners of English

Valeria Peretokina[1], Catherine T. Best[1,2], Michael D. Tyler[1,3], Bruno Di Biase[2]
[1]MARCS Institute, Western Sydney University, Australia
[2]School of Humanities and Communication Arts, Western Sydney University, Australia
[3]School of Social Sciences and Psychology, Western Sydney University, Australia
{V.Peretokina, C.Best , M.Tyler, B.Dibiase}@westernsydney.edu.au

Abstract

The current study examined on-line processing of second-language (L2) morphophonology by Mandarin learners of English, as compared to native English listeners, in a cognitively demanding self-paced listening task. Participants' listening times (LTs) to target singular and plural nouns that varied in phonological complexity and grammaticality were measured and analysed. The phonological representation of the targets was revealed to influence L2 listeners' processing speed the most, while morphological plurality and grammaticality did not seem to modulate LT in native or L2 participants. Thus, listeners appeared to disregard grammatical violations in favour of utterance comprehension.

Index Terms: speech perception, self-paced listening, morphophonology, grammaticality, processing load.

1. Introduction

Certain second language (L2) morphological structures, such as grammatical plurality, are challenging to acquire. Adult L2 learners often demonstrate inconsistent use of these units in speech production [1, 2] and insensitivity to inflectional errors in on-line reading comprehension [3, 4]. For instance, Mandarin Chinese (ManC) learners of L2-English are reported to show variable L2 production of English plural morpheme <-s> [1] as well as unresponsiveness to the omission of plural inflections in real-time reading tasks [4]. Research on L2 acquisition has linked these findings to representational gaps in learners' native language (L1) morphology [2, 5] (e.g., lack of grammatical plurality in Mandarin [6]) or phonology [7] (e.g., phonotactic restriction against /s/ in coda position or consonant clusters in Mandarin [8]). Thus, according to the representational gap account, a morphophonological feature is assumed to be unattainable for native-like mastery if it does not have a counterpart in L1 grammar or if the inflection of an L2 word stem yields L1 phonotactic violations.

It has been found, however, that adult ManC learners of L2-English are able to show sensitivity to plural marker <-s> and to its omission in a phoneme monitoring task [9]. In that study, ManC intermediate learners of Australian English (AusE) and AusE monolinguals were presented with audio recordings of English utterances and instructed to detect a target phoneme /s/ (realised as [s]) in various morphological (singular vs. plural nouns) and phonological (singleton vs. cluster codas) contexts. In 'catch' utterances, /s/ indicating plurality was omitted. The findings revealed L2 listeners' sensitivity to violations of English morphophonological alternations, and native-like processing patterns of word-final /s/. Interestingly, the presence of a morpheme boundary in plural nouns was observed to quicken L2 perception of /s/, even though Mandarin lacks both morphological marking of grammatical plurality and syllable-final /s/ as singletons or in clusters. This indicates that L2 learners were able to use morphological information in comprehension similarly to native listeners, implying that difficulties with L2 morphophonology in speech production and on-line reading tasks may not inevitably result from gaps in L1 morphology.

The inconsistency in previous findings may stem from the range of experimental procedures used to test acquisition of L2 morphophonology, with the majority of past studies examining L2 speech production or reading comprehension and only investigating perception of spoken L2 using phoneme monitoring. Varying cognitive demands of the tasks may also contribute to the discrepancy in the observed results. Spontaneous speech production and self-paced reading tasks targeting on-line reading comprehension put participants under a high processing load by introducing time pressure and requiring an extensive working memory use. Phoneme monitoring, on the other hand, tests perception of a single feature and does not explicitly check stimulus comprehension, which reduces the amount of cognitive resources needed to accurately perform the task and allows participants to access their L2 morphophonological knowledge, even for structures that are not represented in the L1.

Given the lack of empirical evidence on perception of spoken L2 under high processing load, the present study seeks to test ManC listeners' sensitivity to L2 morphophonological variations in a self-paced listening task. Self-paced listening is comparable to phoneme monitoring due to the use of *spoken* rather than *written* stimuli, but is more cognitively demanding as, similarly to self-paced reading, it investigates L2 learners' real-time speech processing. To explore perception of L2 morphophonology depending on phonological representation, morphological complexity, and grammatical violations of the target words, singular nouns ending in [Vs] (e.g., *house*) and [Cs] (e.g., *chance*), plural nouns ending in [Cs] (e.g., *cats*), and ungrammatical plural nouns with omitted inflection (e.g., She has many *cat*) were included in the design. Listening time (LT) to each word group was measured, with longer LT signifying processing delays, and in case of ungrammatical plurals, reflecting sensitivity to morphological violations.

ManC participants are hypothesised to find singular nouns ending in [Vs] the easiest for processing, even under a heavy cognitive load. Even though this group of singular nouns deviates from Mandarin phonotactic rules by having /s/ in codas, it complies with Mandarin morphology and does not contain L1 phonotactically impermissible coda clusters. By contrast, singular nouns ending in [Cs] are presumed to induce longer LT due to a more complex phonological representation that is inconsistent with ManC phonotactic rules. As for L2 listeners' perception of plurals, competing predictions can be made. On the one hand, if morphological awareness of L2-

English grammatical plurality is not obscured by the cognitive demand of the task, ManC listeners are expected to exhibit a difference in the LT to plural nouns in comparison to both singular noun groups. As L2 listeners have previously demonstrated faster processing of plural than singular nouns in phoneme monitoring [9], similar LT patterns may occur in the present experiment. Alternatively, retrieval of knowledge about L2-English grammatical plurality, coupled with performing a presumably challenging on-line task, could be manifested by longer LT to plurals relative to both groups of singular nouns. If ManC participants are sensitive to grammaticality of the targets, L2 processing of the omission errors is hypothesised to result in the longest LT. Otherwise, if L2 listeners' ability to detect errors is hindered by the processing demands of the task, LT to ungrammatical plurals is expected to be comparable to the LT to correct plurals.

To manipulate the amount of cognitive load within the task, targets appeared in sentences, either finally or medially. A longer LT to targets in final relative to medial position is expected, as it will reflect the increase in processing demand associated with participants retaining preceding words of the utterance in their working memory.

2. Method

2.1. Participants

Forty-eight participants were recruited for the present study in Sydney, Australia. The test group comprised 24 ManC speakers with a length of residence (LoR) in Australia of less than one year (15 females, 9 males; M_{age} = 24 years, Range: 18–32 years; M_{LoR} = 5 months, Range: 1–10 months). The English proficiency level of ManC participants was categorised as intermediate, as determined by their IELTS scores. Twenty-four AusE monolingual speakers (15 females, 9 males; M_{age} = 25 years, Range: 18–39 years), who were born and raised in Australia and did not indicate proficiency in any languages other than English, acted as the control group. None of the participants had any hearing, speech or language impairments.

2.2. Materials

A list of 272 experimental and filler five-syllable stimulus sentences was created for this experiment (examples presented in Table 1). Experimental sentences contained target monosyllabic singular and plural nouns ending in /s/ that appeared utterance-medially or finally. Using the CELEX database [10], we selected only high-to-medium frequency words as targets to ensure that targets' semantic difficulty would have minimal effect on participants' performance in the task ($M_{frequency}$ = 93 per million, Range: 10–528 per million).

Each group of target nouns was divided into two predicted processing difficulty levels, 'easy' and 'difficult'. Difficulty level of the target words was determined separately for each noun category. Singular nouns were categorised to be 'easy' or 'difficult' depending on their phonological representation with the singular nouns ending in [Vs] classified as 'easy', while those ending in [Cs] identified as 'difficult'.

Plural marker <-s> always appeared in [Cs], given that when it is preceded by a vowel it is realised as the allophonic variant [z] (e.g., *keys*) rather than [s]. As /z/ does not exist in the ManC phonological inventory, the [z] allophone of <-s> would have presented additional difficulty for L2 processing. Thus, within the category of plural nouns the difficulty level distinction was based on grammaticality of the targets. Correct plurals were labelled as 'easy' in comparison to ungrammatical plural nouns with omitted inflections, predetermined as 'difficult'. The 'easy' targets were presumed to require shorter processing time than the 'difficult' targets.

Filler sentences either contained no target phoneme /s/ or /s/ appeared in any word/utterance position. Participants' responses to fillers were not included in statistical analyses.

A 26-year-old female monolingual AusE speaker read the sentences word by word in a neutral voice. The session was recorded in a soundproof booth at 44.1 kHz sampling rate using a Shure SM10A-CN headset microphone, a MOTU Ultralite-mk3 audio interface, and Cool Edit Pro 2.1 software.

Royalty-free photographs were found on-line for presentation with the audio stimuli. Images were processed to be approximately the same size, with the long side set to 400 pixels and preserved aspect ratio.

2.3. Design and procedure

Testing sessions were conducted individually in a sound-attenuated booth using an Acer TravelMate P653 laptop and Sennheiser HD 650 headphones connected to an Edirol UA-25EX external sound card. Each participant completed a background questionnaire and a computer-based self-paced listening task, programmed and presented in DMDX [11].

At the beginning of each trial participants were simultaneously presented with the first word of an utterance and a picture depicting the meaning of the sentence on a computer screen. The participants' task was to pace themselves through each utterance by pressing a space-bar on a computer keyboard, indicating that they understood a word and were ready to move on to the next one, until the end of an utterance. The picture was present on the screen throughout the duration of the utterance. Each trial was followed by a 'yes/no' comprehension question asking if the utterance and the picture matched. Mismatched trials constituted 20% of the total number of trials and were distributed evenly across different sentence types. Mismatches were introduced to ensure that participants were attending to comprehension when performing the task. Trial presentation order was randomised for each participant. Comprehension rate was calculated and LTs were analysed only for trials with correctly answered comprehension questions.

3. Results

3.1. Comprehension rate

Responses to picture-sentence matching questions were recorded and percentage of correct responses was calculated for each participant. Then, the AusE and ManC groups' mean comprehension rate scores were derived from these values.

Table 1. *Stimulus sentence types.*

Stimulus Sentences	Morphological Structure	Difficulty Level	Utterance Position	
			Medial	Final
Experimental	Singular nouns	**Easy:** ending in [Vs]	Her new *house* looked clean.	I would like more *ice*.
		Difficult: ending in [Cs]	Take a *chance* to learn.	Buy that candy *box*.
	Plural nouns	**Easy:** correct plurals	Fluffy *cats* looked cute.	I came for five *nights*.
		Difficult: omitted inflection	Many *book* lay there.	She has many *cat*.
Fillers			We had a picnic. The *sky* went dark grey.	

While all listeners showed a high comprehension rate (M_{AusE} = 90%, M_{ManC} = 85%), which was clearly above chance in matched and mismatched trials (M_{AusE} = 91%, M_{ManC} = 88%; M_{AusE} = 86%, M_{ManC} = 78%, respectively), an independent-samples *t*-test revealed a significant difference between the L1 and L2 groups, $t(382) = 4.07$, $p < .001$. Lower than native comprehension performance of ManC listeners was expected due to their intermediate L2-English proficiency. That is, we infer that the group difference in picture-sentence matching accuracy did not indicate L2 participants' lack of utterance comprehension and did not prevent us from obtaining a sufficient number of correctly perceived trials to conduct a reliable analysis of LT data.

3.2. Listening time (LT)

After discarding trials with incorrect responses to picture-sentence matching questions, the remaining LT means, shown in Figure 1, were submitted to a 2 (Language group: ManC, AusE) x 2 (Utterance position: medial, final) x 2 (Morphological structure: singular nouns, plural nouns) x 2 (Difficulty level: 'easy', i.e., singular nouns ending in [Vs], correct plural nouns; 'difficult', i.e., singular nouns ending in [Cs], plural nouns with omitted inflection) analysis of variance. Language group was a between-subject factor; utterance position, morphological structure, and difficulty level were within-subject factors. Analyses were conducted by subjects (F_1) and items (F_2).

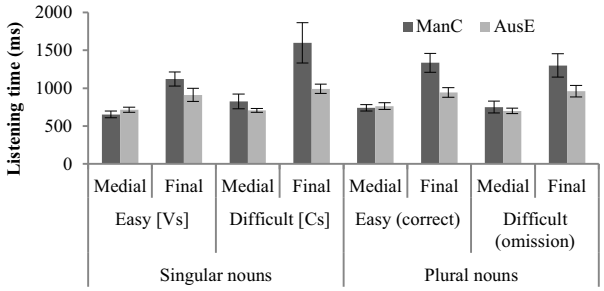

Figure 1: *Language group differences in listening time to target words as a function of utterance position, morphological structure, and difficulty level. Error bars represent standard error of the mean.*

As shown in Table 2, there were significant main effects of language group, with AusE participants exhibiting shorter LT (M = 837 ms) than ManC participants (M = 1041 ms), and utterance position, with longer LT to targets in final (M = 1147 ms) than medial position (M = 731 ms). A significant interaction between these two factors suggests that the utterance-final targets induced longer processing delays in L2 listeners than in L1 listeners (M_{ManC} = 1341 ms; M_{AusE} = 953 ms), while there were no significant group differences utterance-medially (M_{ManC} = 742 ms; M_{AusE} = 721 ms). A main effect of difficulty level reached significance in the subjects analysis, tentatively confirming our prediction that processing of 'difficult' targets would result in longer LT (M = 979 ms) than processing of 'easy' targets (M = 899 ms).

Two-way interactions between language group and difficulty level, and between morphological structure and difficulty level are also presented in Table 2. However, they will not be discussed in detail as the three-way interaction among language group, morphological structure, and difficulty level, shown in Figure 2, provides the most comprehensive account of the data. To examine the differences that led to the three-way interaction, simple effects tests with a Bonferroni correction were performed on the two simple two-way

Table 2. *Significant main effects and interactions (in bold) in the analysis of variance.*

Main effects	
Language group	**$F_1(1, 46) = 4.27$, $p = .044$** **$F_2(1, 136) = 3.21$, $p < .001$**
Utterance position	**$F_1(1, 46) = 49.4$, $p < .001$** **$F_2(1, 136) = 108.84$, $p < .001$**
Difficulty level	**$F_1(1, 46) = 4.45$, $p = .04$** $F_2(1, 136) = 2.53$, $p = .114$
Interactions	
Language group * Utterance position	**$F_1(1, 46) = 9.68$, $p = .003$** **$F_2(1, 136) = 29.59$, $p < .001$**
Language group * Difficulty level	$F_1(1, 46) = 3.85$, $p = .056$ **$F_2(1, 136) = 4.03$, $p = .047$**
Morphological structure * Difficulty level	**$F_1(1, 46) = 10.45$, $p = .002$** **$F_2(1, 136) = 7.41$, $p = .007$**
Language group * Morphological structure * Difficulty level	**$F_1(1, 46) = 5.01$, $p = .03$** **$F_2(1, 136) = 5.67$, $p = .019$**

interactions between morphological structure and difficulty level for each language group. The data were collapsed across the utterance position factor, as it did not contribute to the interaction, and split by language group.

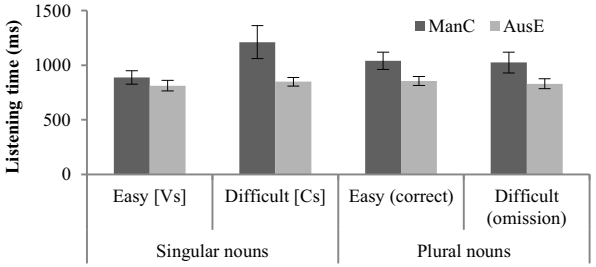

Figure 2: *Language group differences in listening time to 'easy' and 'difficult' morphological structures. Error bars represent standard error of the mean.*

For ManC participants (the dark grey bars in Figure 2), the simple two-way interaction between morphological structure and difficulty level was significant, $F_1(1, 23) = 8.37$, $p = .008$, $F_2(1, 136) = 7.54$, $p = .007$. Simple effects analyses revealed that the L2 group's LT to 'difficult' targets remained similar regardless of morphological structure, $F_1(1, 23) = 2.39$, $p = .136$, $F_2(1, 136) = 1.86$, $p = .175$, but was statistically different for 'easy' targets, $F_1(1, 23) = 25.63$, $p < .001$, $F_2(1, 136) = 5.76$, $p = .018$. That is, the LT to 'easy' singular nouns, i.e., singular nouns ending in [Vs] (M = 888 ms), was significantly shorter than the LT to 'easy' plural nouns, i.e., correct plurals (M = 1041 ms). Furthermore, a significant difference in L2 participants' LT was found when comparing the two difficulty levels within singular nouns, $F_1(1, 23) = 8.23$, $p = .009$, $F_2(1, 136) = 8.16$, $p = .005$. ManC listeners took more time to respond to 'difficult' singular nouns ending in [Cs] (M = 1212 ms) relative to 'easy' singular nouns ending in [Vs] (M = 888 ms). Hence, the shortest LT in ManC group was observed for singular nouns ending in [Vs], with this difference being confirmed both against singular nouns ending in [Cs], which constitutes the other level of the morphological structure factor for singular nouns, and against correct plurals, which represents the other pre-identified 'easy' target word type. As shown by the light grey bars in Figure 2, a simple two-way interaction between morphological structure and difficulty level for AusE participants was not significant, $F_1(1, 23) = 1.23$, $p = .14$, $F_2(1, 136) = .71$, $p = .4$. This indicates that there were no variations in L1 participants' LT depending on morphophonology and grammaticality of the targets.

The simple effects analyses allowed us to directly compare the ManC group's LT to correct plural nouns and singular nouns ending in [Vs] and revealed that plurals induced longer processing time in L2 listeners. However, it is unclear whether the increase in LT to plural nouns was observed due to their morphological complexity or their phonological representation (i.e., coda cluster). Therefore, a subsequent paired-samples *t*-test was carried out to compare ManC group's LT to plural and singular nouns ending in [Cs], as these two word groups differ only in morphological complexity. If it is indeed the grammatical inflection encoded by /s/ that makes plural nouns difficult for processing, we would expect ManC participants to demonstrate a longer LT to plural nouns than singular nouns ending in [Cs]. A lack of difference in LT to these two word groups, however, would suggest an L1 phonological rather than L1 morphological influence on L2 processing. No significant difference was uncovered by the *t*-test ($M_{\text{correct plurals}}$ = 1041 ms, $M_{\text{nouns [Cs]}}$ = 1212 ms), $t(47) = 1.38$, $p = .176$, which is consistent with the idea that it was the L1-violating phonological representation of the target words and not the L1-absent grammatical plurality that predominantly modulated the ManC group's processing speed.

4. Discussion

The aim of this study was to examine on-line processing of L2-English morphophonology by ManC listeners under a high cognitive load of a self-paced listening task. In particular, we investigated participants' LT to target singular and plural nouns depending on their phonological representation, morphological complexity, and grammaticality.

Our findings provide clear evidence for the cognitive load effect on perception of spoken L2. As anticipated, both participant groups experienced delays when processing utterance-final targets. Slower LT was presumably caused by the necessity to retain an entire utterance in working memory in preparation to answer a comprehension question, which in turn resulted in the increased cognitive load and hindered processing. The utterance-final delays were more prominent in the performance of the ManC than the AusE group. This observation highlights the difference between L1 and L2 speech processing, with the former being more automated and the latter requiring more cognitive effort.

Unlike the results of the previous phoneme monitoring study [9], which had demonstrated near-native performance of L2 learners, the present experiment revealed distinct differences between the processing patterns of the two participant groups. While AusE listeners did not show any LT variations depending on morphophonological complexity of the target words, ManC participants' processing was observed to be modulated by the phonological representation of the target nouns. In line with our hypothesis, L2 listeners demonstrated the shortest LT to singular nouns ending in [Vs], confirming L1 phonological and phonotactic influences on L2 speech perception. In comparison to singular nouns ending in [Vs], longer LTs were found for both singular and correct plural nouns ending in L1-impermissible [Cs], with no significant differences between the two. This finding suggests that it is indeed phonological representation that determined L2 processing speed, whereas grammatical plurality did not appear to contribute significantly to LT variability.

In accordance with previous studies on on-line reading comprehension [3, 4], L2 learners did not differentiate between grammatical and ungrammatical targets, which could potentially suggest a representational gap in their L2 grammatical system or difficulty accessing L2 morphological knowledge under time pressure. However, this interpretation is contradicted by the fact that, surprisingly, native listeners also did not demonstrate any changes in LT to plurals with omitted inflections in comparison to their correct counterparts. This unexpected finding could have been due to the experimental procedure itself, as participants appeared to concentrate on the comprehension of the utterances and responses to picture-sentence matching questions instead of attending to grammatical structures. Also, given that AusE participants come from an area of Sydney that shows high multi-cultural representation, it would seem likely that they have had substantial exposure to L2 learners of English of various backgrounds and proficiency levels prior to participating in the study. This could have made native listeners less sensitive to inflectional errors and potentially explain the lack of response to grammatical violations.

In sum, the present study provides an extensive account of L2 morphophonology perception in spoken language and verifies that the processing load of the task and the phonological representation of the targets have a pronounced effect on L2 learners' perceptual patterns. Contrary to the findings of the phoneme monitoring experiment [9], morphological structure and grammaticality did not influence L1 or L2 participants' LTs in the current study, indicating that processing of grammatical inflections and error identification are largely determined by the cognitive demands of the task. This also might suggest that listeners ignored grammatical errors and focused on attending to the meaning of the utterances. Our subsequent studies will test ManC participants with a longer LoR in Australia to explore whether L2 exposure improves morphophonology processing.

5. References

[1] G. Jia, and A. Fuse, "Acquisition of English grammatical morphology by native Mandarin-speaking children and adolescents: Age-related differences," *J. Speech, Lang. Hear. Res.*, vol. 50, no. 5, pp. 1280–1299, 2007.

[2] R. Hawkins and S. Liszka, "Locating the source of defective past tense marking in advanced L2 English speakers," in *The Interface between Syntax and Lexicon in Second Language Acquisition*, R. van Hout, H. Aafke, F. Kuiken, and R. Towell, Eds. Amsterdam: Benjamins, 2003, pp. 21–44.

[3] N. Jiang, E. Novokshanova, K. Masuda, and X. Wang, "Morphological congruency and the acquisition of L2 morphemes," *Lang. Learn.*, vol. 61, no. 3, pp. 940–967, 2011.

[4] N. Jiang, "Selective integration of linguistic knowledge in adult second language learning," *Lang. Learn.*, vol. 57, no. 1, pp. 1–33, 2007.

[5] I. M. Tsimpli and M. Dimitrakopoulou, "The interpretability hypothesis: Evidence from wh-interrogatives in second language acquisition," *Second Lang. Res*., vol. 23, pp. 215–242, 2007.

[6] C. N. Li and S. A. Thompson, *Mandarin Chinese: A functional reference*, Oakland, CA: University of California Press, 1989.

[7] D. Lardière, "Attainment and acquirability in second language acquisition," *Second Lang. Res.*, vol. 22, pp. 239–242, 2006.

[8] S. Duanmu, *The phonology of Standard Chinese*, New York: Oxford University Press, 2007.

[9] V. Peretokina, C. T. Best, M. D. Tyler, J. A. Shaw, and B. Di Biase, "Perception of English codas in various phonological and morphological contexts by Mandarin learners of English," in *Proc. of the 18th Int. Congress of Phonetic Sciences, ICPhS 2015, 10–14 August 2015, Glasgow, Scotland* [Online], Available: https://www.internationalphoneticassociation.org/icphs-proceedings/ICPhS2015/Papers/ICPHS0603.pdf.

[10] R. H. Baayen, R. Piepenbrock, and H. Van Rijn, *The CELEX lexical database*, Philadelphia: University of Pennsylvania, 1993.

[11] K. I. Forster and J. C. Forster, "DMDX: A Windows display program with millisecond accuracy," *Behavior Research Methods, Instruments & Computers*, vol. 35, pp. 116–124, 2003.