NOVITAS ROYAL
RESEARCH ON YOUTH AND LANGUAGE

"I STILL CAN'T QUESTIONS": ISSUES AFFECTING EFL DEVELOPMENT IN AN IMMERSION ENVIRONMENT

Clare WRIGHT*

Abstract: This study examined the development of EFL proficiency in an immersion environment. Adult Chinese speakers of English were tested at the beginning and end of a ten-month period of immersion in the UK on their acquisition of English question forms using a timed grammaticality judgement task. Participants showed significantly faster response times after ten months, but no significant difference in accuracy of target-like judgements, suggesting that immersion benefits fluency more than accuracy.

Key words: adult EFL learners, instruction, immersion, Chinese L1

Özet:. Bu çalışma yabancı dil olarak İngilizce'nin gelişimini yoğun bir pratik ortamında incelemiştir. İngilizce öğrenen yetişkin Çinli öğrencilere, zaman bazlı dilbilgisellik değerlendirme görevi kullanılarak, İngiltere'deki on aylık zaman diliminin başında ve sonunda İngilizce soru şekillerini edinimleri üzerine bir test uygulanmıştır. Katılımcılar 10 ayın sonunda belirgin bir biçimde daha çabuk tepkiler vermiş, ancak hedef dil bazlı doğrulukta belirgin bir fark görülmemiştir; ki bu hedef dil ortamında yoğun pratiğin yabancı dilde akıcılığı doğruluktan daha çok etkilediğini önermektedir.

Anahtar Sözcükler: Yetişkin İngilizce öğrencileri, öğretim, yoğun pratik, Anadili Çince olan öğrenciler

Background

The study discussed here examines individual variation in adult acquisition of a second language (L2), in the context of proficiency in English wh-questions (wh-movement) by instructed Chinese speakers of English.

Wh-movement has been long identified as an area of individual variation in grammatical proficiency, and target-like proficiency is deemed to be difficult, if not impossible, for adult learners (Johnson & Newport 1989), where adult is usually taken to mean over the age of around ten years. Age and length of exposure is commonly argued (as by Johnson & Newport, see also DeKeyser 2000 for an overview) to be the key factor affecting such variation, but other explanations have been put forward (Moyer 2004) including learner-internal factors such as motivation, attitude, gender, memory capacity and learning style, and learner-external factors such as instructional background, amount and type of interaction, and native-language (L1) transfer. The task thus facing most typical teenage learners of English can seem a daunting one, with so many issues potentially affecting how they become proficient in English.

In addition to these general factors affecting L2 acquisition, different types of English questions create specific asymmetries in acquisition. Simple questions – "What did you see?" are seen as acquired earlier than complex questions -"What did you think you saw?" (Pienemann 1998). An

*NewcastleUniversity, Newcastle upon Tyne, UK, clare.wright1@newcastle.ac.uk

asymmetry has also been found between complex object questions -"What did Mary say Tom liked?" and complex subject questions - "Who did Mary say liked the song?" – where objects are seen as quicker to process than subjects (Schachter & Yip 1990, White & Juffs 1998), but it has not, to my knowledge, been established whether this asymmetry is also found in simple questions.

A number of studies have suggested that instructed learners can become highly proficient after explicit or interactional classroom exposure (Ellis 1994, Ellis et al. 2009, Doughty 2001, Mackey et al. 2002, Sanz & Morgan-Short 2005); some researchers suggest that immersion may not even be necessary to achieve very advanced or native-like levels of proficiency (White & Juffs 1998). The majority of studies find wide individual variation and fossilisation at advanced levels, even in cases of long-term residence in the L2 environment (see, e.g. Han 2004, Birdsong 2005, Wright 2006, Lardiere 2007).

The question of how adult L2 proficiency develops during immersion is thus still the subject of debate, and longitudinal research on the effect of immersion exposure is sparse, in particular to assess the individual variation in amount and rate of development arising from the change from an instructed environment to an immersion environment. This study seeks to address some of these issues, with particular reference to Chinese learners of English, whose exposure to interactional or native-like input is seen as limited (Gu 2003), and whose L1, Mandarin, lacks overt wh-movement. Therefore studying the impact of immersion on Chinese learners of English should provide useful empirical evidence on how native-language transfer and limited-input exposure affect learners when they find themselves in an immersion environment. The findings should also shed light on how to focus attention best in the classroom on different grammatical structures to ensure learners can make maximum progress, even at more advanced levels, in the most effective way.

Study Hypotheses

This study examines the issues discussed above through a longitudinal study of Chinese speakers of English (CSE) during a period of ten months' immersion in the UK.

Two hypotheses were tested:

1) Advanced instructed learners of English would show significant changes in grammatical proficiency during immersion (measured as faster speed and greater target-like accuracy on a timed grammaticality judgement task);

2) Proficiency on two types of questions would show asymmetric development during immersion: improvements in proficiency would be greater for simple questions vs. complex questions, and for object questions vs. subject questions.

Methodology

Participants

Thirty-two volunteer participants were recruited among Chinese-speaking postgraduates studying at British universities (twenty-four female and eight male), all with an IELTS score of 5.5 or 6. Bio-data on learning background and exposure to input were gathered via a questionnaire, to test for inter-learner variation in exposure to English prior to immersion (Dornyei 2003). Mean age of learning (AOL) was 11.41 (range: 7-14); length of learning (LOL) was 11.77 years (range: 5 to

18 years). No significant effects on IELTS score were found for gender, AOL or LOL, so participants' proficiency was assumed to be homogenous before the period of immersion.

Task Design

A Reaction Time judgement task was devised containing 76 tokens, with 40 simple and 36 complex questions, which comprised 28 object and 28 subject constructions, and 20 other items (not reported on here), and equal numbers of grammatical and ungrammatical items. The items were balanced for lexical content, and between six and ten words long (see Appendix for sample items). The items were computerised using DMDX[1] for randomised timed presentation via a laptop. An introductory screen explained that the participants were to see sentences which they were to judge as grammatically acceptable or not, on a Likert scale of -2 (unacceptable) to +2 (acceptable). They were told they should respond as quickly as possible with their initial instinctive response. Their choice was activated by pressing pre-programmed and labelled buttons easily accessed at either edge of the laptop keyboard (Left Control, Left Shift, End, Right Arrow, programmed as the -2, -1, +1 and +2 buttons respectively). Each item appeared in full, and centred on the screen. The task was self-timed: the clock (recording in milliseconds) started after the item appeared on the screen, and continued until one of the labelled buttons was pressed; this action also then generated the next item to appear, in a different random order generated by the software for each participant. Three practice items were presented to ensure the participants understood how to use the buttons appropriately. Pressing the space bar started the experiment which continued until the words "Test finished – Thank you" appeared on the screen accompanied by a beep. The output recording speed and selection of button was generated as a .zil file encoded in SPSS.

Participants were seen individually in a quiet room for data collection on two occasions: Time 1, within two weeks of arrival in the UK, and Time 2, after ten months' immersion. The Reaction Time (RT) task was administered both times under similar conditions. The information on AOL, LOL and other bio-data referred to above was collected at Time 1.

Results

All the data on speed and accuracy were encoded using SPSS. Descriptive statistics were collected for speed and accuracy of response at both Time 1 (T1) and Time 2 (T2) to see how far these two measures changed during immersion. The data were then analysed using related-samples t-test, to check for significant differences between the two times of data collection.

For the first research hypothesis, overall reaction time and accuracy were analysed first. All scores for speed (i.e. reaction times) are reported in seconds rather than milliseconds for ease of reference. Accuracy is shown in raw scores out of a possible maximum of 76 (an accurate response was scored when +2 or -2 was pressed for grammatical or ungrammatical items respectively). Mean time at T1 was 506.37 seconds; mean time at T2 was 432:81. Mean accuracy at T1 was 40.13; mean accuracy at T2 was 39.69. The mean scores for overall speed and accuracy at T1 and T2 are shown in graph form in Figure 1a and 1b below. The y-axes on the speed the range of minimum to maximum actual individual scores.

---

[1] DMDX software, an alternative to PsyScope or E-prime, was developed by Ken and Jonathan Forster at Monash University and the University of Arizona, and is freely available to download (http://www.u.arizona.edu/~jforster/dmdx.htm)

Figure 1a: overall time

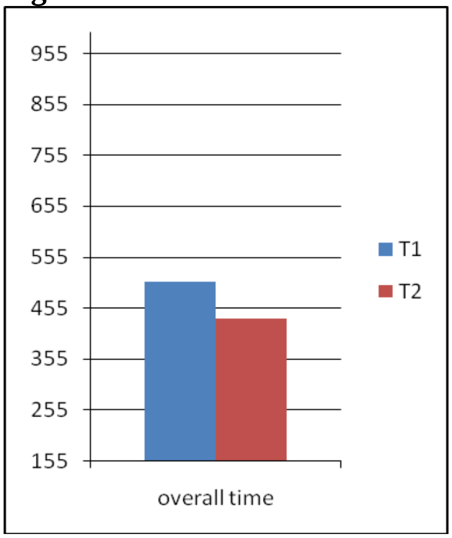
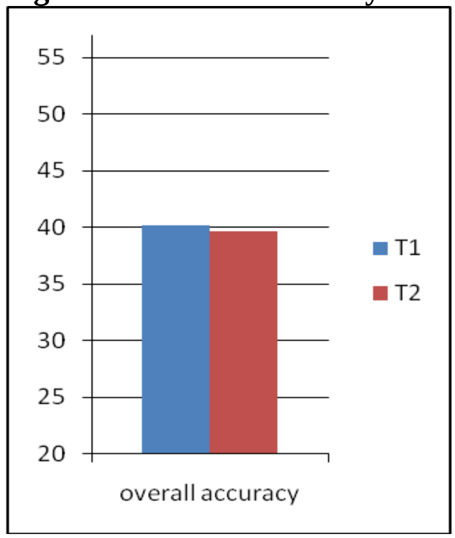


Figure 1b: overall accuracy



It is evident that participants speeded up in their reaction times by T2, an improvement which was significant ($p<.05$). However, accuracy did not show any improvement – in fact mean scores very slightly and non-significantly decreased ($p>.05$), providing conflicting evidence for the first research hypothesis that proficiency would significantly improve during immersion.

The data were then analysed in more detail by question type (40 simple vs. 36 complex questions, 28 object vs. 28 subject questions) to check for the expected asymmetries outlined in the second research hypothesis. As above, reaction time is shown in seconds; accuracy scores are shown out of a possible maximum of 40.

For speed of response of simple questions, mean time at T1 was 239.74 seconds; mean time at T2 was 204 seconds. For speed of complex questions, mean time at T1 was 265.79 seconds; mean time at T2 was 228.82 seconds. For accuracy of simple questions (out of 40), mean score at T1 was 22.97; mean score at T2 was 23.47. For complex questions (out of 36), mean score at T1 was 17.1; mean score at T2 was 16.22. These results are illustrated in Figure 2a and 2b below (as above, y-axes represent the range of individual actual scores from minimum to maximum).

Fig 2a: Time (simple vs. complex)

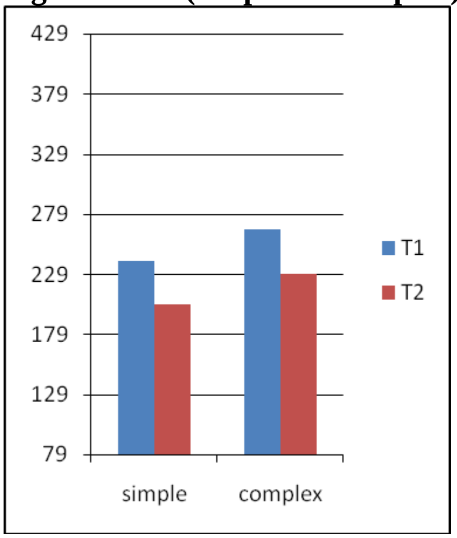
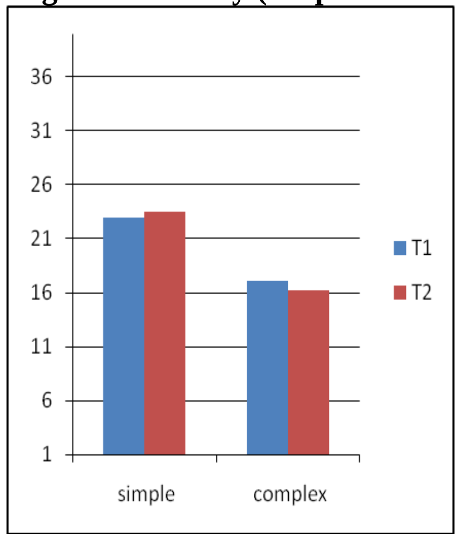


Fig 2b: Accuracy (simple vs. complex)



As predicted, simple questions were judged significantly faster than complex questions at both T1 and T2 ($p<.001$), and speed at T2 was significantly quicker than at T1 for both simple questions ($p<.01$) and complex questions ($p<.05$). Participants showed a significantly greater improvement by T2 ($p<.001$) for reaction times on complex questions (36.97 seconds) than simple questions (35.74 seconds). For accuracy, the results showed, as expected, that simple questions were judged significantly more accurately than complex questions, at both T1 and T2 ($p<.001$). However, simple questions showed only a slight and non-significant improvement between T1 and T2. Additionally, accuracy on complex questions actually decreased between T1 and T2 (though again only slightly, and non-significantly), as implied by the lack of improvement for overall accuracy shown above.

Turning now to check the predicted asymmetry between object and subject questions, mean time at T1 for object questions was 174.89 seconds; mean time at T2 was 143.56 seconds. For subject questions, mean time at T1 was 181.7 seconds; mean time at T2 was 155.34 seconds. For accuracy (out of 28), mean accuracy for object questions at T1 was 16.94; mean accuracy at T2 was 16.81. For subject questions, mean accuracy at T1 was 13.47; mean accuracy at T2 was 14.84. These results are shown in Figure 3a and Figure 3b below.

Fig 3a: Time (object vs. subject)

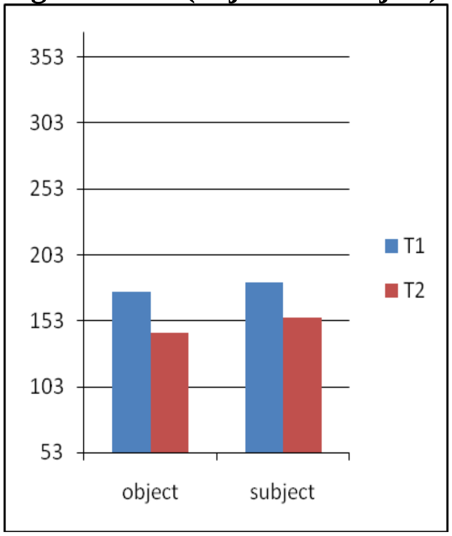
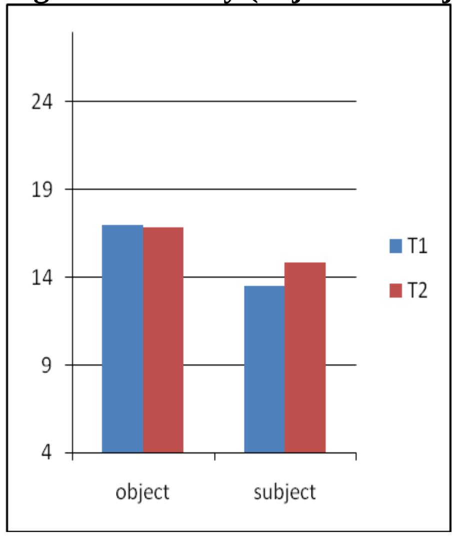


Fig 3b: Accuracy (object vs. subject)



As predicted, object questions were processed more quickly than subject questions at both T1 and T2, though non-significantly. But object speed showed greater improvement by T2 (mean decrease in time 31.33 seconds) than subject speed (26.36 seconds, a difference which was significant ($p<.005$).

For accuracy scores, as expected, objects were judged more accurately than subjects at T1 and T2 (non-significant), but, surprisingly, object accuracy did not improve between T1 and T2 (showing a trivial decrease of .13 out of 28, or less than 1%), whereas subject accuracy did improve between T1 and T2 (by 1.37 out of 28, or by 5%). However, as inferred from the lack of significant improvement in overall accuracy seen above, these changes were not statistically significant.

Given the lack of clear patterns of development in mean scores for the different question types between Time 1 and Time 2, I examined the individual patterns of variation at Time 2 for accuracy on simple, complex, object and subject questions, to assess how these compared. There was a greater range found for simple questions (55%) than for complex questions (50%), and much higher maximum scores on simple questions (85%) than complex questions (66%). The range was greater on subject questions (64%) than on object questions (50%), although maximum scores were similar (86% for object questions; 82% for subject questions). These scores are shown in raw form in Table 1 below.

Table 1: Individual variation in question types at Time 2

Time 2 Accuracy (/max)	*Min*	*Max*	*Range*	*SD*
Simple (/40)	12	34	22	4.88
Complex (/36)	6	24	18	4.62
Object (/28)	10	24	14	3.62
Subject (/28)	5	23	18	4.10

Thus, across the pool of participants, there was evidence of a wide range of proficiency on all four types, but in predictable patterns following the asymmetries shown above, with simple and object questions showing the highest maximum levels of accuracy, and subject and complex questions showing the lowest minimum levels of accuracy.

Discussion

A group of advanced Chinese instructed learners of English (IELTS 5.5 or above) were tested for development in their proficiency in English question forms over a period of 10 months' immersion in the UK. The research hypotheses in this study examined how much their overall scores on a timed grammaticality judgement task changed in speed and native-like accuracy during this period (Time 1 to Time 2), and whether these changes were predictable, according to suggested asymmetries between simple and complex questions, and between object and subject questions.

The results showed that there was a significant improvement in their overall speed, and that speed at both Time 1 and Time 2 reflected quicker times for simple questions compared to complex questions, and for object questions compared to subject questions, as predicted. However, the improvement in time for complex questions was greater than for simple questions. I suggest that reaction times for simple questions were near optimum even before immersion, as could be expected given the high salience of such question types in a standard instructed curriculum. Therefore increased exposure would not make any significant impact on reaction times. Complex questions are argued to be more difficult to process and later acquired (Pienemann 1998), and the improvement in times shown here suggest that immersion markedly facilitated acquisition of complex questions, to the extent that reaction times for complex questions by Time 2 nearly equalled those for simple questions at Time 1.

The expected asymmetry between object and subject questions was also found, but the greater change in reaction times for object questions than for subject questions suggest that both types of question were not yet optimally proficient. The easier processing load predicted for object questions (White & Juffs 1998) is argued to explain the greater improvement for object questions than for subject questions found here.

Therefore, in terms of reaction times, the data from this study show robust evidence for greater proficiency in speed of response as a result of immersion, in a predictable pattern or implicational hierarchy favouring simple over complex questions, and object over subject questions.

However, in terms of accuracy, the data show less clear patterns of evidence. Overall accuracy did not significantly change during immersion, remaining just under chance at both Time 1 and Time 2. Subject questions showed the most improvement, and simple questions showed slight improvement by Time 2, although both types remained at or just above chance. Object questions are argued here to be acquired (where 60% accuracy is deemed to show acquisition, see, e.g. Vainikka & Young-Scholten 1998), but they showed no improvement in accuracy beyond this point between Time 1 and Time 2; complex questions were judged below chance at both times, and showed some decline in accuracy by Time 2. This puzzling discrepancy between improvement in time but decline in accuracy for complex questions at Time 2 was hard to explain so I looked more closely at individual scores to see if there were any difference between participants which could explain the results shown in the statistical analysis.

Closer investigation of individual variation in accuracy by time 2 confirmed that for some individuals within the pool of participants, simple questions, object questions and subject questions could be judged very accurately (above 80%), but even for the most proficient individuals, complex questions were still the most difficult (maximum was 64%). This suggests that the implicational hierarchy outlined above still holds even for more proficient L2 users.

However, the wide range on all four question types showed that there were participants who were well below chance even at Time 2, although even for these less proficient participants, the implicational hierarchy outlined above still applied. Minimum scores were highest for object questions (36%), then simple questions (30%), then subject questions (18%) and lowest for complex questions (17%). It could be argued that immersion should have had greatest impact on these lower-proficiency participants. In order to assess if this wide range of proficiency affected the overall findings, the participants were split into three groups according to Accuracy scores from Time 1, and an ANOVA was run to compare Accuracy scores by group at T2. No significant between-group differences were found ($p>.05$), suggesting that individual variation in accuracy was spread across all groups, and immersion did not have a differential impact according to level of proficiency on arrival.

It could be argued, additionally, that immersion was not consistent for all participants, and some may have engaged far more than others in active use of L2 over and above the standard content-based input arising from their postgraduate studies, which would have helped their individual rates of improvement. Diary data was collected at Time 2 from eighteen out of the thirty-two participants showing their use of English over the period of a week, in order to calculate the average daily use of English (mean daily use: 7.77 hours per day, ranging from 3.3 hours minimum to 14.7 hours maximum). Correlational analysis however showed no significant correlation between daily use and accuracy at Time 2 ($r=-.130$, $p>.5$).

A number of methodological issues were explored to see if there was an aspect of the study design which explained the difference between the clear improvement shown in reaction times and the lack of clear improvement seen in accuracy scores, especially for complex questions.

Firstly, looking at the evidence from both reaction times and accuracy, it could be argued that some participants may have traded accuracy for speed of response, particularly for complex questions. However, there was no correlation found (using Pearson correlational analysis) between time and accuracy overall ($r=.057$, $p>.05$) nor for complex questions ($r=.059$, $p>.05$), implying that there was in fact no trade-off.

Secondly, the reaction time task design may have confounded the results by creating differences in reading time that this study was not able to factor out. Although item length was balanced as far as possible, simple questions were on balance shorter than complex questions, thus facilitating a possible “built-in” faster speed for simple questions. However, there were four more simple items than complex items which was designed to offset any potential inequality. Furthermore, the significantly greater improvement in speed for complex questions by Time 2 suggests that any “built-in” advantage for simple questions did not distort the real pattern of changes in reaction times.

Thirdly, grammaticality judgement tasks themselves have been argued to be an unreliable method of testing linguistic knowledge, particularly in the context of what are seen as opposing sides of a divide between explicitly taught or more salient linguistic structures versus structures that cannot be taught or are rare in the input (Mandell 1999, Bialystok 2002, Sorace 2003, Paradis 2004). However, L2 developmental studies do not always clearly show evidence of this theoretically-motivated divide between types of knowledge, but more of a spectrum of linguistic knowledge, accessed through a "coalition of resources" (Herschensohn 1999: 220). The task in the study described here used a balance of both more and less salient structures in order to avoid any confounding effect of linguistic structure.

Fourthly, reaction time tasks have also been argued to over-simplify the analysis of what linguistic knowledge of what is being tapped, by normally using a two-way distinction of "acceptable" or "unacceptable" which can lead to purely random choices. A four-way judgement scale was used here minimise any impact of random guessing, although of course, I cannot be completely sure that guessing was not used as a strategy to some degree.

Comparing which types of question showed most improvement or decline in either speed or accuracy by Time 2, it was shown that complex subject questions were judged the slowest and least accurately at Time 2, and that simple object questions were judged the quickest and most accurately. Due to test design, it was not possible to confirm if the improvement in subject question accuracy (the only significant improvement in accuracy of any measure) lay in simple or complex questions, but since complex questions declined in accuracy, I infer that this improvement lay in simple subject questions.

While there is no obvious explanation for the lack of improvement in accuracy in complex questions, I suggest that participants' existing knowledge of complex questions was based on a number of holistically stored chunks of object question structures (Myles 2004), derived from typical classroom instruction at intermediate level and above (e.g. Acklam 1996). These chunks provided a basis for participants to show improvement in response times following immersion, but would perhaps also undergo some decomposing and restructuring, to allow the underlying grammatical rule to be abstracted and then generated more accurately, in a so-called "U-shaped" developmental pattern (Kellerman 1985). The same argument could apply also to complex subject questions; however, such structures appear to be less common in the input (almost none are provided in a common intermediate level text book such as Acklam 1996). I suggest then that complex subject questions are less likely to exist as stored holistic structures, and would need to be acquired more or less from scratch, which could require more than ten months' immersion

To conclude, predicted asymmetries were found between simple and complex questions and between object questions and subject questions, confirming previous evidence that simple questions are acquired before complex questions, and object questions are easier to process than subject questions. The data shown here suggest that simple object questions are acquired first, and can processed at near optimum speed and accuracy without immersion, since these question types showed least change during immersion. Complex subject questions are argued to be acquired last, and are deemed to be the hardest to process, since levels of accuracy and speed changed little and were far from target-like even after immersion. However, it was shown that accuracy on complex object questions also went down slightly during the period of investigation, arguing that for complex questions, ten months' immersion may not be long enough to restructure linguistic knowledge to target-like levels. It is not clear why this is so, although some change, perhaps in a “U-shaped” development, was underway where some complex items, perhaps processed before immersion as chunks (Myles 2004), were being reanalysed prior to more accurate generation.

The implication for teachers and students of the asymmetries shown here, and of the difficulties in restructuring complex questions in particular, is that later acquired structures could be promoted in the input through a wider range of question forms, and made more salient through explicit presentation of the more difficult complex forms, especially subject questions. However, this should not be done until the earlier acquired forms are securely in place, to avoid too much reliance on un-analysed chunks.

For learners seeking to improve their proficiency by making the effort to immerse themselves in an English setting, it could avoid frustration to recognise that immersion for less than one year may not always be sufficient to trigger significant improvement in accuracy, particularly for complex questions. Previous research has found that even decades of immersion do not necessarily guarantee target-like accuracy (e.g. Lardiere 2007), suggesting that exposure in itself is necessary but not sufficient factor in L2 acquisition. It seems that for advanced instructed learners, immersion has most effect on improvements in efficient processing of existing instructed knowledge rather than developing greater proficiency in more difficult or new linguistic knowledge.

Conclusion

Increased exposure arising from one year’s immersion in an L2 environment can significantly facilitate improved proficiency in that L2, as shown here in increased speeds of processing on a timed grammaticality task. Four types of questions were examined, revealing an asymmetry between simple and complex questions, and between object and subject questions. An implicational hierarchy is suggested that simple object questions are processed most quickly, then simple subject questions, then complex object questions, and complex subject questions are processed most slowly. However, accuracy on these four structures did not significantly change during immersion, particularly for complex questions, possibly due to reliance on “chunks” learned during previous instruction.

These findings suggest that immersion helps learners process what linguistic knowledge they already have with greater efficiency, rather than lead to acquisition of new linguistic knowledge. Current theories of L2 acquisition do not yet have a comprehensive model of how L2 reaches an end-state (Birdsong 2005), and it is clear that further research is required to explain more clearly how the acquisition/processing interface operates in that transition.

References

Acklam, R. (1996). *First Certificate Gold Coursebook*. Harlow, Longman.
Bialystok, E. (2002). Cognitive processes of L2 users. In V. Cook (ed.). *Portraits of the L2 User*. Clevedon: Multilingual Matters. 145-166.
Birdsong, D. (2005). Nativelikeness and non-nativelikeness in L2A research. *International Review of Applied Linguistics, 43*, 319-328.
DeKeyser, R. (2000). "The robustness of critical period effects in second language acquisition." *Studies in Second Language Acquisition* 22: 499-533.
Dornyei, Z. (2003). *Questionnaires in second language research*. Mahwah, NJ: L. Erlbaum.
Doughty, C. (2001). Cognitive underpinnings of focus on form. In P. Robinson (ed.). *Cognition and second language instruction.* Cambridge, Cambridge University Press: 206-257.
Ellis, R. (1994). *The Study of Second Language Acquisition*. Oxford, Oxford University Press.
Ellis, R., S. Loewen, R. Erlam, J. Philp, C. Elder, H. Reinders (eds.) (2009). *Implicit and Explicit Knowledge in a Second Language*. Clevedon, Multilingual Matters (forthcoming).
Gu, Y. (2003). "Fine brush and freehand: the vocabulary-learning art of two successful Chinese EFL learners." *TESOL Quarterly* 37(1): 73-10.
Han, Z. (2004). Fossilisation in adult second langage acquisition. Clevedon, Multilingual Matters.
Johnson, J. and E. Newport (1989). Critical period effects in second language learning: the influence of maturational state on the acquisition of English as a second language. In M. Johnson (ed). *Brain Development and Cognition*. Oxford, Blackwell. 248-282.
Kellerman, E. (1985). If at first you do succeed . . . In S. Gass and C. Madden (eds.). *Input in second language acquisition*. Rowley: Newbury House. 345–353.
Lardiere, D. (2007). Ultimate Attainment in Second Language Acquisition: a Case Study. New York, Routledge.
Mackey, A., J. Philp, et al. (2002). Individual differences in working memory, noticing of interactional feedback and L2 development. In P. Robinson (ed.). *Individual differences and instructed language learning.* Philadelphia, John Benjamins: 181-209.
Mandell, P. (1999). On the reliability of grammaticality judgement tests in second language acquisition research. *Second Language Research* 15 (1): 73-99.
Moyer, A. (2004). *Age, Accent and Experience in Second Language Acquisition*. Clevedon, Multilingual Matters Ltd.
Myles, F. (2004). "From data to theory: the over-representation of linguistic knowledge in SLA." *Transactions of the Philological Society* 102(2): 139-168.
Paradis, M. (2004). *A Neurolinguistic Theory of Bilingualism*. Amsterdam/Philadelphia, John Benjamins.
Pienemann, M. (1998). *Language Processing and Second-Language Development: Processability Theory.* Amsterdam, John Benjamins.
Sanz, C. and K. Morgan-Short (2005). Explicitness in pedagogical interventions: input, practice and feedback. In C. Sanz (ed.). *Mind and context in adult second language acquisition.* Washington D.C.: Georgetown University Press. 234-263
Schachter, J. and V. Yip (1990). "Why does anyone object to subject extraction?" *Studies in Second Language Acquisition* 12: 379-392.
Sorace, A. 2003. Near-nativeness**.** In M. Long and C. Doughty (eds.), *Handbook of Second Language Acquisition*. Oxford: Blackwell. 130-152.

Vainikka, A. and M. Young-Scholten (1998). Morphosyntactic triggers in adult SLA. In M.-L. Beck (ed.). *Morphology and its interfaces in second language knowledge.* Amsterdam/Philadelphia: John Benjamins. 89-113.

White, L. and A. Juffs (1998). Constraints on wh-movement in two different contexts of nonnative language acquisition: competence and processing. In S. Flynn, G. Martohardjono and W. O'Neil (eds.). *The Generative Study of Second Language Acquisition*. Mahwah, NJ, L. Erlbaum: 111-129.

Wright, C. (2006). Is interlanguage variation in complex CP structures predictable? *Proceedings of the Fourth University of Cambridge Postgraduate Conference in Language Research.* Cambridge, UK: Cambridge Institute of Language Research. 277-284.

Appendix: sample sentences from the Reaction Time Task

1.	What did Tom buy at the shop?	*Simple Object*
2.	Who ate the cake with his fingers?	*Simple Subject*
3.	*What John eating at the party?	*Simple Object (ungrammatical)*
4.	*Who was arrive later by car?	*Simple Subject (ungrammatical)*
5.	Who do you suppose John wanted to marry?	*Complex Object (finite)*
6.	Who did Ann say liked her friend?	*Complex Subject (finite)*
7.	Who did Tom expect to beat Mary?	*Complex Object (non-finite)*
8.	Who did Ann want to win the game?	*Complex Subject (non-finite)*
9.	*What did books about make Ann happy?	*Complex Subject complement (ungrammatical)*
10.	*What did Mary see the card while John ate?	*Complex Adverbial (ungrammatical)*