## CONTENTS

<table>
<thead>
<tr>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Varieties of English: Taiwanese Attitudes and Perceptions</td>
<td>1</td>
</tr>
<tr>
<td><strong>Shouchun Chien</strong></td>
<td></td>
</tr>
<tr>
<td>Final Devoicing in Central Kurdish: An OT Analysis</td>
<td>17</td>
</tr>
<tr>
<td><strong>Tswana Hamid</strong></td>
<td></td>
</tr>
<tr>
<td>The Prosodic Marking of Givenness in English and Italian: A Comparative Study</td>
<td>28</td>
</tr>
<tr>
<td><strong>Katie Harris</strong></td>
<td></td>
</tr>
<tr>
<td>The Syntax and Semantics of Arabic Spatial Ps</td>
<td>44</td>
</tr>
<tr>
<td><strong>Sameerah Saeed</strong></td>
<td></td>
</tr>
<tr>
<td>Examining the Effects of Study Abroad on Mandarin Chinese Language Development among UK University Learners</td>
<td>67</td>
</tr>
<tr>
<td><strong>Clare Wright &amp; Cong Zhang</strong></td>
<td></td>
</tr>
</tbody>
</table>
EDITORIAL NOTE

Newcastle Working Papers in Linguistics aims to publicise the productivity of research of the Centre for Research in Linguistics and Language Sciences (CRiLLS) and external contributions. The papers collected in this volume have been contributed by staff and postgraduate students affiliated with Newcastle University as well as presenters at the 8th Newcastle Postgraduate Conference in Theoretical and Applied Linguistics.

We would like to thank all of the reviewers (listed below) for their valuable time and constructive comments. Without their detailed suggestions and remarks, it would have been impossible to maintain the high standards of this volume. We would also like to thank the members of CRiLLS for their support as well as all of the contributors to this volume.

Reviewers:

Jalal Al Tamimi
Laura Bailey
Alex Bellem
S J Hannahs
Anders Holmberg
Mimi Huang
Wyn Johnson

Ghada Khattab
Alex Ho-Cheong Leung
Robert M. McKenzie
Geoffrey Poole
Peter Sercombe
Thora Tenbrink
Martha Young-Scholten

The editors
VARIETIES OF ENGLISH: TAIWANESE ATTITUDES AND PERCEPTIONS

SHOUCHUN CHIEN
(University of Glasgow)

Abstract

Non-native speakers (NNS) of English outnumber native speakers (NS) in the 21st Century (Crystal 2003). This shift points to increasing opportunities for the speaker of English as a Foreign Language (EFL) to use English to communicate not only with NS, but also with NNS of English. How does this globalised phenomenon affect people’s attitudes towards the different varieties of English? This research examines the attitudes of Taiwanese people to different varieties of English: Australian English (AE), General American English (GAE), Indian English (IE), Japanese English (JE), Spanish English (SE), Standard Southern British English (SSBE) and Taiwanese English (TE). Analysis of 317 responses showed an overwhelming preference for the native variety GAE in terms of both status and solidarity, which might result from the fact that GAE is the most commonly used ELT model in Taiwan. Additionally, participants demonstrated different evaluations on the dimensions of status and solidarity: where in-group identity is concerned, TE was less stigmatized in terms of solidarity than status. This finding parallels the study of Garrett et al. (2003) who found that listener-judges seem to prefer their own variety on the solidarity dimension and native varieties with higher prestige in terms of status. I compare these findings to previous research and discuss what they might mean for the status of English in an increasingly globalized world.

1. Introduction

Situated in the ‘Expanding Circle’ of Kachru’s (1992) concentric circle model, English is used primarily as a Foreign Language (EFL) in Taiwan (Wu & Ke 2009). However, the emergence of English as a global language has created a huge demand for communication to take place in English. This induces the continual increase and frequent usage of English in Taiwanese society (Chen 2003). With the visible number of Non-native Speakers (NNS) of English now exceeding Native Speakers (NS) (Smith 1992, Crystal 2003), current globalisation has brought increased opportunities for Taiwanese EFL speakers to be exposed to different varieties of English. That is, EFL speakers are more likely to communicate in English with NNS than NS in Taiwan. How, and in what ways, might these globalised contexts affect Taiwanese people’s perceptions towards varieties of spoken English? With the application of the Verbal-Guise Test (VGT), this study will investigate attitudes towards a number of varieties of English in Taiwan.

According to Eagly & Chaiken’s (1993:1) definition, ‘attitude is a psychological tendency that is expressed by evaluating a particular entity with some degree of favour or disfavour’ and ‘evaluating refers to all classes of evaluative responding, whether overt or covert, cognitive, affective or behavioural’. Notably, Garrett (2010:23) highlights the essentialness of assessing attitude intensity (strength), which is the extent of negative or positive feelings of the attitude subject. As attitude intensity has an effect on judgement and behaviour (Perloff 2003:56 cited in McKenzie 2010:25), McKenzie (2010:25) emphasised that attitude studies should not only ‘identify one’s attitude towards an object but also to measure the intensity with which it is held’. In examining attitudes towards languages, many (Fasold 1984, Baker 1992, Trudgill 1992, Garrett 2010) have noted that language attitude...
studies are based on the belief that speech can evoke stereotyped reactions that reflect differentiated views of the language and its speakers. For instance, Giles & Billing (2004:202) contended that ‘listeners can very quickly stereotype another person’s personal and social attributes on the basis of language cues and in ways that appear to have crucial effects on important social decisions made about them’. Although language attitude is an ‘umbrella term’ which contains attitudes towards various aspects of the language, this study, with the aim of investigating Taiwanese attitudes towards varieties of English and the speakers who use them, concentrates on the perspective of ‘attitudes to language variation, dialect and speech style’, ‘attitudes to language groups, communities’ and ‘attitudes to language preference’ (Baker 1992:29).

The paper is divided into the following sections. Firstly, a review of previous research on attitudes towards the different varieties of English is introduced. The general research findings in this area will be covered, but with a main focus on findings relevant to the situation in Taiwan. Next, the methodology adopted will be explained. Lastly, the findings of this research will be discussed, along with further implications for English Language Teaching (ELT) to EFL learners in Taiwan.

2. Literature Review

The literature on attitudes towards the varieties of English is vast, but one of the main findings arising from this body of research is that native varieties tend to receive higher evaluations than their counterparts (Giles & Coupland 1991, Forde 1995, Dalton-Puffer et al. 1997, Garrett et al. 2003, Hiraga 2005, Kim 2007, McKenzie 2008, Zhang 2009). Turning to the situation in Taiwan, a number of studies (Kobayashi 2008, Wong 2011) have shown that Taiwanese listener-judges appear to prefer native varieties like American English over non-native varieties such as those spoken in the Philippines or India, despite the fact that non-native varieties are becoming more prevalent. This preference is also reflected in Taiwanese EFL speakers’ tendency to favour native English-speaking teachers over non-native English-speaking teachers in the ELT industry (Cheng 2009, Wu & Ke 2009).

The phenomenon of EFL speakers preferring ‘Inner Circle’ (IC) varieties could be explained by applying Kachru’s model (1992, 1997). Taiwan is situated in the Expanding Circle (EC) and thus the usage of English in Taiwan can be regarded as ‘norm-dependent’ (Kachru 1985:16-17). In Smith’s (2012:228) words, ‘English has been seen as having a different status from that in the inner and outer circles: learnt primarily as foreign language; and non-dependent, looking towards the inner circle for its linguistic models.’ Along with Kachruvian’s model (1992), a number of other models exist to explain English worldwide, including Moag’s Life Cycle Model (1992) and Schneider’s Dynamic Model of Postcolonial English (2007). Although Kachru’s Circle Model (1992) has been subject to criticism in recent years (e.g. Bruthiaux 2003, Jenkins 2009:17-18), for the present purposes, it provides an adequate description of Taiwan with respect to the use of English.

The research on Taiwanese students learning English in the Philippines, Kobayashi (2008) found that participants seem to regard English spoken in IC countries, such as North America, as the standard in the process of English acquisition. That is, Taiwanese learners perceive General/Standard American English as the sole learning target. In addition, with a lesser preference for non-native varieties such as Philippines English, Taiwanese participants indicated that their main purpose in learning English is to communicate with NS from the IC countries. Similarly, a study of Taiwanese EFL children’s perceptions towards teachers with native and non-native accents demonstrated that native English-speaking teachers who speak General American English are much preferred than non-native English teachers (Cheng 2009). This is supported by the finding that accent was ranked the second highest reason for
choosing a native English-speaking teacher. Interestingly, although elementary level students have only studied English for 3-4 years in school, they have already established their goal of adopting native-like accents. Parallel to the findings of Cheng (2009), when investigating Taiwanese university students’ attitudes, American English was much preferred to Indian English and Mandarin-accented English (Wong 2011). Furthermore, Liou’s (2010) research demonstrated that, for most Taiwanese English learners, the English of native speakers is much favoured even when the notion of English as an International Language (EIL) is concerned.

Additionally, the literature demonstrated the importance of conducting language attitude studies. For example, Ellis’s (2000) study on Second Language (L2) acquisition showed a positive correlation between attitudes towards the L2 and its speakers and the level of proficiency, which indicates that learner attitudes have an impact on L2 learning. According to Garrett (2001), studies of language attitudes are important as they account essentially for language maintenance and change within the sociolinguistic spectrum. Similarly, in the study of EFL learners’ attitudes to English in Brazil, Friedrich (2000) stressed the importance of language attitude studies with NNS. Specifically, as learners’ attitudes towards English involve not only ‘a set of formal features’ but also ‘feelings, stereotypes, expectations and prejudices’, they contribute essentially to the understanding of sociolinguistic parameters (Friedrich 2000:222). McKenzie (2010:73) has suggested that the results of language attitude studies are beneficial to the ‘choice of linguistic model’ for pedagogical application. Al-Tamimi & Shuib (2009:30) have argued that a better understanding of EFL students’ attitudes could help designers of the EFL curriculum to implement ELT programmes that would encourage the attitudes most beneficial to the success of EFL learners.

3. Research Question

Previous research (Kobayashi 2008, Cheng 2009, Wong 2011 among others) shows that Taiwanese speakers prefer native varieties of English when compared to non-native varieties. Given the on-going growth of exposure to non-native varieties in Taiwan, whether the attitudes of EFL speakers in Taiwan have changed is still uninvestigated. The increase in international communication makes varieties of English used widely, it is, therefore, essential to examine how they are perceived by NNS in Taiwan. This research discusses language attitudes towards the varieties of English in Taiwan and examines how different varieties of English are rated across different traits by Taiwanese NNS of English. In doing so, this study aims to answer the question ‘Do EFL speakers in Taiwan prefer certain varieties of English?’ Specifically, the responses of participants will help to demonstrate whether EFL speakers in Taiwan still prefer native varieties of English, as suggested in previous studies (Kobayashi 2008, Cheng 2009, Wong 2011), or whether there is an increasing acceptance of non-native varieties, along with a more positive perception of them.

4. Methodology

4.1. Research Instrument: Questionnaire

The outline of the questionnaire is composed of two sections. The first part of the questionnaire investigates participants’ nationality, gender and etc. The information offers a basic overview of each participant, which helps to look into the effects of certain social constraints on the variation. The second part is the semantic-differential scale of the VGT, which has been used successfully in studies that investigated attitudes towards several English varieties (cf. Bayard et al. 2001, Hiraga 2005, Kim 2007, McKenzie 2008, McGee 2009,
Zhang 2009 among others). According to Obiols (2002), VGT involves informants listening to recordings of natural speech from several anonymous speakers. They are then asked to evaluate each speaker separately on a bipolar semantic-differential scale. The application of VGT works as an indirect method to investigate participants’ language attitudes (Garrett et al. 2003).

4.2. Selection of Traits

Previous studies have suggested that it is important to examine whether the traits in VGT are appropriate in terms of meaning for informants in language attitude studies (Hiraga 2005, McKenzie 2008, Zhang 2009). The six traits of ‘confident’, ‘intelligent’, ‘educated’, ‘authoritative’, ‘friendly’ and ‘lively’ were used for their validity in a pilot study, in which feedback was given from Taiwanese students studying in Glasgow. Moreover, as the on-line questionnaire was conducted in Mandarin for Taiwanese participants, the variable of translating the selected traits into Mandarin was also considered.

In this study, informants were asked to rate each speaker individually, according to the six different traits on a six point semantic-differential scale (see Table 1 for details). This study adopts a six-point bipolar rating scale: Jenkins (2007:152) has argued that the use of an even number for the semantic-differential scale will help to ‘force respondents to evaluate each accent either positively or negatively and prevent them from adopting a neutral position’. By doing so, the study will indirectly get informants to state whether their evaluation of each speaker is closer to the negative or positive evaluation. For the statistical analysis, participants’ positive and negative evaluations on the six traits in the VGT were arranged by the same criteria: 1 is the least favourable evaluation, 6 is the most favourable evaluation.

Table 1 Verbal Guise Test (VGT): The six-point semantic differential scale

<table>
<thead>
<tr>
<th>Trait</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
</tr>
</thead>
<tbody>
<tr>
<td>Unconfident</td>
<td></td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td>6</td>
</tr>
<tr>
<td>Unintelligent</td>
<td></td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td>6</td>
</tr>
<tr>
<td>Uneducated</td>
<td></td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td>6</td>
</tr>
<tr>
<td>Not Authoritative</td>
<td></td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td>6</td>
</tr>
<tr>
<td>Unfriendly</td>
<td></td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td>6</td>
</tr>
<tr>
<td>Boring</td>
<td></td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td>6</td>
</tr>
</tbody>
</table>

Confident
Intelligent
Educated
Authoritative
Friendly
Lively

4.3. Selection of English Varieties

According to Kachru’s three-circle model that represents the spread of English (1992, 1997), this study aims to be inclusive in that it contains both native and non-native varieties in order to give a relatively more comprehensive picture of Taiwanese EFL speakers’ attitudes towards varieties of English. Kachru’s model (1992, 1997) is composed of the ‘Inner Circle’ (IC), known as the circle of ‘norm-providing’ (Kachru 1985:16-17), and includes countries such as the USA or UK, which use English as a native language. These ‘norm-providing’ varieties, American and British English, are accepted as the mandatory ‘standard’ for ELT worldwide (Kachru & Nelson 2006:12). Secondly, in the ‘Outer Circle’ (OC), the ‘norm-developing’ world (Kachru1985:16-17), English is spoken as a second language (ESL). Lastly, the ‘Expanding Circle’ (EC) can be seen as the ‘norm-dependent’ context. The rationale for the selection of each variety was as follows.

Standard Southern British English (SSBE) was chosen for inclusion as one of the reference varieties in the IC instead of Received Pronunciation (RP), spoken by elite social
classes who constitute around three percent of the British population (McArthur 1992:851). The variety of SSBE in this study refers to the broader ‘educated pronunciation’ of British English spoken in southern England that has been widely taught (Kachru & Nelson 2006:94, Bieswanger 2008:30). Since it is highly rated in terms of status (e.g. Ladegaard 1998, Ladegaard & Sachdev 2006), it is worth examining how Taiwanese listeners judge this variety in the study. It was important to include General American English (GAE), as it is the most taught model in the Taiwanese EFL classroom (Cheng 2009, Liou 2010). It will also offer potential comparisons with previous studies, which have demonstrated that Taiwanese EFL speakers have a strong preference for American English (Cheng 2009, Wong 2011, Kobayashi 2012). Australian English (AE) is also classified as a variety from the IC. AE was chosen as it has been perceived as a relatively ‘less standard variety’ among NS countries (Jenkins 2007:150). Therefore, the inclusion of AE will examine how EFL informants evaluate it when compared to GAE and SSBE. Indian English (IE) is the only OC variety of this study. However, Wong (2011) suggests that there is a large population of ESL speakers of IE; it is essential to include IE in the VGT. David Graddol once maintained that ‘no other region has been more affected by the rise of English than Europe’ (2001:47 cited in Modiano 2009:223); the inclusion of Spanish English serves the role of incorporating an English variety spoken in an European country into this experiment as English becomes a lingua franca in Europe (Seidlhofer et al. 2006, Modiano 2009). As the variety of Japanese English (JE) has been widely studied in language attitude studies (e.g. Chiba et al. 1995, McKenzie 2008, Eisenchlas & Tsurutani 2011), JE is chosen so that comparison can be made with previous studies. The inclusion of JE in this study will also offer an insight into how Taiwanese NNS perceive this non-native variety, which is closest to them geographically. Finally, the Taiwanese variety (TE) is chosen with the purpose of discovering how participants perceive their own variety (cf. Cheng 2009, Wong 2011). For instance, Jenkins (2007) demonstrated that, with the exception of their own variety, research participants tend to hold negative perceptions towards NNS of English.

4.4. Selection of Text

The stimuli materials used in this study are taken from the Speech Accent Archive (Weinberger 2012). They are freely available and have been used in other language attitude studies (Hiraga 2005, Cheng 2009, Yook & Lindemann 2013), which provide a possibility of comparing it with previous studies. Moreover, Cheng (2009:32) has suggested that the stimulus is designed phonetically and consists of common English words but with a variety of difficult English sounds and sound sequences. This design helps to reveal each speaker’s original accent and thus make participants’ ratings of the speakers more valid. Furthermore, the text is of a suitable length for the Taiwanese listener-judges, which made the task more manageable as they will have to listen to seven recordings. The text of the passage from the Speech Accent Archive is as follows:

Please call Stella. Ask her to bring these things with her from the store: Six spoons of fresh snow peas, five thick slabs of blue cheese, and maybe a snack for her brother Bob. We also need a small plastic snake and a big toy frog for the kids. She can scoop these things into three red bags, and we will go meet her Wednesday at the train station.

Speech Accent Archive (Weinberger 2012)
4.5. Selection of Recordings

The Speech Accent Archive offers a large selection of speech recordings from various language backgrounds (Hiraga 2005). It provides detailed biographical data of each recorded sample and categorises them according to speaker’s native languages. Moreover, both NS and NNS of English read the same elicitation paragraph (this paragraph is reproduced in 4.4) and their recordings are phonetically transcribed into consonant, vowel and syllable structure. Cheng (2009:31) suggests ‘researchers can use the recording samples in the archive to compare and analyse the accents of different English speakers’. However, although VGT enables each authentic speaker to represent each guise, researchers have less control over other speech-related variables such as voice volume and speed. In recognition of this potential issue and following on from other studies (Hiraga 2005, McKenzie 2008, Zhang 2013), this experiment aims to select speakers who are roughly similar in voice quality to avoid bias.

Most of the recordings for this research were taken from the Speech Accent Archive website¹. The only exception was the IE variety: the recordings in the Speech Accent Archive were unsuitable, as the speakers had all lived away from India for some time. Therefore, a recording of the same text by an IE speaker was recorded in the sound laboratory of the English Language department at the University of Glasgow. This ensured uniformity of voice quality of all recordings used in this study. All recordings were male voice, to avoid any ‘gender-linked language effect’ as suggested by Lambert (1967) and to prevent the possibility of listener-judges’ different reactions to male and female voices (Bradac et al. 2001, Zhang 2009). In addition to the seven recordings representing the seven different English varieties, an extra recording was used as a practice example of VGT. The purpose of the practice example is to help informants familiarise themselves with the instructions of the VGT. Therefore, in the practice example of the VGT, the recording is selected as a female SSBE voice to differentiate from the male voice in the other seven recordings. The recording for the practice example was also made in the same sound laboratory in Glasgow.

4.6. Research Informants

As McKenzie (2010:84) has suggested, a large number of participants in language attitude studies make research findings more likely to be ‘representative of the target population’. In order to conduct multiple surveys efficiently and effectively, questionnaires were collected through a website that was designed and administered by the researcher during a three month period from June to September 2013. Student participants were mainly contacted through lecturers or teachers in Taiwanese colleges, universities and schools. Additionally, employed participants were contacted by a ‘friend-of-a-friend’ method (Milroy 1980). There were 317 valid questionnaires returned in total, 194 are from students and 123 from adults under employment.

4.7. Data Collection Procedure

Before answering the questionnaire, participants were given a brief introduction to the study. On an initial ethics consent page, participants were informed that the study would be completely anonymous. Participants were asked to provide background information in a basic information section before the start of the listening task of the VGT. Finally, the VGT section was divided up, with the practice example followed by the seven different speakers. Participants were given a practice example showing how to complete the evaluations on the

¹ http://accent.gmu.edu/
semantic-differential scale and were told of the importance of giving their intuitive impression when evaluating these seven different accents and there was no right or wrong answer. Informants were asked to listen to each recording once. When listening to the recording, informants were asked to rate the speaker on the VGT scale. Repetition of the recording was controlled on the website—it could only be played once: this forced listener-judges to give a spontaneous response. However, as participants clicked on ‘next page’ to continue answering the questionnaire at their own speed, ample time could be taken between the seven different recordings.

4.8. Data Analysis

The data was analysed using SPSS Version 20.0 (Statistical Package for the Social Sciences). Following previous research (e.g. Hiraga 2005, Kim 2007, McKenzie 2010, Yook & Lindemann 2013), this study also conducted a one-way repeated measure analysis of variance (ANOVA) test to assess the difference between the mean ratings of the VGT to gain an overall view of the participants’ opinions towards different varieties of English (See section 5 for details). When a repeated measures analysis is undertaken, SPSS executes ‘Mauchly’s test’, which tests the sphericity assumption that these variances are not significantly different (Kerr et al. 2002:120-121, Pallant 2010:280). If the significance value associated with Mauchly’s test is greater than 0.05, the null hypothesis is accepted and it is concluded that the sphericity assumption is met (Kerr et al. 2002:121).

5. Results and Discussion
5.1. Preliminary Analysis

The first step in analysing the results of the VGT is to carry out a Principal Component Analysis (PCA) to examine whether the traits are clustered into different groups (Bayard et al. 2001, McKenzie 2008, Eisenchlas & Tsurutani 2011). According to Miller et al. (2002:174-175), PCA is a ‘data reduction’ technique that aims to condense a larger set of variables down to a smaller number of components. The components extracted thus summarise the ‘common features’ between the variables within a dataset. Following Kaiser’s criterion (Pallant 2010:184), the analysis demonstrated the loading of 6 traits on the two components with eigenvalues greater than 1.0 which together account for 81% of the variance.

Table 2 Varimax-rotated factor analysis with Kaiser normalization (N=317; loadings less than 0.5 are not listed)

<table>
<thead>
<tr>
<th>Traits</th>
<th>Component 1 (Status 64%)</th>
<th>Component 2 (Solidarity 17%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>confident</td>
<td>0.844</td>
<td></td>
</tr>
<tr>
<td>intelligent</td>
<td>0.885</td>
<td></td>
</tr>
<tr>
<td>educated</td>
<td>0.862</td>
<td></td>
</tr>
<tr>
<td>authoritative</td>
<td>0.771</td>
<td>0.858</td>
</tr>
<tr>
<td>friendly</td>
<td></td>
<td>0.927</td>
</tr>
<tr>
<td>lively</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 2 revealed that the traits ‘confident’, ‘intelligent’, ‘educated’ and ‘authoritative’ formed the ‘status’ group and the traits ‘friendly’ and ‘lively’ clustered in a ‘solidarity’ group. The result further impacted on the following ANOVA analysis in the sense that participants’
evaluations of each speaker are examined according to the two dimensions of ‘status’ and ‘solidarity’.

### 5.2. Evaluations according to Speaker Status and Speaker Solidarity

The rankings of the mean ratings in descending order of the seven English varieties according to speaker status and solidarity are shown in Table 3. While the status dimension consists of ‘confident’, ‘intelligent’, ‘educated’ and ‘authoritative’, the solidarity dimension is composed of ‘friendly’ and ‘lively’. A higher number indicates a greater degree of favour for each variety. Varieties of English, shown in bold, indicate where there are significant differences ($p<0.05$) between the participants’ ratings.

#### Table 3 Rankings according to speaker status and solidarity (in descending order of mean evaluation; N=317)

<table>
<thead>
<tr>
<th>Status (Confident &amp; Intelligent &amp; Educated &amp; Authoritative)</th>
<th>Solidarity (Friendly &amp; Lively)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Variety of English Mean</td>
<td>Variety of English Mean</td>
</tr>
<tr>
<td>GAE</td>
<td>4.29</td>
</tr>
<tr>
<td>SSBE</td>
<td>4.07</td>
</tr>
<tr>
<td>IE</td>
<td>3.81</td>
</tr>
<tr>
<td>AE</td>
<td>3.65</td>
</tr>
<tr>
<td>JE</td>
<td>3.12</td>
</tr>
<tr>
<td>SE</td>
<td>3.01</td>
</tr>
<tr>
<td>TE</td>
<td>2.91</td>
</tr>
</tbody>
</table>

In terms of the status dimension, participants rated standard varieties from the IC (GAE, SSBE) more positively than the OC variety (IE) and the NS variety (AE) or those varieties spoken at the EC (JE, SE and TE). The result of a one-way repeated measure of ANOVA demonstrated that Mauchlay’s Test=0.821: consequently sphericity was assumed. ANOVA results indicated that there were significant differences between informants’ ratings of the seven speakers’ status: $F (6, 1896)=243.816, p<0.05$. When the mean evaluations for speaker status were analysed in the pairwise comparison test (with Bonferroni correction), the results showed speakers of GAE, SSBE, IE and AE reached statistical significance ($p<0.05$; shown in bold, in Table 3). In other words, no significant difference was found in the mean rating of speaker status between JE&SE and SE&TE, which indicated that Taiwanese informants seem to make consistently low evaluations of these varieties.

Consequently, participants’ evaluations of the seven varieties can be divided into three distinct hierarchies. Speakers of the prestigious NS varieties (GAE, SSBE) are much preferred, followed by the variety of the OC (IE) and the variety of the IC (AE), with the NNS varieties (JE, SE and TE) least preferred. This result is similar to that of previous studies (Kobayashi 2008, Cheng 2009, Wong 2011), which demonstrated that Taiwanese participants prefer IC varieties over NNS varieties, with particular preference for GAE. As American and British English of the IC are commonly regarded as ‘legitimate’ varieties for educational purposes, this might lead to participants’ preference for GAE and SSBE in terms of status (Kachru & Nelson 2006:12). The strong preference for GAE might also be affected by the ‘high vitality of American English’ suggested by Ladegaard & Sachdev (2006:105-106). Similarly, it could be explained that familiarity with GAE as the ELT model (Cheng 2009, Liou 2010) in
pedagogical settings and increased media exposure to GAE (Bayard et al. 2001) further contribute to participants’ high evaluations. This result also paralleled the argument of Dalton-Puffer et al. (1997) that students who have longer exposure, and thus greater familiarity with specific English varieties, tend to give a relatively positive evaluation of this variety.

Notably, participants’ own variant (TE) received the lowest evaluation. This is similar to previous findings by Dalton-Puffer et al. (1997) in Austria and McKenzie (2008) in Japan, which show that the least favoured is Taiwanese participants’ own variety. This might result from the notion of TE or speaking English with their own accent as having low status, which had been instilled by ‘media-transmitted stereotypes’ and the influence of the ELT classroom setting, as suggested by McKenzie (2008:74-75). That is, Taiwanese EFL speakers were inclined to avoid their own accent when speaking or learning English and thus aimed to pursue the prestigious varieties of the IC, especially American English, which had become the goal of English acquisition.

With regard to the solidarity dimension, the ANOVA results indicated Mauchlay’s Test=0.868: consequently, sphericity was assumed. There were significant differences between the informants’ ratings of the seven speakers’ solidarity: F (6, 1896)=41.295, p<0.05. When the mean evaluations for speaker solidarity were analysed in the pairwise comparison test (with Bonferroni correction), the results demonstrated that there were significant differences between GAE&IE speakers as well as JE&SE speakers (p<0.05; shown in bold, in Table 3). In other words, no significant differences were found in IE, AE, SSBE, TE or SE, which indicates that participants shared similar perceptions in terms of solidarity among these varieties.

Participants’ evaluations of the seven varieties could be further divided into three distinct hierarchies. Firstly, the finding that an NS variety (GAE) instead of an NNS variety was the most highly rated is in contrast to the study of McKenzie (2008) in Japan but similar to the finding of Zhang (2013) in Hong Kong. It might result from ‘the strong inclination and the pressure towards conformity to model native English - such as American English’ found in Zhang (2013:13). In other words, this inevitably affected participants’ high evaluations of the prestige varieties of GAE and the relatively less positive attitude towards participants’ own accent. Secondly, the NNS variety of IE and the NS variety of AE were rated more highly than the prestigious IC variety of SSBE. This is in direct contrast to the result on the status dimension. This finding seems to confirm that of previous studies (Giles 1970, Bayard et al. 2001, Hiraga 2005) in which the evidence suggested that non-standard varieties are usually preferred to standard varieties when solidarity is concerned. Finally, varieties of TE, JE and SE of the EC were the lowest rated groups on the solidarity dimension. Lindemann’s finding (2003), whereby there tends to be a close connection between participants’ attitudes towards the speakers and speakers’ accents, might help to explain the low rating these ‘accented’ varieties (TE, JE and SE) received. Nevertheless, it is worth noting that, unlike its evaluations on the status dimensions, TE was not the least favoured variety in terms of solidarity when compared to the rest of the EC varieties. Similar to the finding of McKenzie (2008), this may be a result of informants’ tendency to prefer, or give higher ratings to, their own varieties where the high degree of solidarity and the marker of self-identity are considered.

In summary, except for the overwhelming preference for GAE on both dimensions, participants demonstrated different attitudes on the dimensions of status and solidarity. This result revealed that, while participants preferred the standard native variety more when status is concerned, non-native or non-standard varieties received higher evaluations on the dimension of solidarity.
6. Discussion and Conclusion

In this research, the language attitudes of 317 Taiwanese participants towards native and non-native varieties of English were investigated. In general, NS varieties received higher evaluations than the less prestigious NNS varieties. For instance, the findings revealed that participants hold an overwhelming preference for GAE over EC varieties of JE, SE and TE. Nevertheless, participants demonstrated different evaluations on the dimensions of status and solidarity. For example, when in-group identity is concerned, TE was rated higher in terms of solidarity than status. Moreover, while SSBE received a higher rating than IE and AE in terms of status, it received lower evaluations in solidarity dimension than the NNS variety of IE and the NS variety which is ‘still trying to achieve legitimacy’ – AE (Kachru & Nelson 2006:12). This finding is similar to the previous studies of Giles (1970) and Hiraga (2005), which showed that, when compared to the status dimension, listener-judges rate standard varieties of English lower on the solidarity dimension. Moreover, the positive evaluation IE received in terms of both status and solidarity demonstrated an open acceptance towards non-standard varieties of the OC. Parallel to the study of Eisenchlas & Tsurutani (2011) among native speakers of Australian English, this finding further indicates an increased positive attitude towards NNS varieties of English on the part of participants. The increasing self-confidence and independence of varieties like Australian English and Indian English are possibly ‘promoted by systematic linguistic descriptions and might reduce the influence of the traditional norm of standard native varieties’ (Bieswanger 2008:32).

Nevertheless, limitations are present in interpreting the results based on the VGT. Bias of this study exists in disentangling what it is exactly that listeners are reacting to when they hear different varieties of English speech. Studies have shown that accent may not be the only factor in forming listeners’ attitudes: other variables such as context, the level of speaker’s accent, fluency and message content should be taken into consideration (Cargile 1997, Cargile & Giles 1998). Moreover, possible effects of speakers’ voices on attitudes should not be underestimated. Although the recordings of the speakers representing each variety had been prudently selected, it is still challenging to control ‘prosodic and paralinguistic features of voice (such as pitch, voice quality and speech rate), as well as other aspects of reading style and expressiveness’ (Giles & Coupland 1991:34). Therefore, it would be profitable for future research to include more than one speaker to test whether the accent is representative of speakers of these varieties. As a result, the influence of the voice features and the bias of relying on one speech sample to represent a particular variety can be eliminated (Kim 2007, Zhang 2009).

Kachru & Nelson (2006: 126) once argued that the notion of IC varieties being ‘better’ than non-IC varieties is now ‘empirically invalid’ and perhaps native varieties need no longer be the sole standard in English acquisition. Researchers (Smith 1992, Friedrich 2000, Matsuda 2003, McKay 2003, Deterding 2005, Eisenchlas & Tsurutani 2011) have shared similar views and maintained the importance of equipping EFL speakers with an awareness of different English varieties as this will allow them to cope with the cultural and linguistic bias that appears in both native and non-native varieties of English. Consequently, it is essential that EFL speakers be exposed to varieties of English as increasing exposure would result in familiarity and awareness of linguistic features and cultural information of different forms of English speech that would further contribute to effective international communication (Kirkpatrick 2007). This may be of particular relevance to language planners and educators in Taiwan with respect to the specific choice of linguistic models employed for ELT (Kobayashi 2012). For Taiwanese EFL speakers to communicate successfully with NS and NNS of different ‘regional, social and cultural backgrounds’ and to be ‘linguistically, sociolinguistically and pragmatically’ equipped through exposure to different varieties of
English (Bieswanger 2007:405), the findings of this study might be applicable to the development of EFL learning programmes from the perspectives of pedagogical materials (Matsuda 2003), curriculum design (Deterding 2005, Bieswanger 2008, Wong 2011), sociolinguistic training for teachers (Norrish 1997) and teacher recruitment (Kirkpatrick 2007), etc.

In terms of teaching materials, teachers can introduce speakers of varieties of English by showing movies or video clips to students (Matsuda 2003). As informants of this study demonstrated positive evaluations of AE and IE, which are not the mainstream varieties being applied in schools, it is perhaps worthwhile introducing the linguistic and pragmatic norms of AE and IE to Taiwanese EFL learners. Applying the notion to curriculum design, Wong (2011) demonstrated that, after delivering a lecture addressing the variety of Indian English from a World English perspective, the course had a positive influence on Taiwanese university students’ perceptions of Indian English. This indicates that positive perceptions of different forms of English speech can be achieved through increasing their use in a classroom setting. Moreover, to provide learners of English with an education that enables optimum communicative competence, the incorporation of varieties of English with a particular focus on the linguistic analysis of pronunciation in ELT appears to be vital (Bieswanger 2008:33). For example, Görslach (1999:18 cited in Bieswanger 2008:33) maintained that learners of English, at an early stage, should be exposed to as many accents as possible, with a broad receptive training. By the same token, Deterding’s study (2005) in Singapore suggested that exposing learners to different varieties of English such as Estuary English, which has become more widespread in southern England, would help students to become familiar with various styles of pronunciation. In Taiwan, the majority of the listening materials applied American accents (Jou 2010:7-8). The findings of this study exhibited that, except for GAE, positive attitudes were shown towards other varieties such as SSBE, IE and AE, which suggests the importance of introducing a variety of different accents for Taiwanese EFL students in listening exercises.

In addition, many (Kachru & Nelson 2006, Bieswanger 2008) have identified the need to address the issues related to the varieties of English in training programmes for both prospective and active teachers. Taking the teacher training programme TESOL in Malaysia for example, an understanding of the forms of Malaysian English and its development in terms of context and users is suggested (Norrish 1997). It could be applied to teacher training in Taiwan, with a focus on discussion of the local forms of Taiwanese English since this might be the variant most used by students. Moreover, Kirkpatrick (2007:196) proposed that teachers wishing to teach non-native English speakers in Asia should be able to ‘analyse the specific linguistic difficulties that speakers of the non-Indo-European languages can face in the learning of English and are able to adopt strategies to help their learners overcome these difficulties’. In this case, since Taiwanese ELT teachers have experience of learning English as a foreign language, they would be better placed than native English speaking teachers in understanding the English learning problems that a student might encounter (Kirkpatrick 2007:190). Furthermore, the findings of this study demonstrate that non-native varieties such as IE and TE were better evaluated in the solidarity than the status dimension, which implies the potential acceptance of ELT teachers from the OC countries such as India where English is used dominantly as a L2. Therefore, the belief of ‘native speaker fallacy’ should be cautiously addressed when teacher recruitment is taking place in Taiwan (Phillipson 1992).

In conclusion, this study helps to give insight into Taiwanese attitudes towards different English varieties. The findings of this research are likely to provide implications for the language planning decisions with respect to how varieties of spoken English are perceived by Taiwanese English learners as well as future research in understanding the sociolinguistic repertoire of English in the EFL context of Taiwan.
References


Chien Varieties of English: Taiwanese Attitudes and Perceptions


ShouChun Chien
School of Critical Studies: English Language
University of Glasgow
G12 8QH
United Kingdom

s.chien.1@research.gla.ac.uk
**FINAL DEVOICING IN CENTRAL KURDISH: AN OT ANALYSIS**

TWANA HAMID  
(Newcastle University)

**Abstract**

As Selkirk (1986) states, final devoicing is a pattern of phonological distribution in which both voiced and voiceless obstruents occur in a language, but at the end of a particular prosodic domain only voiceless obstruents occur. The Phonological process of final devoicing has been a well-studied topic dating back to Trubetzkoy in 1933 (translated published in English in 1969) and it has been cited by most phonologists as an example of neutralisation (Brockhaus 1991 inter alia). Many languages disfavour coda voiced phonemes; however, devoicing is subject to parametric variation (Myers 2012). Most languages resolve voiced coda in a similar fashion; by devoicing rather than nasalization, deletion, or epenthesis.

This paper argues that Central Kurdish (henceforth CK) is one of the languages that undergoes final devoicing; it also claims that Kurdish resolves coda voicing by devoicing the voiced coda rather than other means mentioned above. For example, underlying /bəɾd/ ‘stone’ can become [bəɾt] but not *[bəɾ], *[bəɾm], *[bəɾdə]. Further, my analysis makes two theoretical claims: First, the prosodic domain within which coda devoicing occurs in CK includes both syllable and Prosodic word. Second, coda devoicing in CK, like many other languages (Steriade 2001/2008), can pose a hitherto unsolved problem to Optimality Theory. With the total absence of literature on this topic, I will recourse to the intuition of native speakers’ judgement to support the claim that final devoicing is occurring in CK.

1. **Introduction**

The Phonological process of final devoicing has been a well-studied topic dating back to Trubetzkoy in 1933 and it has been cited by most phonologists as an example of neutralisation (Brockhaus 1991). Selkirk (1986) states that final devoicing is a pattern of phonological distribution in which both voiced and voiceless obstruents occur in a language, but at the end of a particular prosodic domain only voiceless obstruents occur. Thus, final devoicing is meant to occur at the end of both syllables and words and it is also limited to obstruents which have voiceless counterparts in the phonemic inventory of the language.

Nevertheless, to the best of the author’s knowledge, no research has ever tried to look into the existence of this phenomenon in CK and there is no CK corpus to be used in search for this phenomenon. I observed the occurrence of final devoicing both in syllable and word final positions and noticed that CK has no voicing contrast in final obstruents, neutralising it towards voicelessness. With the total absence of literature on this topic, I will recourse to native speakers’ intuition judgement on the occurrence of final devoicing. Constraint-based analysis is used to examine the data. The present essay has two main goals: The first is to examine final devoicing in CK and look into its properties. The second is to use Optimality Theory for the analysis and see how it works. This paper falls into the following sections:

2. Discussion of the Data  
3. Descriptive Generalisations  
4. Constraint Interaction  
5. Devoicing as a problem for OT  
6. Conclusions
2. Discussion of the Data

There is no general consensus among phonologists concerning the causes of final devoicing which is quite common among the world’s languages including German, Dutch, Polish, Turkish and Russian, among others (Brockhaus 1991). Earlier accounts of devoicing attributed it to fortition or strengthening, while most recent literature regard it as an instance of weakening (see Harris 2009; Crystal 2008). Iverson & Salmons (2007) use two reasons for regarding final devoicing as fortition: One is based on the observation that obstruent voicing commonly occurs inter-vocally, is clearly a weakening process to assimilate the obstruent consonants with the surrounding vowels. The other argument takes the form of a claim that devoicing strengthens final obstruents in order to demarcate the right edge of words. Harris (2009) uses the same two reasons to argue against the claims that regard final devoicing as fortition in the sense that it becomes more consonantal and turns less sonorous than the underlying voiced consonants.

In the following list of words, the final voiced obstruent phoneme is devoiced (realised and heard by native speakers) as the voiceless counterpart of the same phoneme /b/ to [p], /d/ to [t] and so on (Note that the examples include all the obstruents: plosives and fricatives except [h] and affricates).

(1) Set of Data for CK Final Devoicing

<table>
<thead>
<tr>
<th>1. a</th>
<th>1. b</th>
<th>1. c</th>
</tr>
</thead>
<tbody>
<tr>
<td>[bara:s]</td>
<td>‘pig’</td>
<td>[bara:za:kan]</td>
</tr>
<tr>
<td>[gi:ɾ]</td>
<td>‘tension’</td>
<td>[giɾiz i w ʔa:lzoɾi]</td>
</tr>
<tr>
<td>[ʔaːzaɾ]</td>
<td>‘free’</td>
<td>[ʔaːza:diɾ]</td>
</tr>
</tbody>
</table>

In the examples of (1a) the devoicing includes word final devoicing as (either the words are mono-syllables or the non-final syllables do not give evidence for syllable final devoicing). (1b) shows that the obstruents regain their voicing when they are not final (either followed by another consonant or a vowel. In this case, as Wolf (2008) states, there is no reason to think that any language would only allow devoicing intervocally (given that intervocalic voicing is a typologically common process). Examples of (1b) represent only syllable final (coda) devoicing. Following Kager (1999), I will use coda devoicing instead of word final devoicing as coda devoicing captures greater generalisation. Unlike Japanese, which only allows coda consonants if they share a place of articulation with the onset of the following syllable, consonantal place features of coda in CK can be different from the place of articulation of the following onset (see 1c). Thus, the Coda Condition Constraint is ranked low and violable in CK. Languages vary with regard to syllable final or word final devoicing. While in Slavic, Romance, Germanic, Basque and many others only word-final obstruents are voiceless, in Thai, Vietnamese, Turkish, Malay and many others the syllable coda is devoiced (Myers 2012).

Looking for counterevidence to our analysis, one may look at cross-linguistic data to be cautious of possible pitfalls. For instance, in Dutch, final devoicing only occur within individual words while when we look at phrases, the final obstruents are re-voiced through voicing assimilation of the neighbouring phonemes. In CK, however, the phonological difference between single words and phrases is often not visible for these two reasons: The first
one is that, CK is a morphologically rich language and the different suffixes are mostly null, that is, not affecting the structure of the root word. Owing to this factor, most morphological phrases are phonologically single words, for example, the phrase /mirdl/ ‘he died’ undergoes devoicing [mirt] as any single word. The second factor seems to be that CK syllables are not complex and the syllables strongly abide by onset maximisation in a way that they turn the final obstruent into the onset of the next syllable as in gir:ziw. ʔa:.lb. zi ‘tension and unstable’ which comes from *girf iw ʔa:lbzi when the devoicing occurs before syllabification. Nevertheless, this does not mean voicing assimilation is totally absent in CK. The phoneme /kl/ in the word jak.dʒar ‘once’ is assimilated to /g/ resulting in [jag.dʒar].

Coda devoicing should be kept apart from assimilation of final consonant. In a language like English, for example, certain sounds (voiced fricatives in this case) change into other sounds (voiceless fricatives) under the influence of adjacent sounds (Spencer 1996) as in (2):

(2) a. five past /faɪvpɑːst/ → [faɪpɑːst]
    b. love to go /laʊtvəɡ/ → [laʊtəɡʊ]
    c. breath slowly /briːðsləʊlɪ/ → [briːðsləʊ.lɪ]

In the above examples, the final obstruents undergo regressive assimilation: the final voiced obstruent is devoiced under the influence of the following voiceless phonemes. However, in coda devoicing, in contrast to assimilation, the voiced obstruent is devoiced without being affected by adjacent phonemes. It can be said that final devoicing is both context-free and context-sensitive: it is context free in the sense that the devoicing is not triggered by the adjacent sounds and it is context-sensitive in the sense that only in coda position the voicing is neutralised.

As mentioned in the introduction, no previous studies are available to support the occurrence of coda devoicing in CK. So, solid evidence is needed to prove this point. I will approach this issue using evidences from both native speakers’ judgement and phonological argument. For speakers’ judgement, 10 native central CK adult speakers (7 male) were consulted to decide on whether they perceive the final consonant in (1a) as devoiced and the ones in (1b) regain the voicing. The speakers’ perceptions were also consulted about the devoicing of syllable final consonants. All the speakers acknowledged devoicing of the coda consonants in (1a) and (1c) and voice regaining for (1b). However, they also noticed that they perceive a devoiced coda as a different phoneme from an underlyingly voiceless consonant. This supports the fact that the voiced and voiceless consonants are underlyingly different and behave accordingly. Voiced consonants, for example, cannot shorten the preceding vowels in contrast to underlyingly voiceless phonemes which shorten the preceding vowels.

Sometimes speaker judgement is not categorical or even may be misleading. For instance in the examples given above, it can be argued that the underlying phonemes in the first place are actually voiceless. Hence, a competing analysis arises as a result of two competing rules: The first one, coda voiced obstruents undergo devoicing as this paper claims while according to the second analysis, underlying voiceless obstruents surface as its voiced counterpart before a vowel. As far as the list of words in (1) is concerned, both of them are possible. To solve this enigma, as Davenport & Hannahs (2010) state, when more than one analysis is possible for a set of data, it is one of the tasks of the phonologists to evaluate competing analysis and choose between them.

Looking a bit further for more evidence in CK, many words like [twpaka] ‘the ball’, [kɪl k i rɛwɪ] ‘tail of fox’ [dasɪm] ‘my hand’, [kafə] ‘it is foam’ can be found. Words like these and multiple other examples can be used as a counterargument for the second analysis and at the same time they can be used as an admissible evidence to support position one, i.e. an underlying voiced obstruent becomes voiceless at the end of a word. If position two were

Newcastle and Northumbria Working Papers in Linguistics 20, 2014

Hamid
correct, these words would have to appear as *[tvbaka], *[kilg i rewi], *[dazi min], *[kava], but voiceless consonants never surface as their voiced counterparts intervocally as far as I am aware. So, underlying voiceless consonants tend not to undergo voicing assimilation. Rather, it is the underlying voiced consonants that undergo devoicing in coda position as in (1a and 1c). Examining the full set of CK data we never find words like *[bard], i.e. a word ends with and pronounced with a voiced stop, although there are CK words which have voiced coda obstruents in the spelling.

Thus far, perceptually and phonologically, it has been established that coda devoicing is really attested in CK. This achievement helps us to account for the correct input and output in the next section when through constraints different competing candidates are evaluated and the optimal output is chosen. In the next section, an OT analysis is developed to account for the set of data presented in this section to account for the coda devoicing of obstruents in CK.

3. Descriptive Generalisations

In rule-based phonology, there is no correlation between the rules as rules need only concern themselves with their immediate outputs; they are blind to the overall actions of other rules and to the ultimate outcome of their own application. Constraint-based analysis, on the other hand, considers the broader implications of any specific operation or the effects of any particular process (Davenport & Hannahs 2010).

In phonological alternations, it is inevitable that faithfulness constraints should conflict with markedness constraints. In our case, devoicing in CK seems to be clear that no phonemes are deleted or epenthesised, but rather, a feature is changed and thus IDENT (F) is violated. However, thinking in terms of markedness constraints, the task seems to be daunting and not as easy as it looks from the beginning. Any faithfulness constraint that is violated must be dominated at least by one markedness constraint.

Markedness constraints can only do one thing: prohibit certain output configurations. But this command is too general to be understood or applied easily on phonological alternations as Gen can produce any candidate. Chomsky (1995) criticises this aspect of OT stating that “there seems to be no barrier to the conclusion that all lexical inputs yield a single phonetic output, namely, whatever the optimal syllable might be (perhaps /ba/).”

Chomsky was looking at one side of OT in his assessment. It is true that Richness of the base and Freedom of Analysis (two properties of OT) allow all lexical items to produce anything including /ba/, but he ignored two important points: the first one is that Gen can produce any item from the input before the Eval where through a set of ranked constraints, output candidates are evaluated to their harmonic values and the optimal one is selected. The second point, though more crucial, is that Chomsky does not take into consideration lexicon optimization which roughly states that underlying forms should always match surface forms in the absence of evidence to the contrary. However, lexicon optimization should not be understood in a way that all optimal candidates are faithful to input.

For the coda devoicing in CK, the following descriptive generalisations are proposed:

(3) a. No voiced obstruent in coda position. (If the syllable includes complex coda, final phoneme in the coda is devoiced).
   b. (a) is enforced by devoicing.
   c. (a) is not enforced by coda deletion.
   d. (a) is not enforced by epenthesis.
The descriptive generalisations indicate that the CK words obey a markedness constraint that are violated by _____voiced obstruent≠ or ____ voiced obstruent, where (≠) indicates word boundary and (.) is used for syllable boundary.

4. Constraint Interaction

The decision of which constraints to be used and how they should be ranked in an analysis depends mostly on the input elicited from the data. Sometimes the given data is not sufficient for constructing a solid OT analysis that works beyond the given data. A major reason for that is probably the data set was constructed with a focus on forms that alternate, it may be biased toward input that map to unfaithful output forms. The list of words and phrases of (1b) is good evidence that the input should be of a voiced coda.

Based on the descriptive generalisations in the previous section, a marked constraint is needed for (3a) to disallow no voiced coda which is *VOICED-CODA.

(4) *VOICED-CODA

Obstruents must not be voiced in coda position.

This constraint is well established in OT literature (see Kager 1999, Steriade 2001, Lombardi 2001, McCarthy 2008 inter alia). The basic tenets of OT requires another constraint for the first constraint to conflict with about output forms. The second constraint should be a typical faithfulness constraint requiring the input value of the feature voice to be preserved in the output which is IDENT_IO(voice).

(5) IDENT_IO(voice)

The specification of the feature [voice] of an input segment must be preserved in its output correspondent.

The constraints including the above constraints are supposed to be universal but it is the rankings that are subject to language particulars. For example, in English, the faithful constraint IDENT_IO(voice) outranks *VOICED-CODA resulting in a word in which final voicing is pronounced.

IDENT_IO(voice) » *VOICED-CODA

While in CK, the constraints are ranked in a reversed order; resulting in voice neutralisation in coda.

*VOICED-CODA » IDENT_IO(voice)

To draw a tableau for this ranking, two candidates are needed and since ranking arguments are based on comparing candidates, we need a winner and a loser. The winner is already chosen [ʔaːzaːt] ‘free’ which satisfies *VOICED-CODA and derived from the input lʔaːzaːd/. The loser, on the other hand, should do better than the winner on the IDENT_IO(voice) and worse than the winner on *VOICED-CODA. A loser that meets both these criteria is *[ʔaːzaːd].

Following Brasoveanu and Prince’s (2005) comparison between ranking tableau and violation tableau, I use the ranking tableau as in the ranking tableau the winner is already known and we are trying to figure out which ranking will produce that winner. Now after the winner, the loser and the input are known; the ranking argument can be easily drawn as illustrated in tableau (6).
(6) Ranking argument: *VOICED-CODA » IDENT_IO(voice)

<table>
<thead>
<tr>
<th></th>
<th>*VOICED-CODA</th>
<th>IDENT_IO(voice)</th>
</tr>
</thead>
<tbody>
<tr>
<td>a.</td>
<td></td>
<td>*</td>
</tr>
<tr>
<td>b.</td>
<td>*W</td>
<td>L</td>
</tr>
</tbody>
</table>

A careful look at the markedness constraint *VOICED-CODA reminds us that this constraint can be satisfied in different ways. Lombardi (2001) assumes that *VOICED-CODA can be permuted in final devoicing languages with respect to other faithfulness constraints. She states that different permutations of the constraints either reveal different grammars or explore gaps in the typology of the languages as shown in (7-9). The resulting typology has two significant gaps: languages that avoid voiced codas by deletion as shown in the tableau (7), and languages that avoid voiced codas by vowel epenthesis (8).

(7) Max at bottom (Not attested yet)

<table>
<thead>
<tr>
<th></th>
<th>*VOICED-CODA</th>
<th>IDENT_IO(voice)</th>
<th>DEP</th>
<th>MAX</th>
</tr>
</thead>
<tbody>
<tr>
<td>a.</td>
<td></td>
<td>*</td>
<td></td>
<td></td>
</tr>
<tr>
<td>b.</td>
<td>*W</td>
<td>L</td>
<td></td>
<td></td>
</tr>
<tr>
<td>c.</td>
<td>*W</td>
<td>L</td>
<td></td>
<td></td>
</tr>
<tr>
<td>d.</td>
<td>*W</td>
<td>L</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

(8) DEP at bottom (unattested though Kwakwala may be an example of this ranking)

<table>
<thead>
<tr>
<th></th>
<th>*VOICED-CODA</th>
<th>IDENT_IO(voice)</th>
<th>MAX</th>
<th>DEP</th>
</tr>
</thead>
<tbody>
<tr>
<td>a.</td>
<td></td>
<td>*</td>
<td></td>
<td></td>
</tr>
<tr>
<td>b.</td>
<td>*W</td>
<td>L</td>
<td></td>
<td></td>
</tr>
<tr>
<td>c.</td>
<td>*W</td>
<td>L</td>
<td></td>
<td></td>
</tr>
<tr>
<td>d.</td>
<td>*W</td>
<td>L</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

(9) IDENT (voice) at bottom (German)

<table>
<thead>
<tr>
<th></th>
<th>*VOICED-CODA</th>
<th>MAX</th>
<th>DEP</th>
<th>IDENT_IO(voice)</th>
</tr>
</thead>
<tbody>
<tr>
<td>a.</td>
<td></td>
<td>*</td>
<td></td>
<td></td>
</tr>
<tr>
<td>b.</td>
<td>*W</td>
<td></td>
<td>L</td>
<td></td>
</tr>
<tr>
<td>c.</td>
<td>*W</td>
<td></td>
<td>L</td>
<td></td>
</tr>
<tr>
<td>d.</td>
<td>*W</td>
<td></td>
<td>L</td>
<td></td>
</tr>
</tbody>
</table>

The inevitable question arises here is that why certain constraint ranking seems to be unattested or even impossible? Lombardi (2001) approaches this question by revising the theory of faithfulness suggesting that deleting a voiced consonant is basically less faithful than devoicing it. When a voiced consonant is deleted, IDENT_IO(voice) is vacuously satisfied as the optimal candidate in [ba] in (7) shows. This analysis is parallel with the well-established view in OT which states that violation to input is always minimal and always happens with a good reason. Steriade (2001) takes a similar approach but with a different reasoning. She argues
that relative unfaithfulness is determined by perceptual similarity. Speakers choose [bat] over [ba] because [bat] is more similar perceptually to [bad].

The example from CK supports the view that the candidate that incurs the minimum violation to the input will be the optimal candidate. That is, *VOICED-CODA is satisfied through the violation of IDENT_IO(voice) rather than violating MAX or DEP. However, this analysis is far from convincing (for more discussion, see section (4) of this essay).

(10) IDENT (Voice) at bottom as in CK

<table>
<thead>
<tr>
<th>/ʔaː.zaːd/</th>
<th>*VOICED-CODA</th>
<th>MAX</th>
<th>DEP</th>
<th>IDENT_IO(voice)</th>
</tr>
</thead>
<tbody>
<tr>
<td>a. ☞ ʔaː.zaːt</td>
<td>*W</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>b. ʔaː.zaːd</td>
<td>*W</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>c. ʔaː.zaː</td>
<td>*W</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>d. ʔaː.zaː.da</td>
<td>#W</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

One essential factor for valid ranking that can be used as a test for supporting an analysis is to look for constraints that may do the same job as the dominating constraint. A constraint that seems to do the similar job is the one which states that no obstruent should be voiced; these kinds of constraints are called context-free constraint. This constraint is already available in OT literature under the name VOICED OBSTRUENT PROHIBITION VOP (Kager 1999).

However, CK final devoicing is an example of positional neutralisation: obstruents are devoiced in certain contexts of the syllable coda, while their voicing features are retained elsewhere. This means, devoicing is context sensitive in CK. The three constraints involved in the final obstruent devoicing are ranked in the following way: (these constraints are also used by Kager 1999).

(11) Markedness Constraint-sensitive » Faithfulness » Markedness Constraint –free
    *VOICED-CODA » IDENT_IO(voice) » VOP

The rankings in (11) states that a voiceless realisation of obstruents in coda position outranks voice preservation in coda obstruents. While faithfulness on input feature [voice] outranks the complete devoicing of obstruents. The word /bard/ [bart] ‘stone’ can be a concrete example for the ranking in (11) as the final obstruent in [bart] is unfaithful to the input while the onset /b/ retains its voicing feature required by the context free markedness constraint VOP. The tableau (12) shows the interaction between the constraints.

(12) Tableau format for *VOICED-CODA» IDENT_IO(voice) » VOP

<table>
<thead>
<tr>
<th>/bard/</th>
<th>*VOICED-CODA</th>
<th>IDENT_IO(voice)</th>
<th>VOP</th>
</tr>
</thead>
<tbody>
<tr>
<td>a. ☞ bart</td>
<td>*</td>
<td></td>
<td>*</td>
</tr>
<tr>
<td>b. part</td>
<td>**W</td>
<td></td>
<td>L</td>
</tr>
<tr>
<td>c. bard</td>
<td>W</td>
<td>L</td>
<td>**W</td>
</tr>
<tr>
<td>d. pard</td>
<td>W</td>
<td></td>
<td>*</td>
</tr>
</tbody>
</table>

The candidates (12c, d) are eliminated by the highest ranking constraint *VOICED-CODA, as each contains a voiced obstruent in coda position. The other two remaining candidates (12a-b) satisfy *VOICED-CODA and they are passed to the lower ranking constraints for evaluation. Both (12a and b) violate IDENT_IO(voice), but (12a) is chosen on the basis that its violation to IDENT_IO(voice) is minimal compared to (12 b).
A striking property of constraint ranking in (12) is the candidate (12d). This losing candidate [pard] is harmonically bound which means that cannot win no matter how the constraints are ranked. [pard] preserves the input value for [voice] in its coda consonant, but is unfaithful to [voice] in its onset consonant. This incurs violations for each of the three constraints in the tableau: it violates *VOICE-CODA as it has a voiced coda obstruent [d]. It also violates IDENT-IO(voice) as it is unfaithful to the input value of [voice] in one of its consonants (the onset [p]). In addition, it violates VOP because it contains a voiced obstruent [d]. Kager (1999) gives a similar account for devoicing in Dutch.

5. Is Coda Devoicing a problem for OT?

In rule based phonology, the grammar is an input–output relation as in OT. An element (A) changes to an element (B) in the context of (X) and (Y) as shown here:

\[ A \rightarrow B / X \rightarrow Y \]

Each rule makes only one structural change (A \( \rightarrow B \)) to the input. For every rule to apply, a certain context is needed known as structural description (XAY). Both aspects of rule-based phonology (structural change and structural description) have counterparts in OT. The trigger of the negative constraint (*XAY), can simply be regarded as the structural change and there is an anti-change constraint in the same time represented by faithfulness constraints (*A \( \rightarrow B \)).

However, a basic difference between Rule based theory and OT is that in OT different phonological processes may have a unitary function while rule based theory cannot relate different rules as it has no means of expressing output goal. This functional unity which is known as conspiracy was signalled as early as Kisseberth (1970) who observed that in Yawelmani, (syncope, epenthesis, allomorphy) all act to create or preserve well-formed structure. Before OT, phonologists noticed the role of output in grammatical theory. Hence, mixed models were invented which led to complicated interaction of rules and constraints (Kager 1999).

On the other hand, OT avoids such interactional complexity by limiting grammatical interactions to constraints. It can also account for conspiracy in which different means are used to achieve the same goal. In Welsh, for example, Epenthesis, deletion and metathesis all serve to resolve potential sonority sequencing violations (see Hannahs 2011). While in rule-based theory, the structure is blind to the output and therefore cannot find the coherence relation between the processes.

Moreover, it is assumed that different ranking of phonological processes yield different grammars. For example, *VOICE-CODA dominates IDENT-IO(voice) in German, Polish, Russian, CK and other languages, In languages like Arabic and English, the reverse ranking is seen. Nevertheless, Factorial typology reveals gaps in the ranking, that is, languages which are unattested but might be expected to exist. This means, certain markedness constraints can only be satisfied in only one way and final devoicing seems to be a good example in this respect as tableaux (7) and (8) show.

OT suggests that in any phonological analysis two types of language should be attested: the ones which satisfy the constraint by changing the marked element and the ones where position of a marked element is modified. Yet for most such constraints, the languages which modify the marked element are attested but the ones modifying the position are not (Staroverov 2010). McCarthy (2008) uses the term too-many-solutions problem for this phenomenon which has too many solutions and only one of them is applied cross-linguistically. For example, in the case of coda devoicing, the problem of no voiced coda can be solved in several ways as shown in (13) for bard ‘stone’ but only (13a) is attested in all the languages that devoice coda. Namely, constraint ranking, as expected, is not attested.
(13) a. devoicing: /bard/ → [bart]
b. nasalization: /bard/ → [barm]
c. lenition: /bard/ → [barz]
d. deletion: /bard/ → [bar]
e. feature reversal: /bard/ → [part]
f. metathesis: /bard/ → [darb] or [rabd], etc.
g. epenthesis: /bard/ → [bardə]

Thus, it seems, the predictions of the theory do not match up with the attested typology of languages which claims that OT is inherently typological (McCarthy 2002). This problem has been approached in different ways. First, Steriade (2001) proposes a move towards abandoning the original OT assumption that every constraint may trigger multiple repairs. Second, Lombardi (2001) adopts a modified theory of faithfulness to features which allows them to account for the unavailability of repairs in (13b-d) as well as (13 e-f). Finally, Staroverov (2010) uses Serial OT to tackle this problem by replacing relevant OT constraints with constraints that specify position in the output of the previous derivational step. However, this problem, which is the subject of on-going debate in the literature, is beyond the scope of this essay.

6. Conclusions

CK undergoes coda devoicing: the devoicing includes plosives, fricatives and affricates. This devoicing is neither assimilation nor Coda Condition similar to what happens in Japanese. It is similar to coda devoicing in Dutch and German. Constraint-based analysis can be used to explain devoicing. The constraints used for evaluating the candidates produced by Gen are established in the OT literature.

However, coda devoicing poses a problem for the basic principles of OT in the sense that it does not allow different ranking of the faithfulness constraints. Namely, what is expected by factorial typology (all the possible orders for a given set of constraints) to be attested cross-linguistically has not been met. This problem is approached but has not received a solution with convincing evidence yet.

References


Twana Hamid
School of English Literature, Languages and Linguistics
Percy Building, Newcastle University
Newcastle upon Tyne, NE1 7RU
United Kingdom

t.s.hamid@ncl.ac.uk
THE PROSODIC MARKING OF GIVENNESS IN ENGLISH AND ITALIAN: A COMPARATIVE STUDY

KATIE HARRIS
(University of Cambridge)

Abstract

This paper compares the prosodic marking of repeated (given) information in English and Italian, with a view to addressing the role of cross-linguistic variation in second language acquisition. The investigation considers the prosodic marking of givenness in monolingual Neapolitan Italian (NI) and Standard Southern British English (SBE) speech, in addition to that of a native NI subject who speaks English at an advanced level (NI-EL2). New and given tokens were generated by a semi-spontaneous card game which repeated referents within noun phrases. While the SBE speakers were found to signal given information through accent distribution and duration, the NI speakers did not appear to encode givenness prosodically: accent distribution, pitch contour and duration were not found to be predictive of givenness. The NI-EL2 speaker showcased accent distribution and pitch contours typical of NI in English tokens, indicating that prosodic transfer is persistent at advanced levels.

1. Introduction

The prosodic marking of given referents (which have been previously mentioned or are expected from the discourse) is characteristic of ‘plastic languages’ such as English and Dutch (Vallduvi 1991). The most notable prosodic correlate of givenness is the absence of a pitch accent on a word which would otherwise be accented, known as deaccentuation (Ladd 1996). Contrastingly, ‘non-plastic languages’ such as Italian and Spanish, which encode information status syntactically, tend to resist the deaccentuation of given referents (Cruttenden 1993, Avasani and Vayra 2005). Cross-linguistic variability in the prosodic marking of information status is problematic in second language acquisition as speakers have been found to transfer native language (L1) prosodic strategies to second language (L2) productions (Raiser and Hiligsmann 2009). The current study presents two investigations: a comparative study of the prosodic marking of givenness in Standard Southern British English (SBE) and Neapolitan Italian (NI)1, followed by an analysis of prosodic transfer in the English productions of a native NI speaker.

The first section addresses previous research regarding both the prosodic marking of givenness and prosodic transfer in plastic and non-plastic languages. Subsequent sections present the aforementioned investigations. Finally, results are considered in light of implications for further research.

1 The current study investigates Standard Italian spoken with a Neapolitan accent, rather than the Neapolitan ‘dialect’ which is recognised as a distinct language (UNESCO 2000).
2. Research context
2.1. Plastic languages

In plastic languages, accent distribution functions as a ‘prosodic pointer’ to highlight new/contrastive information and render given information less salient (Avasani and Vayra 2005), as showcased in the following example:

(A) Would you like some CHOCOLATE?
(B) No thanks, I don't EAT chocolate.

In (A) ‘chocolate’ appears as a new referent, which is signalled through accentuation. Conversely, the repetition of ‘chocolate’ in (B) determines its status as given, which is marked accordingly through deaccentuation.

While a large body of research attests the link between deaccentuation and givenness in plastic languages (Terken 1984, Hirschberg 1993, Gussenhoven 2007), a small number of studies reveal conflicting results (e.g. Terken and Hirschberg 1994, Sityaev 2000). This discrepancy may be attributed to the lack of a ‘simplistic binary division’ between new and given information (Baumann 2005). For instance, considerable distance between given co-referents may cause the reintroduction of entities no longer considered given in the immediate context (Sityaev 2000). Similarly, changes in grammatical function and word order may cause accentuation related to syntactic newness, irrespective of discourse related givenness (Terken and Hirschberg 1994).

These studies highlight the need for a more articulated view of what defines ‘givenness’ in intonational analyses: newness and givenness are not dichotomous entities, but points on a scale which accounts for discourse proximity and syntactic relations. Accordingly, in Swerts et al. (2002) Dutch speakers systematically deaccented given information when co-referents were characterised by close proximity, shared syntactic function and surface expression.

2.2. Non-plastic languages

In the aforementioned study, Swerts et al. (2002) compared Italian accentuation strategies to those observed in Dutch and found that given information was not deaccented. This is in line with Ladd’s (1996) observation that non-plastic languages have a fixed accent distribution which tends not to convey information status. Consider the following quote from ex-Italian president Scalfaro.

Italian: [le inchieste] servono a mettere a POSTO cose andate fuori POSTO
English: [the investigations] serve to PUT in place things gone OUT of place

Ladd (1996)

Note that while the given referent posto ‘place’ is accentuated in Italian, its English counterpart is obligatorily deaccented.

It is important to note that Italian may deaccentuate given information in certain contexts, for instance, when expressed as a full noun phrase (NP) or clause (Avesani et al. 1995, Farnetani and Zmarich 1997). However, the systematic accentuation of referents which are repeated within NPs has been attested in a number of studies, indicating that NP-internal deaccentuation of givenness is generally avoided in Italian (Swerts et al 2002, Avesani and Vayra 2005).
2.3. Alternative cues

The prosodic encoding of information structure is not limited to accent distribution cross-linguistically (c.f. Nolan and Jónsdóttir 2001, Zerbian et al. 2010). In English, shorter vowel length is associated with given information (Fowler and Housum 1987), although this is perhaps predictable as a correlate of stress (Fry 1958). With regard to Italian, while Avesani (1997) found that given referents were marked with a specific pitch accent (L*), Avesani and Vayra (2005) found that neither pitch accent nor duration acted as an acoustic correlate of givenness. With specific reference to NI, contrastive/narrow focus has been associated with an L+H* pitch accent (D’Imperio 1999). However, in line with previous observations, the prosodic marking of information structure exhibits variation between full constituents and NP-internal referents: as such, phrasal constituent marking is not necessarily a predictor of NP-internal encoding. The current study is unaware of any investigation which considers accentuation strategies pertaining to NP-internal givenness in NI.

In sum, previous studies report varied results regarding the systematicity of the marking of givenness in English and whether it is marked at all in Italian. This may be attributed to differences in the definition of givenness, in addition to methodological inconsistency and regional variation. Rigorously controlled methods such as those outlined in Swerts et al. (2002) provide support for the deaccentuation of givenness in plastic languages and the lack of associated deaccentuation in non-plastic languages. Accordingly, given referents in the current investigation correspond to the discourse and syntactic criteria outlined in section 2.1. With regard to the presence of additional acoustic cues, it is possible that NI signals givenness through alternative prosodic strategies, although research on this point is for the moment inconclusive.

2.4. Prosodic transfer

The acquisition of L2 prosodic strategies has received limited attention in the literature thus far. Amongst studies which have investigated prosodic acquisition, there is a general consensus that non-target like productions persist amongst advanced learners (e.g. Grosser 1993, Raiser 2003). A further point of concurrence is the clear influence of L1 in L2 prosodic learning (e.g. Backman 1979, Jenner 1976, Lepetit 1989). The current study is unaware of any investigations which specifically address prosodic transfer in the marking of information structure for native Italians acquiring English. Nonetheless, a study of native French speakers learning Dutch (non-plastic and plastic languages respectively) revealed a high degree of prosodic transfer of L1 French accentuation strategies in L2 Dutch speech (Rasier and Hiligsmann 2009). These results indicate that the acquisition of plastic prosodic strategies is problematic for native speakers of non-plastic languages. Moreover, a number of studies highlight the role of L1 transfer in other L2 prosodic features, such as duration and PA (Lepetit 1989, Ueyama 2000).

However, defining the precise role of L1 transfer has proved somewhat elusive. A survey of prosodic acquisition across different language backgrounds (Trouvain and Gut 1997) reveals parallels such as reduced pitch range and inappropriate PA assignment, indicating that universal development patterns may play a role in L2 prosodic acquisition. Yet, as Mennen (2006) notes, isolating developmental phenomena from L1 transfer is often problematic. Non-target patterns which are not clearly derived from L1 strategies would provide support for such universals, although as yet there is no strong body of evidence to sustain such claims.
3. Motivation of the current study

The following research questions will be investigated:

1. Do SBE and NI speakers signal givenness through accent distribution? In line with Swerts et al. (2002) it is predicted that SBE and NI productions will pattern with plastic and non-plastic languages respectively: namely, that SBE subjects will deaccent given tokens whereas in NI speakers will accent both new and given tokens.

2. The second question investigates alternative acoustic correlates of givenness:
   (a) Are durational differences a predictor of givenness in English and Neapolitan Italian? In line with Fowler and Housum (1987) it is predicted that vowel duration in SBE given tokens will be shorter than that of new tokens. In contrast, it is predicted that NI will exhibit no durational differences, as in Avesani and Vayra (2005).
   (b) If NI tokens are accented irrespective of information status, is givenness marked through a specific pitch accent type? In line with Avesani and Vayra (2005), it is predicted that new and given tokens will exhibit no qualitative difference in pitch accent.

3. How does cross-linguistic variation between NI and SBE influence the prosodic treatment of English givenness for a native Neapolitan who speaks English as a second language? There are three potential outcomes:
   (a) The speaker may implement strategies similar to those observed in native SBE speech, showcasing no prosodic transfer.
   (b) The speaker may implement strategies similar to those observed in native NI speech, showcasing prosodic transfer.
   (c) The speaker may exhibit strategies different to those of L1 and L2, representative of a developmental stage.
   In line with Raiser and Hiligsman (2009) and Lepetit (1989) it is predicted that the NI-EL2 speaker’s English productions will be strongly characterised by transfer of NI prosodic strategies.

4. Experiment 1

This experiment addresses question 1 and 2a by comparing accent distribution and vowel duration for new and given tokens produced by native SBE and native NI speakers. Question 2b will be investigated through the comparison of new and given pitch contours.

4.1. Method

Procedure

5 NI monolinguals and 5 SBE monolinguals were recorded in their homes in Naples (Italy) and Cambridge (UK). Participants were naïve as to the purposes of the experiment. Tokens were elicited in a semi-spontaneous card game in which participants were required to match colours and numbers selected by a native speaker researcher. Participants were given 5 coloured cards with different numbers on them. The researcher initiated each turn by announcing the colour and number on a card. Participants were then required to match the
card in colour or number and announce their choice. The game continued until participants had used all of their cards. Each participant played 10 rounds.

Materials

Tokens selected for analysis were the English number ‘two’ and the Italian colour blu ‘blue’ due to their monosyllabic structure and similar vowel qualities. Cards were manipulated in order to elicit 20 phrase-final tokens for each participant: 10 new and 10 given (table 1). Other colours and numbers served as fillers.

Table 1: Eliciting ‘two’ and blu ‘blue’ in new and given contexts

<table>
<thead>
<tr>
<th>Language</th>
<th>Researcher</th>
<th>Participant</th>
<th>Information status</th>
</tr>
</thead>
<tbody>
<tr>
<td>English</td>
<td>Green three</td>
<td>Green two</td>
<td>New</td>
</tr>
<tr>
<td></td>
<td>Red two</td>
<td>Green two</td>
<td>Given</td>
</tr>
<tr>
<td>Italian</td>
<td><em>Sei verde</em> ‘Six green’</td>
<td><em>Sei blu</em> ‘six blue’</td>
<td>New</td>
</tr>
<tr>
<td></td>
<td><em>Due blu</em> ‘Two blue’</td>
<td><em>Sei blu</em></td>
<td>Given</td>
</tr>
</tbody>
</table>

Measurements

Tokens were recorded in mono at 16-bit resolution and re-digitized on Praat (5.3.35) at a sampling rate of 44100 HZ. Native speaker researchers listened to unlabeled tokens in order to determine accent distribution. English and Italian pitch accents were classified in accordance with the ToBI annotation system (figure 1 row 3). Vowel duration was measured by examining the waveform and spectrogram in parallel (figure 1 row 2). The English onset was classed as the beginning of F2 after the stop release in conjunction with the appearance of periodic energy in the waveform. In Italian tokens, the onset was determined as the post-lateral rapid rise in amplitude with a shift in formants in conjunction with a change in period shape in the waveform. Vowel offset was measured in all tokens as the drop in amplitude after the termination of vocal fold striations on the spectrogram.
4.2. Results

Accent distribution and duration

The percentage of accentuated tokens for new and given conditions is presented in table 2. While NI speakers accented 100% of tokens in both new and given conditions, SBE speakers accented 94% of new tokens, and only 6% of given tokens.

Table 2: Accent distribution of given and new tokens in SBE and NI productions

<table>
<thead>
<tr>
<th>Accent distribution</th>
<th>SBE</th>
<th>New</th>
</tr>
</thead>
<tbody>
<tr>
<td>Given</td>
<td>94%</td>
<td>6%</td>
</tr>
<tr>
<td>NI</td>
<td>100%</td>
<td>100%</td>
</tr>
</tbody>
</table>

Duration values are presented in figure 2. The NI speakers exhibited no significant durational differences between new and given tokens. For the SBE speakers, mean vowel duration of given tokens (mean = 154.911 ms SD = 25.73) was markedly shorter than that of new tokens (mean = 178.961 ms SD = 25.62). An independent t-test revealed these differences to be significant \[t (8) = 1.99, p = 0.041\] (one-tailed) \(r = 0.5\). To verify that intonation phrase-final vowel lengthening did not neutralise durational differences in Italian, vowel duration was compared in phrase initial position, \(sei\) ‘six’: table 1. However, no significant difference was observed between given (mean, 137.25 ms, SD = 8.7) and new tokens (mean = 136.27 ms, S.D. = 6.45) \[t (6) = 0.182, p = 0.26\] (one-tailed). One finding which was not part of the original investigation but is nonetheless noteworthy from the above figures, is the longer vowel duration in English compared to Italian. This difference was significant for new tokens \[t (8) = 3.41, p = 0.009\] (two-tailed) \(r = .7\), but neutralised by the reduced English vowel duration in the given condition. NI speech rate (210 wpm) was slightly faster than that of SBE (193 wpm).
Pitch contours

As English given tokens were not accented, their contour will not be considered further here. English new tokens were characterised by an H* followed by a low boundary tone (L%). Accents with a rising boundary tone were observed in 8/100 tones resembling ‘list intonation’ which may reflect the repetitive nature of the task (c.f. Swerts et al. 2002).

Figure 3: A typical realisation of English H* corresponding to new tokens.

With regard to the Italian tokens, both given and new conditions were characterised by a prominent H* on the first word sei ‘six’ and an accent of reduced prominence on the second word blu ‘blue’ (figure 4). This contour was verified against filler tokens viola ‘purple’ and rosso ‘red’ in order to control against the effect of consonantal perturbation. The peak appeared consistently across tokens, indicating that the contour is unlikely to result from perturbation related to segmental material. The majority of tokens exhibited an L% boundary
tone at the right edge of the IP typical of declarative phrases in NI (Grice et al. 2005). 5/100 tokens were marked with rising boundary tones, in line with the aforementioned list intonation in SBE monolingual productions.

**Figure 4: A typical NI pitch contour for given and new tokens**

![Pitch contour](image)

### 4.3. Discussion

As the great majority of English tokens were accented, these results support hypothesis 1 which states that givenness is marked through deaccentation in SBE. The small number of accented tokens may have been due to the nature of the task, where repetition caused participants to lose focus of the information status. Accordingly, these tokens appeared towards the end of the task. In line with hypothesis 2a, SBE given tokens were marked by vowel duration which was significantly shorter than that of the new tokens.

In support of hypotheses 1 and 2, NI did not exhibit prosodic marking of givenness: both new and given tokens were systematically accented and characterised by similar vowel durations. Moreover, tokens carried the same pitch accent across new and given conditions. The following section aims to classify the observed pitch contours.

**Italian pitch contours**

The Italian F0 contours resemble the pitch accents reported in Swerts et al (2002), who investigated native Italian production of noun + adjective NPs similar to those in the current investigation. Although speakers accented both words in all contexts, the second accent was always less prominent and carried the same pitch accent irrespective of information status. Thus, in both Swerts et al. (2002) and the current study, NP-internal intonation curves exhibit sensitivity to phrasal position rather than information status.

The first accent in the present study corresponds to the pre-nuclear H* typical of NI and other varieties of Italian (Grice et al. 2005). Classification of the second accent is more problematic due to a dearth of literature on Neapolitan intonational phonology. The most comprehensive work on NI to date has been carried out by D’Imperio (e.g. 1997; 1999;
2001). These studies primarily focus on pragmatic encoding such as broad, narrow and contrastive focus in addition to declarative-interrogative status. Assignment of the pitch accent based on the current inventory of NI intonational phonology is problematic considering the observed sensitivity to structural, rather than pragmatic factors. Nevertheless, previous research on NI intonational phonology provides a useful point of reference: the following section presents a brief overview of the literature related to NI pitch accents for the purpose of classifying the phrase-final pitch contour observed in the current study.

**Contrastive focus**

Previous research indicates that contrastive/narrow focus phrasal constituents are encoded with an L*+H accent in NI (figure 5b) (Grice et al. 2005). The current study found that such marking was not applicable to NP-internal referents as the pragmatic contrast was not encoded through pitch accent type. Moreover, the pitch contour observed did not appear to feature L* as a phonological target preceding the rise.²

**L***

Given the shallowness of the rise, it is possible that the observed excursion is not relevant to the phonology of the pitch accent, and that the pitch accent corresponds to the monotonal L* observed in Avesani (1997). However, the rise appears systematically enough to be considered a potential phonological target. Moreover, although L* is part of the Neapolitan Italian inventory, it is typically associated with pre-nuclear material.

**Downstep**

In NI it has been proposed that accents with reduced prominence are subject to downstep (!). The shallow pitch contour of the current study is coherent with this interpretation. Consequently, the pitch accent could correspond to !H, although this accent is generally associated with interrogative phrases (figure 5c) (Grice et al 2005). Alternatively, the contour could correspond to the !H+L* accent typical of broad focus sentences in NI (figure 5a). These accents are less prominent than the preceding pre-nuclear H* and ‘merged’ with the L% boundary tone (D’Impero 1999). Given the similarities between the !H+L* accent and that of the current study, it seems possible that the NPs were analysed as broad focus constituents. More precisely, the lack of NP-internal sensitivity to information structure may have caused the phrases to be accentuated according to the broad focus NI declarative pattern. The pitch accent observed here will be described as !H+L* for the purposes of the current study, although a more detailed analysis is left open for exploration in further research.

---

² A valley was observed in 18/100 cases, which was deemed too sporadic to constitute a phonological target. This effect may be a consequence of the drop in transglottal pressure characteristic of voiced plosives, although a detailed analysis of this effect is beyond the scope of this study.
Figure 5: Intonation curves in Neapolitan Italian (Adapted from D’Imperio 1999)

Declaratives
A. Broad focus

[Mamma ballava da Lalla]$_F$ ‘Mum used to dance at Lalla’s’

B. Narrow focus

Mamma ballava [da Lalla]$_F$

C. Y/N questions

Mamma ballava [da Lalla]$_F$?

Duration

One final observation regards the differences in vowel duration between SBE and NI speakers. Although comparative conditions are not ideal due to segmental variation of the preceding sounds (/bl/ in NI and /t/ in SBE) the difference in vowel length observed here has potential implications for NI-EL2, which will be considered in the following section.

5. Experiment 2

This experiment addresses question 3 which explores prosodic transfer in the marking of English givenness by an NI-EL2 speaker. Accent distribution will be measured in 4 conditions: NI new/given and English L2 new/given. Deaccentuation of given tokens (in line with the native SBE speakers in experiment 1) will be taken to indicate L2 prosodic acquisition whereas accentuation of given tokens (in line with the native NI speakers in experiment 1) will point to prosodic transfer. In addition, pitch contours will be compared to English H* and NI !H+L* (experiment 1) in order to identify whether the speaker showcases English pitch accents which reflect prosodic acquisition, NI contours reflective of L1 transfer, or a pattern not directly attributable to either L1 or L2, indicative of a developmental stage.

5.1. Method

Procedure

The NI speaker was recorded at his home in Naples. He had studied English to degree level and had lived in England for a period of 18 months, 1 year prior to the study. He continued to have regular contact with English and his pronunciation level fell within the range of C1 as specified by the Common European Framework of Reference (www.cambridgeesol.org). Procedure, materials and measurements were identical to those used in experiment 1.
5.2. Results

Accent distribution and duration

The percentage of accentuated tokens is presented in table 3 where native values from experiment 1 are reproduced for ease of comparison. All tokens were accented in each condition, in line with the native NI speakers from experiment 1.

Table 3: Accent distribution of given and new tokens for NI-EL2 in SBE and NI speech

<table>
<thead>
<tr>
<th></th>
<th>New Monolingual</th>
<th>New NI-EL2</th>
<th>Given Monolingual</th>
<th>Given NI-EL2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Distribution</td>
<td>NI 100%</td>
<td>100%</td>
<td>100%</td>
<td>100%</td>
</tr>
<tr>
<td></td>
<td>SBE 94%</td>
<td>100%</td>
<td>6%</td>
<td>100%</td>
</tr>
</tbody>
</table>

Duration values are presented in figure 5. English L2 new (mean = 173.34ms SD = 44.2) and given tokens (mean = 174.57ms SD = 27.74) were longer than NI tokens in both new (mean = 119.99ms SD = 27.6) and given conditions (mean = 112.99ms SD = 28.97). However, given tokens were not shorter than new tokens in English productions. High standard deviations were attributed to the fact that recording conditions at times rendered the speaker overly attentive to his speech, causing durational variability. NI vowel duration was shorter than monolingual NI values (experiment 1) which may be attributed to a faster speech rate (NI-EL2 = 240 wpm vs. native NI 210 wpm). Speech rate in English productions (200 wpm) was similar to native SBE values (193wpm).

Figure 6: Duration of new and given tokens in English (GIVEN-SBE, NEW-SBE) and Neapolitan (GIVEN-NI, NEW-NI) for NI-EL2. (NI and SBE from experiment 1)

Pitch accent

The pitch contour corresponded to the !H+L* observed in NI productions in experiment 1. This contour was observed in all 4 conditions (figure 7). As in experiment 1, a rising boundary tone resembling list intonation was observed in 4/40 tokens.
Figure 7: Pitch contours in NI-EL2 tokens. Top row, from left to right: English new, English given; bottom row, from left to right: NI new, NI given.

5.3. Results and discussion

In line with hypothesis 3, the NI-EL2 speaker did not appear to have acquired SBE prosodic strategies for the marking of givenness, nor did he exhibit any prosodic patterns indicative of a developmental stage: new and given tokens were accented irrespective of information structure (in line with native NI distribution) and given vowels were not marked with a shorter duration (in contrast to the SBE speakers in experiment 1). With regard to the general durational differences between NI and SBE in experiment 1, it is worth noting that the NI-EL2 speaker produced longer vowels in English than Italian in both new and given conditions. It appears that the speaker had acquired the increased vowel length characteristic of English but not the prosodic marking of information status. However, these conclusions must remain tentative due to segmental asymmetry. With regard to pitch accent type, the speaker exhibited prosodic transfer from L1 as the H* followed by !H+L* contour resembled that of the NI monolinguals in new and given contexts.
6. General discussion and implications for further research

This study investigated the NP-internal prosodic marking of givenness in Standard Southern British English (SBE) and Neapolitan Italian (NI). In line with previous observations regarding plastic languages (Ladd 1996; Fowler and Houssum 1987), given referents were deaccented and characterised by reduced vowel duration. Conversely, NI given tokens were not distinguished from new tokens in any of the prosodic conditions investigated: tokens were systematically accentuated in both conditions and duration and pitch accent type did not exhibit variation. It is possible that givenness is marked through other prosodic strategies such as pitch range or alignment, though impressionistic views of the current data do not support this notion.

The investigation also explored the nature of prosodic transfer in the marking of givenness in the English L2 productions of a native Neapolitan Italian speaker. As deviations from L1 prosodic patterns were not observed, these results do not support the notion of developmental stages in L2 prosodic acquisition. However, the current study was limited to one speaker who is not necessarily representative of other NI-EL2 learners. Moreover, as only a small selection of prosodic phenomena was examined, these findings do not preclude the existence of a developmental stage.

The observed prosodic transfer corresponds to previous findings regarding the persistence of prosodic transfer at advanced levels (Rasier 2003). It is worth noting that while the speaker exhibited acquisition of the increased English vowel duration, he did not adhere to the durational contrast associated with givenness in SBE. Thus, while the speaker may have acquired prosodic cues pertaining to phrase final lengthening, he did not appear to have acquired the pragmatically governed durational cues. Once more, sample size and segmental asymmetry render such conclusions tentative.

The prosodic transfer of the current study can be accounted for by either a simplistic transferal of L1 prosodic strategies, or by inherent difficulties in the acquisition of the prosodic marking of information structure in plastic languages. The first analysis predicts two effects: 1) speakers of plastic languages such as German and Dutch will exhibit simplistic positive transfer; 2) the acquisition of accentuation strategies will be equally problematic for native speakers of plastic languages acquiring non-plastic languages and vice versa. However, the current body of research on prosodic transfer does not substantiate these predictions. For example, over-accentuation has been identified even amongst native speakers of plastic languages acquiring other plastic languages (Grosser 1993, Wieden 1993). Moreover, Rasier and Hiligsmann (2009) found that native Dutch speakers acquiring French non-plastic strategies were more successful than native French speakers acquiring Dutch plastic strategies. Consequently, they conclude that ‘pragmatic accentuation rules are more difficult to acquire than their structurally-motivated counterparts.’

Such findings are in line with the aforementioned phrase final lengthening observed in the NI-EL2 productions. Taken together with the observed absence of the durational encoding of information structure, these findings support the notion that the acquisition of pragmatically governed prosodic strategies may be more problematic than those which are structurally motivated. This notion is in line with a recent line of research which links the observed variability in the acquisition of certain L2 features to an interaction between the interfaces (c.f. Lozano and Mendikoetxea 2009, White 2011).

Future research will consider the prosodic acquisition of native English subjects acquiring Italian with that of native Italians acquiring English, with the aim of identifying whether the patterns observed in Rasier and Hiligsmann (ibid) are generalizable to other plastic/non-plastic language combinations. While bi-directional studies at the prosodic level
are still relatively scarce, (but c.f. Lepetit 1989, Ueyama 2000) such studies have the potential to offer valuable insights into more general principles governing prosodic acquisition across languages.

References


Harris  The Prosodic Marking of Givenness in English and Italian: A Comparative Study


Katie Harris  
Department of Applied and Theoretical Linguistics  
University of Cambridge  
9 West Road  
Cambridge  
CB3 9DB

klh58@cam.ac.uk
THE SYNTAX AND SEMANTICS OF ARABIC SPATIAL PS*

SAMEERAH SAEED
(Newcastle University)

Abstract

This paper explores the internal syntax and semantics of Arabic spatial prepositions. It is shown that Arabic prepositional elements can be divided into the two main spatial domains: place and path. Moreover, the categorial status and semantics of the two main classes of Arabic prepositions, true prepositions and semi-prepositions, show differences according to their realization and distribution in the spatial P projection hypothesis. Elements within the semi-prepositions, for example, seem to lexicalise different functional heads within a place P projection. The goal of the paper is to propose an extended spatial P projection model for Arabic spatial Ps based mostly on Svenonius (2010) and Pantcheva’s (2011) proposals, which will help further distinguishing between the true and semi-prepositions on one hand and the two subclasses within the semi-prepositions on the other. Along these lines, certain modifications to Svenonius’ (2010) will also be suggested in terms of the way elements used in a place and path domain should be understood.

1. Introduction

Arabic P(reposition)s can be divided into two main classes: true prepositions and semi-prepositions (Badawi, Carter & Gully 2004, Ryding 2005, 2014, Abu-Chacra 2007, inter alia). This division is constructed on a lexical-syntactic basis. The internal syntax of Arabic Ps in general and spatial Ps in particular, has not been paid attention to or examined in the literature of Arabic, leaving lots of gaps in our understanding of the prepositional system of Arabic. For example, in a PP construction as in (1), what does each of min and xəlf lexicalise in a fine-grained PP structure and to which class each of them belongs.

(1) min xəlf fəʃ-ʃə
from behind DEF-screen
‘from behind the screen’

In this paper, I will try to look at the morphological, syntactic and semantic properties of Arabic spatial Ps, in an attempt to define their positions or distributions within a spatial P projection. For this purpose, I will adopt the hypotheses of P projection made by Svenonius (2010) and Pantcheva (2011). The decomposition model suggested by Svenonius (2010), in particular, helps further characterising and distinguishing the elements that belong to the category P in Arabic. For example, true prepositions can lexicalise one specific functional

---

* I would like to thank William van der Wurff, Anders Holmberg, two anonymous reviewers and the audience at the 9th NCL PG conference for their useful comments and suggestions. All remaining errors are mine. The research underlying this paper is supported by the HCED scholarship office in Iraq.

1 My Arabic data are from Modern Standard Arabic (MSA). The examples are either constructed or cited from the arabiCorpus. The examples are transcribed according to the International Phonetic Alphabet (IPA).

2 I use the following abbreviations in the glosses: ACC = accusative case, DEF = definite article, F = feminine, GEN = genitive case, IMP = imperative, M = masculine, NOM = nominative case, PC = pronominal clitic, PST = past, PL = plural, POSS = possessive,PRS = present, 1 = first person, 2 = second person, 3 = third person, SG = singular, Ø = No corresponding element. Dashes (-) are used to separate both affixes and clitics from the stems and periods (.) to separate multiple categories represented by one morpheme.
head, which is Loc in Svenonius (2010), while elements within semi-prepositions are subdivided between Loc and AxPart based on their nominal features. Furthermore, while I agree with Svenonius’ (2010) P projection, I make a few arguments and proposals which capture the way elements used in a place and path domain should be understood and defined. I argue that elements used in a path domain should be viewed as Path Relators while those used in a place domain as Place Relators. As a result I present a fine-grained decomposition model that can work for Arabic spatial Ps. The main functional heads that will be recognised on a morphological and semantic basis are PathRel, PlaceRel and AxPart.

The paper is organised as follows. Section 2 presents a general overview of Arabic prepositional system and its classification. Section 3 presents a detailed syntactic analysis of Ps used in a place domain. The analysis includes discussing their semantic properties and function, morphological make-up and syntactic decomposition. In section 4, I carry out a similar analysis for Ps used in a path domain. A summary and conclusion is presented in section 5.

2. Arabic prepositional system

For the purpose of setting the scene, in this section I will present a brief overview of the prepositional system of Arabic. The overview will include a general sketch of the syntactic and semantic characteristics of the two classes of Arabic Ps as well as the types of complements involved in a PP construction. As mentioned earlier, the two main classes of Arabic Ps are: (a) true prepositions; this is the mono-functional category which includes items that can function only as prepositions; and (b) semi-prepositions; this is a multi-functional category and includes items that can function as adverbs, nouns and prepositions (Badawi, Carter & Gully 2004, Ryding 2005, 2014, Abu-Chacra 2007). This division is constructed on a lexical-syntactic basis; that is, while the true prepositions display all the unique properties of prepositions, the semi-prepositions do not. The true prepositions can be further subdivided into two categories on an orthographic basis: separable and inseparable. The separable Ps are independent elements, e.g. fi ‘in’, ʃəla ‘on’, ʔila ‘to’. The inseparable prepositions, of which there are only few, are prefixed to their complements, e.g. bi- ‘at/in’, lî- ‘to’. See appendix 1 for representative examples.

Exploring the grammatical structure of the true prepositions and the semi-prepositions reveals certain similarities and differences. Consider the examples below:\(^3\)

\[\begin{align*}
(2) & \text{ a. waðəʃ-tu-hu} & \text{ʃəla} & \text{l-minðədø} \\
& \text{put.PST-1SG-3SG} & \text{on} & \text{DEF-table} \\
& \text{‘I put it on the table.’} \\
& \text{b. waðəʃ-tu-hu} & \text{fawq-ə} & \text{l-minðədø} \\
& \text{put.PST-1SG-3SG} & \text{above-ACC} & \text{DEF-table} \\
& \text{‘I put it above the table.’}
\end{align*}\]

Semantically, ʃəla ‘on’ and fawq ‘above’ express the spatial notion of location, and syntactically, in both cases, the following noun is in the genitive case.\(^4\) However, in Arabic

---

\(^3\) In the translation of the Arabic prepositions, I will give the most representative English equivalent(s).

\(^4\) The genitive case on the DP complement takes different surface realizations depending on the noun type. For example, in classical Arabic, where it is mostly marked, the genitive case is usually marked by -i for singular nouns, such as fawqo ʃmiðəbi-dat-i ‘above the table’. Since this is not the main goal of the paper, I refrain from presenting or discussing further forms of genitive case in Arabic (for a detailed list the reader is referred to Ryding 2014: 149-155).
grammar books, ʕəla ‘on’ is categorised as a preposition, and fawq ‘above’ as a noun or adverb of place (see e.g. Abi Asbar 1968, Abdul Hameed 1980, Al-Shumasan 1987). A basic difference between them involves inflection; while prepositions are not inflected, nouns are. Thus, due to their nominal properties, the semi-prepositions can receive inflectional cases such as accusative and genitive markers according to their syntactic functions and positions in the sentence. For example, fawq in (2b) receives the accusative case marking -ə due to its function as an object in the sentence. The case-marking sensitivity of these prepositional elements, however, is more apparent in classical Arabic.5

Another nominal property displayed by the semi-prepositions is that some of them can function as DP complements. See examples below:

(3) a. xəlf ʃ-ʃəfə
behind DEF-screen
‘behind the screen’

b. minə l-xəlf
from DEF-back
‘from the back’

In (3a), xəlf functions as a preposition, while in (3b) it is a DP complement of the preposition minə ‘from’.6 So xəlf in (3b) has totally shifted its category. In addition, some of the semi-prepositions show further nominal properties beside definiteness, such as diminutiveness, e.g. qəbl ~ qubeil ‘a little before’ and bədəl ~ buəeid ‘a little after’. However, despite their nominal features, the semi-prepositions do not accept modification by adjectives or occur with numerals and quantifiers, a feature shared by the preposition al class.

To conclude the discussion so far, words such as fawq ‘above’, xəlf ‘behind’, qəbl ‘before’ and the like are similar to the true prepositions fi ‘in’, bi ‘at/in’ and ʕəla ‘on’ syntactically and semantically, yet not identical due to their nominal origin. They are followed by nouns which are in the genitive case and denote spatial and temporal meanings mostly. Accordingly, I argue that elements such as fawq ‘above’, xəlf ‘behind’, qəbl ‘before’ are prepositions that have been grammaticalised from nouns. To reflect their nominal behaviour in some cases, I refer to them as semi-prepositions, following Ryding’s (2005: 367) terminology.7

As to the type of DP complements, Arabic prepositions can take a range of different complements including noun phrases and clauses (Badawi, Carter & Gully 2004, Ryding 2005). Nouns are the most common complement type and all above examples are of this type. As to pronoun complements, in Arabic these can only be in the form of a pronominal clitic, as in (4):

(4) səʔədəhəb-u ʔhəi-him
will-go-1SG.NOM to-PC.M3PL
‘I will go to them.’

Another complement type is that of clause. The clauses include those introduced by the subordinate marker ʔən ‘that’ followed by verbs in the subjunctive case. Another clause

---

5 In general case endings in MSA are usually pronounced by newscasters and speakers of al fuṣḥa Arabic.
6 When min is followed by definite noun complements, it ends with the vowel -ə.
7 Other suggested terms are ‘prepositionals’ (Badawi, Carter & Gully 2004) and ‘secondary prepositions’ (Abu-Chacra 2007).
type is that initialised by *ma* ‘what’ which is rather a nominal clause. Examples are given in (5a-b):

(5) a. ṭawṣawṣ-tu-hum b.ḏnaḥ bi-ʔan ʔadrus-ʕ bi-djīd
promise.PST-1SG-NOM-M3PL to-that study.PRS-ACC with-hard
‘I promised them to study hard.’

Finally, in Arabic, prepositions can be deleted in certain cases without affecting the grammaticality of the sentence. For instance, prepositions can be optionally deleted when followed by complements in the form of question words, such as *kām* ‘how much’, and also when they are used with a motion verb which can express the path notion. Examples (6a-b) illustrate these two cases:

(6) a. (bi-) ḫam r-ruṣ
by how much DEF-rice
‘How much is the rice?’

b. daxil-na (li-) l-hadīqī
enter.PST-1PL to DEF-garden
‘We entered the garden.’

In the following sections, I examine the distribution of true and semi-prepositions in a spatial P projection. Thus, I will use only those prepositions with spatial uses and/or meanings, using examples of the types made up of [Ps + DP complements] only.

3. Arabic Ps in a place domain

Among the prepositions that are used in static locative relations are *fi* ‘in’, *bi-* ‘at/in’, ḥāla ‘on’, *qurb* ‘near/beside’, *bein* ‘between/among’, ṭawṣawṣ ‘in front of’, *muqabil* ‘opposite’, *ṣalīf* ‘behind’, *fawq* ‘above’, *tūlt* ‘below’, ṭabl ‘before’, *bāṣīd* ‘after’, *jāmin* ‘right’, ḫasar ‘left’, *wasāṭ* ‘middle’, *daxil* ‘inside’, *xarīd* ‘outside’, ṭašla ‘up’, ṭaṣḥal ‘down’. Morphologically, these Ps do not seem to have a complex structure. They are all monomorphemic words, most of which are free independent morphemes while a few are bound morphemes prefixed to their DP complements such as *li-* ‘to’. Some of these elements belong to the true-prepositions and some to the semi-prepositions. The question that arises here is how the true and semi-prepositions are distributed in a place P projection. Before answering this question, a brief overview on place P projection is due.

In the literature on the internal syntax of spatial adpositions, several proposals have been made and attested across languages (see Riemsdijk 1990, Kracht 2002, Svenonius 2008, 2010, Koopman 2010, Dikken 2010, Terzi 2010 among others).\(^8\) The analyses are based on the cartographic approach to phrase structure pioneered by Cinque (1999) and further developed in Cinque (2002), Rizzi (2004) and Cinque and Rizzi (2008). Within this framework, it is argued that phrases and clauses have a complex rich internal structure which can be broken down into several functional elements. Among these, I adopt the syntactic

---

\(^8\) I use the term ‘adposition’ when I abstract over pre- and postpositional elements, otherwise I use the term ‘preposition’, especially when discussing Arabic data.
model of place P projection developed by Svenonius (2010), because it has been applied across several languages with promising results, e.g. Persian (Pantcheva 2006), Hungarian (Hegedűs 2006) and Serbian (Bašić 2007). Besides, as will later be shown, Svenonius’ model works well for Arabic data, although a few modifications and proposals are necessary.

Svenonius (2010) decomposes an English P such as in front of into three functional heads, each of which has a definite semantic function. These are Loc, AxPart and K. The semantic function of Loc is to map regions onto vector spaces.9 For example, in the interpretation of above the window, he proposes a bunch of vectors that project from the window and point upward. AxPart is a function from the set of points occupied by the Ground object in space to some other regions or axes of the Ground such as its top, bottom, front, sides, edges, proximity, etc. (Svenonius 2006, 2010). Thus it hosts nominal elements such as front in in front of, top in on top of, etc. As to the functional head K, semantically it is a ‘function from a Ground DP to a region’ (Svenonius 2010: 132). That is, Svenonius (2010) assumes that K is the element that returns the set of points occupied by the Ground and he refers to this set of points as eigenplace, following Wunderlich (1991). An illustrative example is given in the structure in (7) for the English PP in front of the museum:

(7)  
```
     LocP  
    /    
  Loc   AxPartP  
     |       |  
    AxPart KP  
       |   |  
      K     DP  
     |     |  
    of the museum  
```

While I agree with Svenonius (2010) in terms of the functional sequence these heads maintain across languages, I assume slightly different semantic functions for them, Loc and K in particular. In Saeed (in preparation), I suggest that elements such as in/on/at relate Figures to a specific space with reference to a Ground, and refer to them as Place Relators, hence the functional head PlaceRel.10 For instance, in relates a Figure to an inner space of the Ground, while on relates it to a surface space, and so on. That specific space to which a Figure is related represents the AxPart, which forms a part-whole relationship with the Ground. Finally, following Romeu (2014), I assume that K has a possessive function. It just defines the possessive construction or the part-whole relationship that holds between AxPart elements and the Ground. In English, K can be lexicalised by of, otherwise it is null mostly. In Arabic, K is null and will always be null as Arabic does not spell it out.11 Accordingly, the maximal structure I propose for Arabic PPs used in a place domain is as in (8):

---

9 Svenonius (2010) follows the theory of vectors proposed by Zwarts (1997) and Zwarts and Winter (2000). The latter propose a vector space theory for place adpositional phrases and their modifiers. Vectors are ‘one-dimensional objects with direction and length which define points in a space when they are drawn from a region’ (Svenonius 2006: 52).

10 The terms Figure and Ground are adopted from Talmy (1975). The Figure is the entity whose location is determined, and the Ground is the entity or the location with reference to which a Figure’s location is defined.

11 It is worth mentioning that, as far as Arabic data reveal, when PlaceRel elements are present and AxParts are absent (phonologically), it is PlaceRel which assigns genitive case to the DP complement. In contrast, when AxParts are lexicalised, the genitive case of the DP complement is assigned in the construct state configuration that holds between AxParts and DP complements. This observation leads satisfactorily to the non-necessity of having a K head in Arabic place P projection.
Next I examine the lexicalisation of the heads PlaceRel and AxPart among the Arabic Ps used in a place domain. At first glance, the place expressions within the true prepositions could be said to lexicalise PlaceRel, while those within the semi-prepositions class may lexicalise the AxPart head due to their nominal properties presented above. However, for elements to be assigned to the PlaceRel or AxPart node, certain characteristics should be met. Below I will discuss the properties of the Arabic place-domain Ps in more detail, in terms of (1) their main meanings, (2) co-occurrence with each other, (3) allowing null DP complements, and (4) compatibility with modification expressions.

To start with, the true prepositions fi ‘in’, bi- ‘at/in’ and ħəla ‘on’ relate a Figure to a specific space with reference to a Ground. It is the inner space in case of fi ‘in’ and bi- ‘at/in’,

Illustrative examples are:

(9) a. kan-u fi l-məlsəb
be.PST-3PL in DEF-stadium
‘They were in the stadium.’

b. dˁəʕ-ləʕ-mən
put.IMP.2SG-3SG on DEF-table
‘Put it on the table.’

In addition to these simple uses, these true prepositions can also precede a number of semi-prepositions, such as ħəsfəl ‘down’, wəsəf ‘middle’ and daxil ‘inside’. For example:

(10) a. fi wəsəf l-beit
in middle DEF-house
‘in the middle of the house’ (arabiCorpus, Watan02)

b. bi ħəsfəl l-qəʔim l-ʔeimən
at bottom DEF-port DEF-right
‘at the bottom of the right port’ (arabiCorpus, Ghad01)

c. ħəla ħəsfəl səʔ-śınduq
on bottom DEF-box
‘on the bottom of the box’ (arabiCorpus, Thawra)

Therefore, based on their semantic properties and the word order they have when appearing with the semi-prepositions, these true prepositions seem to lexicalise the syntactic head of PlaceRel. Thus, a PP construction such as fi wəsəf lbeit ‘in the middle of the house’ has the following tree structure:

---

12 These Ps show some allomorphy. The preposition ħəla ‘on’ has the basic forms ħəla in MSA when followed by complements of different types. However, when followed by pronominal clitics, it is ħəlei- in MSA.
As can be seen, I have inserted wasat ‘middle’, which is a semi-preposition, under the AxPart node. This is due to its semantic property of identifying a specific relative frame of reference, the middle space or part of a Ground. Next, I will examine in detail the syntactic and semantic properties of the semi-prepositions.

The semi-prepositions seem to be of two types in terms of displaying nominal features, frames of reference and co-occurrence with true prepositions. For example, wasat ‘middle’, ḥisla ‘up’, ḥasf ‘down’, ḥomam ‘front’, ḥal ‘back’, ḥaxil ‘inside’, ḥarij ‘outside’, ḥomin ‘right’, ḥasar ‘left’ and ḥanib ‘side’ seem to display properties that justify treating them as AxParts, whereas ḥawq ‘above’, ḥeq ‘below’, ḥur ‘near/beside’, bein ‘between/among’, ḥabd ‘before’ and ḥṣd ‘after’ are not likely to be AxParts. For easy reference, I will refer to the former elements as Group A and the latter examples as Group B.

For elements to be categorised as AxParts, they should display specific syntactic and semantic patterns. Syntactically, they should display specific nominal properties and be licit below PlaceRel in a prepositional hierarchy. Semantically, they should define a relative frame of reference – a specific space or part of a Ground. Examples of Group A meet these two conditions. They can be used as nouns and AxParts. As nouns, as in (12a), they can function as a DP Ground, while as AxParts, as in (12b), they define a specific region projecting from a DP Ground, e.g. ḥasf ‘down’ denotes the bottom space, ḥisla ‘up’ the top, ḥal ‘back’ the back and so on.

(11) PlaceRelP
PlaceRel AxPartP
fi ‘in’ AxPartP
wasat ‘middle’ DP
lbeit ‘the house’

In addition, Group A can co-occur with the true prepositions fi ‘in’, bi- ‘at/in’ and ḥisla ‘on’, as shown in examples (10a-c). In fact, a search in the arabicOrpus reveals further examples made up of ḥisla ‘on’, fi ‘in’ and a semi-preposition (some are given in (13)).

(12) a. wasat-tu-hu fi l-ḥasf
put.PST-1SG-3SG at DEF-bottom
‘I put it at the bottom.’

b. wasat-tu-hu fi ḥasf l-bab
put.PST-1SG-3SG at bottom DEF-door
‘I put it at the bottom of the door.’

(13) a. ḥādu-ha n-nahil ẓāqābi ḥisla ḥasf l-ḥwāḏh
hand-POSS.F3SG DEF-thin hold.PRS on down DEF-face
‘Her thin hand was holding the bottom of her face.’ (arabiCorpus, Hayat96)

b. tāmtād ḥisura-ha ḥisla ḥisla qimām l-ḏījāl
stretch.PRS.3SG fences-POSS.3SG on top peaks DEF-mountains
‘Its fences stretch out on top of the mountains.’ (arabiCorpus, Hayat96)
The semi-prepositions in Group B are *fawq ‘above’, *taht ‘below’, *qurb ‘near/beside’, *bein ‘between/among’, *qəbl ‘before’ and *bəd ‘after’. They share a single feature with nouns, which is case inflection. As mentioned earlier, the semi-prepositions can receive inflectional cases such as accusative and genitive according to their syntactic positions in the sentence (recall the example in (2b)). These case markers, however, are mostly apparent in classical Arabic, not the colloquial varieties of Arabic (see the footnote in 5). Thus, these semi-prepositions seem to be dropping their nominal properties and shifting class historically. Moreover, none of them suggest a space or subpart of a Ground or co-occur with any of the true place prepositions. As a result, Ps within Group B cannot be said to lexicalise the AxPart head.

Instead, I assume that these semi-prepositions are more like the true prepositions *fi ‘in’, *bi- ‘at/in’ and *ṣəla ‘on’, syntactically and semantically. They relate Figures to specific spaces with reference to a Ground, such as a relative vertical position in case of *fawq ‘above’ and *taht ‘below’, closeness such as *qurb ‘near’ and *ṣind ‘at’, and so on. Table 1 shows the categorisation of the true and semi-prepositions among the PlaceRel and AxPart projections.

Table 1: Distribution of place-denoting Ps in MSA

<table>
<thead>
<tr>
<th>PlaceRel</th>
<th>AxPart</th>
</tr>
</thead>
<tbody>
<tr>
<td>*fi ‘in’</td>
<td>*ʔəmam ‘in front of’</td>
</tr>
<tr>
<td>*bi- ‘at/in’</td>
<td>*xəlf ‘behind’</td>
</tr>
<tr>
<td>*ṣəla ‘on’</td>
<td>*wəsəf ‘middle’</td>
</tr>
<tr>
<td>*fawq ‘above’</td>
<td>*daxil ‘inside’</td>
</tr>
<tr>
<td>*taht ‘below’</td>
<td>*xarıdʒ ‘outside’</td>
</tr>
<tr>
<td>*qurb ‘near/beside’</td>
<td>*ʔəʃəla ‘up’</td>
</tr>
<tr>
<td>*bein ‘between/among’</td>
<td>*ʔəʃəf ‘down’</td>
</tr>
<tr>
<td>*qəbl ‘before’</td>
<td>*jəmin ‘right’</td>
</tr>
<tr>
<td>*bəd ‘after’</td>
<td>*jəsar ‘left’</td>
</tr>
<tr>
<td>*ṣind ‘at’</td>
<td></td>
</tr>
</tbody>
</table>

However, the semi-prepositions in the PlaceRel column differ from the true prepositions listed there in several respects. First, unlike true prepositions, they do not combine with elements that are AxParts, hence the ungrammaticality of *fawq ʔəmam ‘above
front’ and *\textit{qurb xal}\textit{f} ‘near behind’. This is probably due to their original nominal features, i.e. the ungrammaticality of $[N + AxPart + DP]$. Second, none of the true place prepositions can occur without a phonologically realised DP complement. The complement can be a full DP (14a-b) or a pronominal clitic (14c), yet not a null element.

(14) a. kitab-uk *ʕala *(r-raf)
    book-POSS.2SG on DEF-shelf
    ‘Your book is on the shelf.’

    b. kitab-uk bi-*ʕala (s-ʕinduq)
    book-POSS.2SG in-DEF-box
    ‘Your book is in the box.’

c. wāqaf-na ʕalei-him
    stand.PST-1PL on-PC.3PL
    ‘We stood on them.’

On the other hand, some PlaceRel semi-prepositions, such as \textit{fawq} ‘above’ and \textit{təht} ‘below’ can occur without a DP complement. In such case, they are mostly treated as locative adverbs (see e.g. Badawi, Carter & Gully 2004, Ryding 2005, 2014). Illustrative examples are:

(15) a. kitab-uk fawq
    book-POSS.2SG above
    ‘Your book is above.’

b. sar-u ʕənqən
    walk.PST-3PL east
    ‘They walked to the east.’

With regard to modification, the true and semi-prepositions used in a place domain can be both preceded by modification expressions. However, the syntactic position of the degree modification seems to differ in each case. Consider the examples below.

(16) a. sə-yazid ʕərtifaʕ l-məbna ʔəmanijətə ʔəmətər ʕala
    will-increase height DEF-building eight metres on
    ʔəʕələ məbna fi l-ʕalam
    higher building in DEF-world
    ‘The building will be eight metres higher than the highest building in the world.’
    (arabiCorpus, Hayat97)

b. ʕala ʕənqən səbʕət ʔəmətər fi l-bəhr
    on depth seven metres in DEF-sea
    ‘At a depth of seven metres in the sea.’ (arabiCorpus, Hayat96)

In (16a), \textit{ʔəmanijətə ʔəmətər} ‘eight metres’ is specifying the height difference between the new building and the currently highest building. In (16b), səbʕət ʔəmətər ‘seven metres’ specifies the depth of something that will be made in the sea. Although these examples may superficially look like having modified Ps, they are actually not. Rather the modifiers are in affiliation with the preceding constituents. The PPs are just identifying a locative space. Contrary to these, consider:
(17) a. 407 ʔəmtar  fawq-ə  mistəwa  saťh  l-bəhr
    407 metres  above-ACC  level  surface  DEF-sea
‘407 metres above the sea level’ (arabiCorpus, Hayat96)

b. tamtəd  li-ʔəmtar  təht-ə  l-ʔərdˁ
    stretch.PRS.3SG  for-metres  under-ACC  DEF-earth
‘It stretches for metres under the earth.’ (arabiCorpus, Masri2010)

In (17a-b), the preceding measure phrases seem to define the length of the upward and downward vectors suggested by *fawq* and *təht*, respectively. It can, therefore, be said that the projection Deg (for degree modification), following Svenonius (2010), can be present phonologically (or morphologically) in an Arabic PP made up of PlaceRel and DP Ground, provided the PlaceRel is lexicalised by a semi-preposition.13 (However, there are exceptions; Ps such as *qurb* ‘near’ and *ʕənd* ‘at’ do not allow modification). The position of the Deg is above PlaceRel in a PP structure. This can be illustrated in the following structure for the PP in (17a):

(18)

```
                  DegP
                        Deg
                        407 ʔəmtar
‘407 metres’
                  PlaceRelP
                        PlaceRel
                        AxPartP
                            fawqə ‘above’
                        AxPartP
                            Ø
                        DP
‘mistəwa saťh l-bəhr
the sea level’
```

The differences between the Place Relator true prepositions *fi* ‘in’, *bi*- ‘at/in’ and *ʕala* ‘on’, on the one hand, and the Place Relator semi-prepositions, on the other hand, are summarised in table 2.

---

13 Degree modification is also allowed in some [AxPart + DP Ground] constructions (with the PlaceRel being not lexicalised). For example:

(i) ʾamsat ʔəmtar ʔəmam l-bab
    five metres  front  DEF-door
‘Five metres in front of the door’ (arabiCorpus, Hayat96)
Table 2: True and semi-prepositions: Place Relators

<table>
<thead>
<tr>
<th>Property</th>
<th>True Ps</th>
<th>Semi-Ps</th>
</tr>
</thead>
<tbody>
<tr>
<td>Co-occurrence with AxParts</td>
<td>+</td>
<td>-</td>
</tr>
<tr>
<td>Null DP Complement</td>
<td>-</td>
<td>- (+ in case of fawq ‘above’ and toh ‘below’)</td>
</tr>
<tr>
<td>Degree modification</td>
<td>-</td>
<td>+ (- in case of qurb ‘near’ and Sind ‘at’)</td>
</tr>
</tbody>
</table>

In sum, Arabic seems to have a morphological representation for the functional projections PlaceRel and AxPart. The true prepositions and some of the semi-prepositions lexicalise PlaceRel (represented as Group B), while semi-prepositions in Group A are lexical representations of AxPart. With regard to Deg, it is present in case of semi-prepositions only.

4. Arabic Ps in a path domain

In this section, I analyse the prepositional elements used in the path domain in Arabic. The analysis will capture their typology, semantic properties and internal syntax. However, I first examine the components involved in a path domain and the role of prepositions in such a domain, both syntactically and semantically.

4.1. Path domain: Background

A path is made up of a set of contiguous points (Herweg & Wunderlich 1991, Nam 1995, Krifka 1998). The components or elements involved in a path are a direction, a starting point, an end point, some middle points and a moving object (cf. Zwarts 2005 and Piñón 1993). Thus, a schematic representation of a path would look like the one in figure 1. A represents the starting point, B represents the middle points, C stands for the end point, X is the object that undergoes movement and the arrow signals the direction followed by the object in the specified path (in this case it is from left to right).

Figure 1: The schematic representation of path and its components

\[
\begin{array}{c}
X \\
A \rightarrow B \quad C
\end{array}
\]

Based on these observations, one conclusion is that direction is a component of path and not the reverse. This, however, does not imply that every path necessarily involves a direction, or even a specified starting point or an end point. It rather depends on the type of adposition in use, e.g. to suggests a direction and an end point; from suggests a starting point, and so on.

In the literature on adpositional elements that denote non-locative spatial meanings, such as English to, from, through and across, several labels have been used, the two main terms being path (e.g. Jackendoff 1983, Zwarts 1997, Gehrke 2008, Svenonius 2010, Pantcheva 2011) and directional (e.g. Riemsdijk & Huybregts 2002, Helmantel 2002,
Koopman 2010, Dikken 2010, Noonan 2010). Moreover, they have been claimed to lexicalise the head of a Path projection, which dominates a Place projection (Dikken 2010, Svenonius 2010, Pantcheva 2011, among others). These elements are usually used in dynamic constructions that include a motion verb and a DP Ground which defines a specific point in a path domain. This can be the end point of a path, (19a), the starting point, (19b), or some intermediate point(s), (19c).

(19) a. She went to the beach.
    b. She came from the beach.
    c. She went through the tunnel.

However, in Saeed (in preparation), I suggest that English elements such as to/from/through are better referred to as Path Relators since their main role or function is to relate a Figure to a specific point with reference to a path. For example, to relates a Figure to a Ground which defines the end point of a path, from relates a Figure to a Ground which defines the starting point of a path and through relates a Figure to a Ground which defines the middle point(s) of a path. This suggestion is based on the fact that a path is entailed, but not lexicalised, neither by non-locative elements such as to/from/through nor by other elements. Accordingly, to/from/through lexicalise a Path Relator projection (hence PathRel). Moreover, the position where such elements (PathRel) are introduced into the syntax of a complex prepositional phrase is above PlaceRelP (cf. Jackendoff 1973, 1983, 1990, Koopman 2010, Dikken 2010, Svenonius 2008, 2010). Thus, the structure of a PP such as to/from/through the jungle can be represented as in (20):¹⁴

(20) PathRelP
    PathRel
      to/from/through
    PlaceRelP
      PlaceRel
        Ø
      AxPartP
        AxPart
          Ø
        DP
          the jungle

It is worth mentioning that the PathRel can be further decomposed into several basic functional heads. This decomposition is based on Pantcheva’s (2011) model of path P projection. Based on morphological evidence drawn from various languages, Pantcheva (2011) decomposes the Path head of a path PP (which corresponds to Saeed’s (in preparation) PathRel) into five functional heads: Goal, Source, Route, Scale and Bound. Each of these forms the head of a correspondent phrase which includes a Spec, head and complement, and serves a syntactic and semantic function. Semantically, each of these functional heads contributes a specific meaning distinct from all others. In what follows, I briefly review

¹⁴ Although it is tempting to assume a silent PATH projection beneath PathRel, I will disregard this idea to avoid complexity and leave it for further research in the future.
Pantcheva’s (2011) analysis for the first three functional heads only (Goal, Source and Route) since they form the three main points involved in a path.  

The decomposition that Pantcheva proposes is based on the assumption that ‘morphological complexity indicates syntactic complexity’ (Pantcheva’s 2011: 63). That is, the morpho-syntactic properties of the PathRelP determine its internal syntactic structure. Pantcheva (2011) assumes that route-denoting paths (in her terms) dominate a source-denoting path, which in return dominates a goal-denoting path. This generalisation is based on cross-linguistic morphological evidence of languages with a rich spatial case system, e.g. Source elements contain Goal elements in the Daghestanian language Chamalal and in Imbabura Quechua (see Pantcheva 2011: 46ff for representative examples), Route elements embed Goal elements as in Slovak, and Route elements also embed Source elements as in the Daghestanian language Akhvakh and Avar. Thus, Pantcheva (2011) assumes the following functional sequence Route>Source>Goal. This can be represented in the structure in (21), adapted from Pantcheva (2011). RouteRel hosts the Route Relators, Source hosts the Source Relators and GoalRel hosts the Goal Relators.

(21) RouteRelP  
     |      |      |  
    RouteRel  SourceRelP  
    |      |      |  
     SourceRel GoalRelP  
     |      |  
      GoalRel PlaceRelP  

Arabic Path Relator Ps in general do not display a complex morphological structure. Thus, I mainly apply the structure in (20) for Arabic relevant elements. The decomposition model given in (21) will be used to analyse cases such as min xilal ‘from through’ and min hawlo ‘from around’ later. In sum, the entities involved in a spatial relationship that includes Path Relators are a Figure and a Ground. The Figure’s location is determined with reference to a Ground. The Ground forms a specific point with reference to a path. It can be a starting point (Source), an end point (Goal) or some intermediate points (Route). The main role or function of Path Relators such as to/from/through is to relate the Figure to one of these points. Syntactically, they lexicalise the projection PathRel. In section 4.2, I examine the typology, semantics and syntax of Path Relators as used in Arabic.

4.2. Path Relators in Arabic

In MSA, there are only a few such Ps that relate Figures to the three canonical points in a path: Goal, Source and Route. These are: ḡila/lij- ‘to’, ḥāta ‘until/up to’, nāhwa ‘towards’, min ‘from’, ṣan ‘away from’, xilal ‘through’, ṣabr ‘across’ and hawlo ‘around’. The distribution of these Ps over the three canonical Path Relator types is given in table 3:

---

15 For a comprehensive account of Pantcheva’s (2011) proposed model, the reader is referred to her work.
16 Although in most of the Arabic references nāhwa ‘towards’ is not categorised as a true or semi-preposition except in Ryding (2005) who lists it among the semi-prepositions, I include it in this paper since it behaves similar to prepositions in terms of allowing a DP complement.
Table 3: Types of Path Relator Ps in MSA

<table>
<thead>
<tr>
<th>Goal Relators</th>
<th>Source Relators</th>
<th>Route Relators</th>
</tr>
</thead>
<tbody>
<tr>
<td>ئلاما ‘to’</td>
<td>min ‘from’</td>
<td>xilal ‘through’</td>
</tr>
<tr>
<td>١ةا ‘towards’</td>
<td>ئمن ‘away from’</td>
<td>ئبرا ‘across’</td>
</tr>
<tr>
<td>١تا ‘until/up to’</td>
<td></td>
<td>١اول ‘around’</td>
</tr>
</tbody>
</table>

While ئلاما ‘to’, ١تا ‘until/up to’, min ‘from’ and ئمن ‘away from’ are true prepositions, ١ةا ‘towards’, xilal ‘through’, ئبرا ‘across’ and ١اول ‘around’ are semi-prepositions (Badawi, Carter & Gully 2004, Ryding 2005). The elements listed within each of the columns, however, do differ in terms of specific properties and need further classification. For this purpose I follow Pantcheva’s (2011) path typology, which as far as I know is the most recent and thorough study of path Ps. Her study is a development of path typologies proposed in Jackendoff (1983), Piñón (1993), Kracht (2002) and Zwarts (2008). For example, Jackendoff (1983: 165) identifies three basic types of path ‘according to the path’s relationship to the reference object or place’: Bounded, Directions and Routes. The first two are subdivided in turn into two types, so the total number of path types in Jackendoff’s (1983) typology of paths is five. However, on the basis of data from approximately 80 genealogically different languages, Pantcheva (2011) identifies eight types of paths divided into three canonical path types (Goal, Source and Route). This division is based on the presence or absence of specific properties: ±TRANSITION, ±ORIENTATION and ±DELIMITATION.

By transition, she means paths may contain a ‘transition from one spatial domain to a complementary spatial domain’ (Pantcheva 2011: 14). Some path adpositions have a transitional property and some do not; moreover, those with the transitional property can include one transition or two. Orientation, on the other hand, refers to presence of direction in the movement denoted by a path adposition. Again, some path adpositions denote a specific direction while some do not. Finally, delimitation is related to the presence of a terminative or starting point in a path. The eight path types are given in (22-24) along with their properties and representative Ps from English (see ibid 31).

(22) Goal
   a. Cofinal (+TRANSITIONAL, +Oriented, -DELIMITED): to the school
   b. Terminative (+TRANSITIONAL, +Oriented, -DELIMITED): up to the school
   c. Approximative (-TRANSITIONAL, +Oriented, -DELIMITED): towards the school

(23) Source
   a. Coinitial (+TRANSITIONAL, +Oriented, -DELIMITED): from the school
   b. Egressive (+TRANSITIONAL, +Oriented, +DELIMITED): starting from the school
   c. Recessive (-TRANSITIONAL, +Oriented, -DELIMITED): away from the school

(24) Route
   a. Transitive (+TRANSITIONAL, -Oriented, -DELIMITED): past the school
   b. Prolative (-TRANSITIONAL, -Oriented, -DELIMITED): along the school

17 The small number of Path Relator Ps in MSA may be due to the fact that Arabic is a verb-framed language. That is, in motion events the path is lexicalised in the verb, a strategy common in French and Spanish also. See Talmy (1985) for the typological distinction between verb-framed and satellite-framed languages.
A general observation is that paths can have different shapes, but not different types. There is no goal or source or route path type as such. Instead Goal, Source and Route can be said to represent the points involved in a path as was discussed in section 4.1. That is, the Goal represents the ending point of a path, the Source represents the starting point of a path and the Route represents the intermediate points. Thus, what Jackendoff (1983) and Pantcheva (2011) refer to as path types should be understood as types of Path Relators; they relate Figures to Grounds which define specific points in a path. Below I examine the list of Ps in table 3 in terms of these properties in an attempt to see how many types of Path Relators exist in Arabic.

The Ps listed in the leftmost column in table 3, which includes ʔila/li- ‘to’ and ʔata ‘until/up to’, belong to the type of Goal Relators. See the MSA examples below (the relevant elements are in bold):

(25) a. ʔasʕal-a-t ʔaxiran ʔila ʔəhatat l-metro
   arrive-PST-3SG finally to station DEF-metro
   ‘At last she arrived at the metro station.’ (arabiCorpus, Chicago)

  b. mənas-u ʔaħad l-mʕəsdə\i\-in minə l-nuzul li-l-
   prevent.PST-3PL one DEF-fans from DEF-descending to-DEF-
   maʃəb stadium
   ‘They prevented one of the fans from entering the stadium.’ (arabiCorpus, Ghad02)

  c. taqʕi ʕəla nahr zaʔər ʔlaði jəʔil ʔəhat
   locate.PRS.3SG on river Zaire which reach.PRS up to
   l-ʕasʕimə kinʃasa
   DEF-capital Kinshasa
   ‘It is found on the river Zaire, which reaches up to the capital city Kinshasa.’
   (arabiCorpus, Hayat97)

  d. ṭumə ʔitaʔəh-a nəhwə s-seijəɾə
   then go-PST.3SG towards DEF-car
   ‘He then went towards the car.’ (arabiCorpus, AhlamFawda)

The basic use of all these Path Relator Ps is to relate a Figure to the end point (the goal) of a path. Thus, they are all goal-oriented elements. Differences among them do exist, though, in terms of Pantcheva’s (2011) other properties: transition and delimitation. For example, ʔila and li- ‘to’, (25a-b), being parallel to English to, are supposed to display the properties +T(transitional), +O(riented) and -D(elimited). That is, e.g. in (25a), the Figure (represented by she) is supposed to undergo a transition from one spatial domain to another. However, ʔila and li- ‘to’ do not suggest the end point represented by the Ground ʔəhatat l-metro ‘the metro station’ to be a termination of a path. Thus, they can be characterised as a Cofinal path type, in Pantcheva’s (2011) typology of path.

ʔata ‘until/up to’, (25c), on the other hand, involves a Figure’s transition to the end point, but contrary to ʔila and li- ‘to’, the end point forms the termination of a path. That is, the Figure’s path ends at the Ground identified. For example, in (25c), the DP Ground l-ʕasʕimə kinʃasa ‘the capital city Kinshasa’ is taken as a boundary to the Figure’s movement in a path. Accordingly, ʔata ‘until/up to’ can be said to display the properties +T, +O and +D, and can thus be considered a Terminative element.

Finally, nəhwə ‘towards’ is non-transitional and non-delimited. The PP nəhwə s-seijəɾə ‘towards the car’ in (25d) neither forms the ending point of the Figure’s path nor delimits its
path. *nḥwā* is thus -T, +O and -D and exemplifies an Approximative path element in Pantcheva (2011).

The Path Relator elements listed in the middle column in table 3, that is *min* ‘from’ and *ṣān* ‘away from’, relate a Figure to the starting point or the source of a path. Thus, they are source-oriented. In terms of transition and delimitation, *min* ‘from’ displays the same properties as its corresponding goal elements *ḥila* and *li-* ‘to’. It suggests a transition of a Figure from the Ground spatial domain to an outer location; besides, the Ground is not set as the initial boundary of the path. Accordingly, *min* ‘from’ is a Coinitial element associated with the properties +T, +O and -D. An illustrative example is:

(26) ṭānr ṭawraḏi ḵawm
not go.PRS.2SG from DEF-house today
‘You will not go out of the house today.’ (arabiCorpus, Madbuli)

As to *ṣān* ‘away from’, it is similar to *min* ‘from’ in terms of being source-oriented and suggesting a non-delimitation. However, it differs with respect to transition. Contrary to *min* ‘from’, *ṣān* does not involve a Figure’s transition from the starting point to an outer location. This entails that *ṣān* is -T, +O and -D, exemplifying thus the Recessive path type in Pantcheva’s (2011) typology of path. 18

I turn now to the elements listed in the third column under Route Relator type. These are *xilal* ‘through’, *ṣābra* ‘across’ and *hawla* ‘around’. Example sentences with these Ps are:

(27) a. ẓәmʃi xilal d-ḍar
walk.PRS.1SG through DEF-house
‘I walk through the house.’ (arabiCorpus, Aghani)

b. ẓīṣt‘aḥab-,nī ẓābǎ ṭudhɑt l-qism
accompany-PST.M3SG-1SG across lobbies DEF-department
‘He accompanied me across the lobbies of the department.’ (arabiCorpus, Chicago)

c. ẓīnḏəməm-tu ẓīla l-dɡalsin hawla r-radjo
join-PST.1SG to DEF-sitting around DEF-radio
‘I joined those sitting around the radio.’ (arabiCorpus, Miramar)

All three relate a Figure to the intermediate points involved in a path; besides, all three lack orientations and delimitations. With respect to the transition property, *xilal* ‘through’ suggests a transition of the Figure from a position outside the Ground to a position inside it and then out of it. Therefore, I assume that it displays transition. The same applies to *ṣābra* ‘across’. The difference between them is in terms of the Ground dimensional type; it is usually bounded in the case of *xilal* and unbounded in the case of *ṣābra*. As to *hawla* ‘around’, the Figure does not undergo a transition as it occupies the whole middle sets of points of the path at some time. Thus, it can be said to be a non-transitional element. Accordingly, while *xilal* and *ṣābra* are Transitive elements, *hawla* is a Prolative element, in Pantcheva (2011). However, a search in the arabicCorpus shows that *xilal*, *ṣābra* and *hawla* are used mostly with atelic verbs, such as *ḥasir* ‘walk’ and *ḥ 参数* ‘run’. A few illustrative examples are given in (28):

---

18 *ṣān* is also used to denote distance, occurring with non-motion verbs, as in *jāḍlisu bəṣi’dan ṣān-hum* ‘He is sitting far away from them’.
Thus, these Ps can be described as unbounded elements and may not involve a transitional property. To avoid drawing premature conclusions, I will disregard the two subtypes of Route elements suggested in Pantcheva (2011) and classify xîlal, ṣəhra and hawla as Route Relator Ps.

To conclude, in MSA, there are lexical representatives of six path types identified in Pantcheva’s (2011) typology of path Ps. These are summarised in (29):

(29) a. ṭilali ‘to’: +T, +O, -D = Cofinal
   b. ḥata ‘until/up to’: +T, +O, -D = Terminative
   c. nəhwa ‘towards’: -T, +O, -D = Approximative
   d. min ‘from’: +T, +O, -D = Cointial
   e. ʕən ‘away from’: -T, +O, -D = Recessive
   f. xîlal ‘through’, ṣəhra ‘across’ and hawla ‘around’: -O, -D = Route

Morphologically, these Ps are simple and thus lexicalise the PathRel functional head. However, in a more fine-grained structure, each of these lexicalises the relevant functional head as suggested in Pantcheva’s (2011) decomposition model, e.g. ṭilali- ‘to’, nəhwa ‘towards’ and ḥata ‘until/up to’ will be hosted by the GoalRel node, min ‘from’ will go under SourceRel and xîlal ‘through’, ṣəhra ‘across’ and hawla ‘around’ will be under RouteRel.

Two interesting cases are the co-occurrence of min ‘from’ with xîlal ‘through’ and hawla ‘around’, forming complex constructions such as min xîlal ‘from through’ and min hawla ‘from around’. See examples below:

(30) a. l-mijah sə-təsir min xîlal s-səd
    DEF-water will-flow from through DEF-dam
    ‘The water will flow through the dam.’ (arabiCorpus, Masri2010)
   b. ṭîltəst-ət ṭîlə ṣəbnaʔi-ha min hawli-ha
    turn.PST-3SG to kids-POSS.f3SG from around-PC.f3SG
    ‘She turned to her kids around here.’ (arabiCorpus, Hayat96)

As can be seen, the order displayed by these Path Relator Ps is the reverse of what I assumed earlier (see the structure in (21)). That is, here the Source element embeds the Route element, where it should be the reverse according to Pantcheva (2011). Thus, I assume that there is a kind of movement yielding the order in (30a-b). That is, min has possibly moved to Spec of RouteRelP, giving the order SourceRel>RouteRel.

Finally, in terms of co-occurrence of Place Relators and Path Relators in Arabic, these elements display specific behaviour. Except for ṭilali- ‘to’ and min ‘from’, the rest of the Path Relators do not allow PlaceRel or AxParts. ṭilali- ‘to’ allows lexicalised AxParts only,
while *min* ‘from’ allows lexicalised PlaceRel and AxParts (one exception being *min fi* ‘from in’). See examples below:

(31) a. ḥila daxil māḥatːat l-metro to inside station DEF-metro ‘to inside the metro station’

   b. ḥila ḏawṣat l-madinā to middle DEF-city ‘to the middle of the city’

(32) a. min ṣala t-tāl from on DEF-hill ‘from the top of the hill’

   b. min fawq t-tāl from above DEF-hill ‘from the top of the hill’

   c. min xaridʒ l-madinā from outside DEF-city ‘from outside the city’

In (31a-b), ḥila lexicalises the PathRel functional head. The elements *daxil* ‘inside’ and ḏawṣat ‘middle’ are semi-prepositions and they lexicalise the AxPart terminal node. In (32a-b), *min* ‘from’ is the PathRel element, while ṣala ‘on’ and fawq ‘above’ are under the PlaceRel node. The difference between the two is that ṣala ‘on’ is a true preposition and fawq is a semi-preposition. In (32c), xaridʒ ‘outside’ is the lexicalisation of the AxPart element. For expository purposes, I provide the tree structures of the examples in (31a) and (32a). These are given in (33) and (34), respectively. For simplicity I assume a non-decompositional analysis of the PathRel projection.

(33) PathRelP
    ├── PathRel
    │    └── PlaceRelP
    │           ├── PlaceRel
    │           │    └── AxPartP
    │           │           └── AxPart
    │           │                   │ daxil ‘inside’
    │           │                   └── DP
    │           │                   └── māḥatːat l-metro ‘the metro station’

(34) PathRelP
    ├── PathRel
    │    └── PlaceRelP
    │           ├── AxPartP
    │           │    └── AxPart
    │           │           │ ḥila ‘to’
    │           │           └── DP
    │           │                   └── māḥatːat l-metro ‘the metro station’

61
In this paper, I have examined the prepositions used in Arabic in the two main spatial domains: place and path. The elements used in a place domain have been referred to as Place Relators, and those in the path domain as Path Relators. First, I examined the internal syntax of Place Relators. This included a discussion of their semantics and the types of elements (morphological or lexical) that lexicalise the functional heads included in an extended spatial P projection. Investigation shows that Arabic has a lexical representation for the functional heads PlaceRel and AxPart. The true prepositions always lexicalise the PlaceRel, while elements of the semi-prepositions are distributed among PlaceRel and AxPart.

The second half of the paper was devoted to elements used in a path domain. Examining the morphological structure of Arabic Path Relators has not revealed a rich or complex syntactic structure. That is, most of the path elements are mono-morphemic and encode a single terminal in a path hierarchy, which is the PathRel. The PathRel projection can be, however, broken down into three main functional heads based on the points to which a Path Relator element relates a Figure: GoalRel, SourceRel and RouteRel. Furthermore, in the spirit of Pantcheva’s (2011) typology of path, MSA has representatives of six types of Path Relators. Finally, combinations of PathRelP and PlaceRelP in Arabic are restricted to a few elements only.

Appendix 1: The prepositions in MSA

<table>
<thead>
<tr>
<th></th>
<th>Separable</th>
<th>Inseparable</th>
<th>Semi-prepositions</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>True prepositions</strong></td>
<td>fi ‘in’</td>
<td>bi- ‘at/in/by’</td>
<td>?əmam ‘in front of’</td>
</tr>
<tr>
<td><strong>Semi-prepositions</strong></td>
<td>?əla ‘on’</td>
<td>li- ‘to/for’</td>
<td>xəlf / wəra? ‘behind’</td>
</tr>
<tr>
<td></td>
<td>?ila ‘to’</td>
<td>tə- ‘by’ (for oath)</td>
<td>fawq ‘above’</td>
</tr>
</tbody>
</table>

19 The list of Ps given in the appendix are by no means exhaustive.
<table>
<thead>
<tr>
<th>Arabic</th>
<th>English</th>
<th>Arabic</th>
<th>English</th>
</tr>
</thead>
<tbody>
<tr>
<td>min 'from/of'</td>
<td>wə- 'by' (for oath)</td>
<td>təḥt 'below'</td>
<td></td>
</tr>
<tr>
<td>ʕən 'away from'</td>
<td>kə- 'like'</td>
<td>qəbl 'before'</td>
<td></td>
</tr>
<tr>
<td>ʕəta 'until/up to'</td>
<td></td>
<td>bəsəd 'after'</td>
<td></td>
</tr>
<tr>
<td>ʕind 'at/with'</td>
<td></td>
<td>bein 'between/among'</td>
<td></td>
</tr>
<tr>
<td>məəɡ 'with'</td>
<td></td>
<td>hawlo 'around/about'</td>
<td></td>
</tr>
<tr>
<td>ʔəndu/μdu 'since/so far'</td>
<td>ləda/lədun 'with'</td>
<td></td>
<td></td>
</tr>
<tr>
<td>hafa 'except'</td>
<td>wəsət 'middle'</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ʔəda 'except'</td>
<td>daxil 'inside'</td>
<td></td>
<td></td>
</tr>
<tr>
<td>xəla 'except'</td>
<td>xaridg 'outside'</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ʔəl 'up'</td>
<td>ʔəsfal 'down'</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ʔəmin 'right'</td>
<td>qurb 'near/beside'</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ʔəsar 'left'</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ʔəbrə 'across'</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>xilal 'through'</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>nμqabil 'opposite'</td>
<td>dun/bidun 'without'</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

References


Sameerah Saeed  
School of English Literature, Languages and Linguistics  
Percy Building, Newcastle University  
Newcastle upon Tyne, NE1 7RU  
United Kingdom  

s.saeed@ncl.ac.uk  
saeedors@gmail.com
EXAMINING THE EFFECTS OF STUDY ABROAD ON MANDARIN CHINESE LANGUAGE DEVELOPMENT AMONG UK UNIVERSITY LEARNERS*

CLARE WRIGHT
(University of Reading)

CONG ZHANG
(University of Oxford)

Abstract

This study tracked ten third-year English students learning Mandarin Chinese as a second language (L2) at a UK university, to examine changes in L2 Mandarin during an eight-month period spent studying abroad (SA). We used three writing tasks and four speaking tasks as measures of writing and speaking proficiency, to assess total output, grammatical accuracy, lexical development, pronunciation and fluency, repeated before and after SA in China. Overall mean oral proficiency scores improved significantly ($p < .05$), especially speech rate ($p < .01$), supporting the claim that SA favours fluency development (Freed et al. 2004), although the measures highlighted difficulties in clarifying precisely how to assess oral proficiency. Written proficiency showed fewer marked improvements: only one writing test (an untimed short essay) significantly improved in length ($p < .05$), and increased complex grammar (use of de-relative clause morphemes, $p < .001$). A sub-group (n=7) provided quantitative data on L2 Mandarin use at different times during SA, showing clear individual differences, highlighting the value of capturing details of students’ experiences during SA (Regan et al. 2009). We also note the lack of standardised linguistically-informed measures for tracking development in L2 Mandarin (Freed et al. 2004; Pallotti 2009, De Jong et al. 2012). Further research is therefore much needed to identify systematic linguistic development in L2 Mandarin, and also to bridge theory and practice in L2 Mandarin language teaching to clarify the interconnecting factors that affect L2 Mandarin language development.

1. Introduction

This small-scale exploratory study focuses on learning L2 Mandarin, which is as yet a radically understudied area in language pedagogy and L2 acquisition research. The study compared changes in oral and written measures of L2 Mandarin before and after a period of Study Abroad (SA) study for ten undergraduates from a UK university, who had started learning Mandarin ab initio on arrival at university two years earlier.

Research on L2 acquisition, especially development of L2 speech, has become a richly informed but often fragmented area, involving different disciplines and sub-fields. There is still relatively little research integrating formal research into L2 acquisition of grammatical or lexical knowledge (linguistic competence) with research into changes in L2 oral interaction (communicative competence), despite this problem being noted many years ago by Dell Hymes (1972). In addition, models of longitudinal stages of L2 development,

* We gratefully acknowledge the support of Newcastle University (HASS Faculty Board) for funding this research, also Linda Cheng of the School of Modern Languages for collecting the data, and Dr Alex Ho-Cheong Leung of Northumbria University and Dr Chi-Wai Lee of Newcastle University for assistance with data analysis. We are also most grateful to colleagues from Newcastle University and to this journal’s reviewers, for their support, advice and comments.
especially oral development, remain under-researched, especially beyond the standard English/European language focus (Pienemann 1998; Vainikka & Young-Scholten 2005).

During the last twenty years, research focusing on L2 development of language learners studying abroad (SA), i.e. spending time in the immersion setting, has become increasingly popular. SA research encompasses a broad agenda covering many aspects of linguistic and socio-cultural issues affecting language learners (Kinginger 2011). SA research can thus be seen as resolving some of the fragmentation noted above, offering a way of capturing valuable data of how language knowledge and language use change, particularly in interaction, when the type of exposure to input changes. There is an added pedagogic driver to the research, in terms of evaluating the specific value of SA programmes as study abroad itself attracts more and more language learners, and more language programmes are set up to promote SA (Wright & Schartner 2013). Yet to our knowledge, a bridge between research into these linguistic, contextual and pedagogic factors has not been widely established in L2 Mandarin development.

Given the scope of the current study, we focus here primarily on research on SA effects on language development, taking SA to be a potential trigger for significant change in language proficiency, given the assumed increase in exposure moving from a foreign-language classroom to immersion. However, the results of many current SA studies vary. Some research findings claim that SA is much more beneficial for students’ language proficiency than other contexts such as classroom teaching in the home country (e.g. Brecht et al. 1995; Davidson 2010; Du 2013). Other studies (e.g. Collentine 2004; Freed et al. 2004; Isabelli-Garcia 2010) argue that study abroad does not guarantee language proficiency as commonly assumed, that immersion is necessarily not as deep and effective as expected (Kinginger 2011), but that even the notion of language proficiency itself is complex, and needs careful analysis (see Kinginger 2011 for an overview).

Research on SA effects remains inconclusive at the level of effects on different aspects of language development, such as grammatical accuracy, written proficiency or oral proficiency. Collentine (2004) concluded that study abroad is not necessarily the best way to improve students’ grammatical accuracy. This study examined 17 morphological, syntactic and morpho-syntactic variables to compare the improvements in a group of study-abroad students with a comparable group of at-home students over a semester. The at-home students had more improvements on grammatical items such as present-tense and indicative accuracy on the verbal level and subordinate conjunction selection on the syntactic level. Isabelli-Garcia (2010) focused on one grammatical item, gender agreement, and did not find significant difference between at-home students and study-abroad students after a period of four months.

A few studies have focused on SA effects on written data in general assessments of language development (e.g. Sasaki 2011; Serreno, Tragant & Llanes 2012), although some discrepancy has been found in the length of SA time needed to show significant progress. Sasaki (2011) investigated improvements in Japanese students’ L2 written English with or without SA, focusing on argumentative compositions. This study discovered that the students who studied abroad for more than four months showed significant improvements on a global measure of written ability assessed by EFL specialists than the other students. Serreno et al. (2012) elicited a 150-word written descriptive piece from participants over a three-semester period. After a semester abroad, the students did not show significant progress in fluency, syntactic complexity, lexical richness or accuracy in this writing task from before going abroad. However, by the third semester, progress in those measures became significant.

By comparison, studies on oral proficiency, especially fluency, find more robust support for the benefits of study abroad, even where time spent abroad is not the same (Brecht et al. 1995; Davidson 2010; Du 2013). Brecht et al. (1995) found marked progress in
oral proficiency in US students on a Russian study-abroad project, as measured by oral test scores. Davidson (2010) replicated the study with 1,881 US-based students, and also found improvements in oral proficiency. Specific aspects of oral proficiency, such as target-like sociolinguistic usage, has also been found to improve after SA: studies have found improvements in grammatical/discourse factors in L2 French, such as the use of formal/informal modes of address or informal forms of the negative (e.g. Regan 1995).

However, the claim that study abroad is the best way of improving oral proficiency is not completely supported. Freed et al. (2004) also tested oral proficiency specifically focusing on temporal measures of oral fluency. They examined the performance of three groups of students, at-home students, students who studied in America but in a summer immersion French camp, and study-abroad students who studied in France. The summer immersion group showed improvements on oral fluency after seven weeks, while the at-home group and study-abroad group did not show significant improvements even after 12 weeks.

The variance across the different studies about the effects of SA may therefore, to some extent, be ascribed to methodological and design complexities in the field, as findings may reflect different amounts of time spent abroad (not always controlled for in SA studies), to different tasks used in different studies, time that students spent speaking or writing in the target language, quality of input, or the starting proficiency levels of the students.

Studies which use generalised measures of proficiency such as an essay or the standardised Oral Proficiency Interview (OPI) are difficult to compare with studies which test more specific targets at various levels or different contexts (e.g. comparing Freed et al. 2004 with Brecht et al. 1995). Quantity and quality of input and output in the target language during SA may also differ greatly: Du (2013) claimed that oral fluency is most influenced by time-on-task, i.e. the amount of time that students use the target language every day. However, Moyer (2011) evaluated the effect of both the quality and quantity of L2 experience on accent as another aspect of oral proficiency, and found that quality of L2 experience is more important than quantity, as measured in terms of significant context-specific interaction. Meanwhile, Wright’s (2013) longitudinal study of oral proficiency among 32 Mandarin learners of English in postgraduate study in the UK found no significant effect on improvement associated with qualitative or quantitative differences in target language use by the study participants.

In addition to quantity and quality of input and output, the starting proficiency level is believed to be a relevant factor too. Davidson (2010) found speaking gains among English L1 students only at advanced levels in a Russian study abroad program. Marqués-Pascual (2011) discovered that advanced learners produced subject-verb inversions after a semester abroad in a Spanish verbal morphology study, while the intermediate-level students did not.

There is clearly great variance across research questions, methodologies and results for SA research generally, and more consistency of research tools is needed to establish greater cross-comparison among different measures of language knowledge and language use. To date most of the research has focused on European languages, which are arguably typologically and sociolinguistically somewhat related. It therefore needs to be established how far standard assumptions about SA drawn from SLA and pedagogic research can be transferred to non-European languages such as Chinese, and L2 Mandarin specifically. In this article, we will use the terms Mandarin throughout to avoid confusion, though we note that the sources referred use either Mandarin or Chinese to some extent interchangeably.

It is clear that SLA research in L2 Mandarin is much needed, given recent rapid increases in numbers of students studying Mandarin, including in study abroad settings. According to official Chinese sources, such as the China Scholarship Council and CUCAS, China’s English-language University and College Admission System, the number of international students in China is rapidly increasing. In 2012, over 320,000 students were
registered in China from over 180 countries to study at both degree and non-degree level (CUCAS 2013), with a projected target of 500,000 by 2020 (China Scholarship Council 2013).

Yet within SLA and pedagogic SA literature, this explosion remains relatively underexplored, and somewhat disparately reported, at least in English-language publications. Two recent papers, Shi and Wen (2009, in Mandarin) and Zhao (2011), provide overviews of linguistic research on L2 Mandarin acquisition, but these typically examined levels of success in acquisition of specific linguistic features, and did not include studies relating to longitudinal development or effects of spending time in the SA context. A brief report assessing SA programmes run by Georgetown University’s Office of International Programs (VandeBerg, Connor-Linton & Paige 2009) included L2 Mandarin SA data, gathered using the standardised ACTFL Oral Proficiency Interview (OPI). Overall, participants across the whole program were found to have significantly improved oral proficiency; however, since this report incorporated all the Georgetown study abroad programs across eight countries, it was impossible to identify the exact improvements for L2 Mandarin learners.

Other studies which specifically examined SA in China include research focusing on sociolinguistic and attitudinal change, such as Jin (2012), Liu (2010) and Yu (2010). Yu (2010) collected questionnaires from 90 L2 Mandarin learners over a period of nine months. The result partially supported the claim that language-related attitudes and motivation would increase while language anxiety would decrease over the period, and concluded that students’ self-ratings on their own language proficiency would increase over time. However, there was no objective measure of language proficiency reported in the study. Liu (2010) studied twelve students’ improvements in a US study abroad program, by integrating different measurements: OPI, Mandarin language standardised assessment tests (SAT College Board 2014), a portfolio of general writing tasks and a survey asking for self-ratings on reading, listening, speaking, writing, cultural awareness, and personal career development. After an 8-week at-home immersion preparation program (living with native Mandarin speakers), an 11-month year of academic study before going to China, and then 4 weeks’ residence in China, the students’ tests scores all showed descriptive increases, and the majority reached advanced level on the OPI ratings after the period of residence in China. However, there was no report of what language features improved the most in each specific context, or any quantitative results to show how the oral and writing task performances compared. Jin (2012) investigated the effect of a study abroad programme; again this paper has a mainly pragmatic or sociolinguistic focus, discussing whether the students could successfully learn the use of compliment words, compared to native speakers. Whether the students’ language fluency or accuracy was improved was not specifically reported.

Du (2013) is one of very few studies to take oral L2 Mandarin as its focus, investigating the development of Mandarin fluency over one SA semester during a 3-year study program. The researcher collected a range of speaking data every month from 29 students during their semester in China, in different contexts on and off campus, using both recorded Mandarin speaking classes for planned instructed output and using Labovian-style individual interviews¹ (Silver & Lwin 2013: 78), to elicit a range of spontaneous output. This study showed significant fluency progress over time but the results were taken from specific 2-minute segments chosen to highlight the students’ most “productive” moments (Du 2013: 135) in terms of numbers of morphemes produced, from each of the four sessions across the semester. This may not therefore be as representative of changes in fluency as at first sight.

¹ A Labovian sociolinguistic interview typically covers a range of topics and activities, including the interviewee retelling an emotionally charged personal experience, producing a rich range of language data of varying degrees of formality and complexity.
Nor was there any evidence from other sources such as written output to provide any comparison of overall language development.

It is clear that there is still much to be known about how L2 Mandarin may be expected to change over time, with no established models of stages of development in grammar or vocabulary, or reliable empirical methods to assess changes in written proficiency and oral proficiency. This paper therefore seeks to add to this new and growing field of research by reporting on an exploratory year-long study of SA effects for a group of intermediate-level students of Mandarin at a UK university, examining aspects of language development in both writing and speaking before and after SA. We assessed writing using four measures for accuracy, length, complexity and optionality. We assessed speaking using six measures for accuracy (in grammar and pronunciation), total output (number of utterances, length of utterances, lexical diversity, hesitations and speech rate).

By combining both written and spoken language changes in a SA context, we aim to provide a holistic perspective on changes in both language knowledge and language use, of benefit for future linguistic and pedagogic research in L2 Mandarin.

2. Study design

The study presented here, for reasons of space, focus on a subset of the whole research project, looking at comparisons before and after SA of assessments of written and oral proficiency. Our research questions were:

(1) How does writing performance change after Study Abroad (SA)?

(2) How does oral performance change after SA?

2.1. Participants

There were ten third-year L2 Mandarin students from a UK university (aged around 20 years old), who went to different cities in China for their study abroad. None had studied Mandarin prior to starting the course at the UK university. All the students during the SA were based in university classes, in Beijing, Shanghai, Xi’an, Chengdu, or Hainan for around eight months (two semesters, October to June); some stayed on longer after classes finished for holiday travel. Most were in mixed-language university dorms throughout their residence in China. This resulted in some lingua franca English use outside class but was comparable across the group. They were therefore judged, as far as possible, to have had comparable experiences in formal exposure to Mandarin during SA.

2.2. Data collection

Given the lack of normalised measures for L2 Mandarin development, and the exploratory nature of the study, the researchers used standard university examination data pre-SA and repeated post-SA, using a battery of writing and speaking tests. These tests (detailed below) provided quantitative data on changes in writing and speaking performance for comparison in a statistically reliable pre-test/post-test design. As language learning context is known to affect language development (Kinginger 2011; Moyer 2004), individual reports of language usage in formal settings in class, and informal settings out of class, were

---

2 One moved into private off-campus accommodation after one semester, but her data did not show significant differences in improvement to the others, so this is not taken to indicate that her living situation impacted on her language development significantly differently to the rest of the group.
also collected throughout the year, to provide some contextual quantitative information about the amount and type of exposure the students had. The language usage reports in this study were not designed to tap exposure in qualitative detail, given the exploratory scope of the study, and the small numbers of participants, but they provide some context for the longitudinal data presented here, comparable to other research conducted using similar tools (e.g. Wright 2013). As noted above, formal exposure to Mandarin was deemed to be reasonably comparable; there may well have been individual differences in informal exposure to Mandarin, but the language usage reports did not appear to capture any significant skewing effect, and so the pre and post-SA test design is argued here to provide a valid set of reliable data for statistical comparison.

2.2.1. Written production data

At both times of testing, the participants were asked to produce two pieces of writing on similar themes in exam conditions, within half an hour – a description and a dialogue; following standard assessments, these were marked out of 15; they also completed one untimed free essay as a classroom-based assessment on expectations of life in China. These three pieces of writing yielded four measures for writing development: one overall measure of grammatical accuracy on the timed work, and three linguistically-motivated measures on the untimed piece of work: length (total characters), complexity and discourse-level optionality.

Accuracy was measured as target-like use of functional morphemes. Complexity was measured by use of two de-morphemes: de-possessive, seen as early acquired, and de-relative, marking relative clauses, seen as late acquired (Zhang 2005). This distinction follows standard models of SLA (and indeed child acquisition) which suggest that late acquisition, either for formal or processing reasons (Pienemann 1998; Vainikka & Young-Scholten 2005), relative clause structures are late acquired. Discourse-level optionality was marked by omission of the shì copula, which is required in L1 English, but can be omitted in L2 Mandarin in certain pragmatically-licensed predicate contexts, and is known to be hard to be target-like even at advanced level (Yuan 2013, p.c.).

The different measures for untimed vs. timed work were required due to lack of consistent production in the timed work (some participants did not complete both sections of the task, or had marked differences in length of writing), which made the specific linguistic analyses of length, complexity and optionality hard to compare statistically across all three pieces of written work on both times of assessment.

2.2.2. Oral production data

At both times of testing, the participants were given four tasks to complete in approximately ten minutes, tapping different aspects of preparation, planning and communicative competence. These four tasks were: a pre-prepared monologue, one of a choice of pre-prepared dialogic role plays, an unprepared, unplanned description of a photograph prompting questions about the content, and a free dialogue about expectations of life in China. The tasks were recorded in examination conditions, then transcribed for further linguistic analysis.

The tasks yielded six measures of oral development: an overall measure of average oral proficiency (aggregating standard assessment ratings out of 15 for accuracy and pronunciation); then five linguistically-motivated measures: total number of utterances, mean length of utterance or turn (MLT), lexical diversity (using a standardised measure of lexical range: D), disfluency, and speech rate.
Accuracy and pronunciation were recorded by the native-speaking examiner, using a subjective but standardised rating scheme as part of the overall university examination scoring procedure. Other measures were assessed by the researchers after transcribing the speaking tests using CHILDES conventions (MacWhinney 2000), using CLAN programs to analyse total number of utterances, mean length of utterance, and lexical diversity using D (Malvern et al. 2004). Transcripts were analysed in terms of single characters, in line with some researchers, who acknowledge that standard definitions of morphemes or words do not easily fit the Mandarin context, but that character calculations can be reliably compared to syllable-based measures of western languages (Du 2013). There is a potential problem when calculating D in Mandarin, as the distinction between single and bi-morphemes could be blurred, and making comparisons with lexical diversity in other language studies harder to manage. However, the researchers were unable to find an alternative normed measure, so D as measured by single characters was used here to aim for greatest consistency. Disfluency was calculated following Wright (2013) as the aggregated total of filled pauses (um, er) and repairs (repeated and reformulated expressions). The word count totals are automatically produced in CLAN analyses using standard computerised programs which measure mean length of turn (MLT) and word frequency (FREQ). The aggregated total was then divided by 100 to present a ratio (between 0 and 1) of disfluency per total length of output.

Speech rate was assessed by two native-speaker research assistants, analysing each of the four speaking tasks (each approximately 2 minutes long), which measured the number of characters produced in a 20-second segment in the middle minute of the task. We appreciate this is not the standard way of establishing speech rate (e.g. De Jong et al. 2012), which would be to use a per-minute ratio (using syllable or word). However, each task used in this study was relatively short, with no official timing constraint across the whole test or within tasks, so we found considerable variability in how each speaking test was conducted both in overall time given to each component, and in interviewer/interviewee utterances at the start and end of each task. Given this variability, we decided that using a standardised selection of the central segment of each task was a sufficiently clear length of run, not confounded by task process, which would be valid and reliable as a measure of speech rate for each participant across all four tasks at both times of assessment.

2.2.3. Data collection

We recorded participants’ scores from the end of second-year university examinations, taking their results from the oral and written components; we also used a final in-class free writing assessment (Time 1). The same students took the same oral and written exam components and writing assignment at October at the start of their final year – i.e. Year 4 (Time 2) following their year in China. All scores from Time 1 and Time 2 were coded using SPSS and compared using statistical analysis. Due to student absences from written task data collection at Time 2, the written analyses refer to nine participants for the timed data and eight participants for the untimed data, while the oral analyses include all ten participants. Given the small participant set, all data are analysed using non-parametric tests.

3. Results and discussion
3.1. Research Question 1 – writing

Our first research question looked at changes in writing performance after Study Abroad. The writing performances were marked with one score for accuracy out of 15. Mean results (with SD, minimum and maximum) comparing Time 1 and Time 2 are shown in Table 1 below; percentages are also shown for ease of comparison. These (and all subsequent)
scores were tested for significant differences using non-parametric Wilcoxon signed rank analysis, due to the small sample size (significance assessed as $p < .05$ or below).

**Table 1: Writing Scores**

<table>
<thead>
<tr>
<th></th>
<th>N</th>
<th>Minimum</th>
<th>Maximum</th>
<th>Mean</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Average timed writing score Time1</td>
<td>8</td>
<td>8.75</td>
<td>13.00</td>
<td>10.75</td>
<td>1.269</td>
</tr>
<tr>
<td>(58.33%)</td>
<td></td>
<td>(86.67%)</td>
<td>(71.67%)</td>
<td>(8.457%)</td>
<td></td>
</tr>
<tr>
<td>Average timed writing score Time2</td>
<td>8</td>
<td>8.50</td>
<td>12.25</td>
<td>10.39</td>
<td>1.377</td>
</tr>
<tr>
<td>(56.67%)</td>
<td></td>
<td>(81.67%)</td>
<td>(69.26%)</td>
<td>(9.171%)</td>
<td></td>
</tr>
</tbody>
</table>

The writing scores reveal that even at Time 1, participants were rated at just above 70% accuracy; there was wide individual range – the highest-scoring participant achieved 87%, compared to the lowest-scoring participant with 58%. By Time 2, there was a slight overall decrease to just below 70% mean accuracy, with slightly larger individual range – the highest score decreased but still was above 80%, while the lowest scored 57%. Investigating further, it was evident that Task 1, the dialogue, showed a decrease from a mean of 11.17 at Time 1 to a mean of 10.33 at Time 2, compared to Task 2, the letter, which showed almost no change from a mean of 10.33 at Time 1 to 10.61 at Time 2.

Measures from the third piece of writing, the untimed class-based assessment, allowed us to examine development in terms of total length (number of characters) and specified target morphemes, shown in Table 2 below. Length of writing showing a marked increase by Time 2, and reduced SD (although the change was not significant, $p > .05$)

**Table 2: Total length of writing assignment (total characters)**

<table>
<thead>
<tr>
<th></th>
<th>N</th>
<th>Minimum</th>
<th>Maximum</th>
<th>Mean</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total length Time 1</td>
<td>8</td>
<td>544</td>
<td>825</td>
<td>661.50</td>
<td>105.93</td>
</tr>
<tr>
<td>Total length Time 2</td>
<td>8</td>
<td>615</td>
<td>904</td>
<td>715.50</td>
<td>98.91</td>
</tr>
</tbody>
</table>

We also analysed changes in three specified morphemes used here to indicate grammatical development. The *de*-possessive marker, and the *de*-relative clause marker, were taken to indicate earlier and later stages of development (following Zhang, 2005). The third indicator of grammatical development was to check for accurate use of the *shi* copula which is optional in Mandarin – whether it is needed or not is pragmatically determined by the context, unlike English copula *be*, which is always required. The total number of characters and the total number of accurately produced target morphemes were noted and compared between Time 1 and Time 2, shown in Table 3 below.

**Table 3: Total production of target morphemes in writing**

<table>
<thead>
<tr>
<th></th>
<th>N</th>
<th>Minimum</th>
<th>Maximum</th>
<th>Mean</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>de</em> possessive morpheme Time 1</td>
<td>8</td>
<td>14</td>
<td>28</td>
<td>20.38</td>
<td>5.24</td>
</tr>
<tr>
<td><em>de</em> possessive morpheme Time 2</td>
<td>8</td>
<td>6</td>
<td>19</td>
<td>10.00*</td>
<td>4.47</td>
</tr>
<tr>
<td><em>de</em> relative morpheme Time 1</td>
<td>8</td>
<td>1</td>
<td>4</td>
<td>2.50</td>
<td>1.07</td>
</tr>
<tr>
<td><em>de</em> relative morpheme Time 2</td>
<td>8</td>
<td>13</td>
<td>24</td>
<td>18.00*</td>
<td>3.59</td>
</tr>
<tr>
<td><em>shi</em> copula morpheme Time 1</td>
<td>8</td>
<td>1</td>
<td>7</td>
<td>4.63</td>
<td>2.56</td>
</tr>
<tr>
<td><em>shi</em> copula morpheme Time 2</td>
<td>8</td>
<td>2</td>
<td>8</td>
<td>3.25</td>
<td>1.98</td>
</tr>
</tbody>
</table>

*Significant difference between Time 1 and Time 2, $p < .05$*
There was a significant decrease in *de*-possessive, from 20 tokens to 10 ($p < .05$), and a significant increase in *de*-relative, from 2 to over 20 ($p < .05$). Although the overall number of tokens is relatively few, the wide use of the *de*-possessive at Time 1 suggests this marker is easily acquirable. The few occurrences of *de*-possessive at Time 1, and the significant improvement in *de*-relative by Time 2 provides some evidence for Zhang’s (2005) argument that *de*-relative is acquired after *de*-possessive. It is likely that some increase in *de*-relative would have been seen in any learning situation, and further study is required to test the claim more reliably that SA fosters acquisition of specific elements of complex grammar. The decrease in *de*-possessive by Time 2 is noteworthy, but is not, of course, to do with acquisition per se. Rather, we argue that the change in distribution may reflect a shift away from over-reliance on early-acquired grammatical forms, as a wider range of complex morphemes are acquired such as aspect markers, passive constructions and other more complex grammatical forms. But future research into this data for greater lexical analysis of morpheme distribution is needed to support this hypothesis.

The cohort’s grasp of the appropriate optionality of the *shi*-morpheme was markedly lower than the other target morphemes at Time 1, with no marked improvement by Time 2. This fitted expectations, since this optional morpheme was targeted to show pragmatic or discourse-level understanding of where grammaticality depends on context, and was predicted to be late acquired. However, the very small number of morphemes produced here are not sufficient evidence to validate our hypothesis; more specific tests are needed to see how far optionality remains a problem during the four years’ formal study of Mandarin and whether SA could make a real difference, as it has been shown to do on other studies of contextually determined optionality, such as Regan’s (1995) study of L2 dropping of *ne* in French negation).

The results here are illustrative rather than generalisable, but indicated that, in general, writing, especially timed, remained problematic for participants, which needs further investigation as to why the mode of assessment is challenging: it could be down to L2 specific problems (e.g. lexical knowledge/character familiarity), or generic problems in foreign language writing at schematic/discourse level (as noted in the L2 English writing literature, e.g. Hamp-Lyons 1991, Ferris & Hedgcock 2005). Focusing on discourse-level writing skills as such was beyond the scope of this study. However, we argue that classroom work could benefit from including specific linguistically-motivated measures to test development. We found improvements in the untimed piece of work in terms of marked increase in length, and evidence of grammatical development with significantly higher uses of the more complex *de*-relative morpheme.

This kind of linguistic analysis provides deeper insights into the specific nature of changing language knowledge in line with other SLA research, but could also be helpful for teachers and language departments to understand better how to assess evidence of progress in using increasingly complex morphemes. The improvement in quantity and complexity would not have been identified by simply comparing exam scores in the timed written tasks, which may commonly be the only formal summative evidence put on record. As this is based on only one institution’s practice, we do not draw any wider conclusions about the nature of summative assessment in L2 Mandarin classrooms in general, but we suggest it would be useful to validate best practice across institutional ways of assessing L2 Mandarin, to ensure sufficient breadth of linguistic change is captured.
3.2. Research Question 2 - speaking

Speaking performances were assessed for accuracy and pronunciation, aggregated together to give an average oral score (out of 10). Again, mean results (with SD, min and max) comparing Time 1 and Time 2 are shown in Table 4 below; percentages are also shown for ease of comparison.

Table 4: Oral Scores from examiner ratings (aggregating pronunciation and accuracy)

<table>
<thead>
<tr>
<th></th>
<th>N</th>
<th>Minimum</th>
<th>Maximum</th>
<th>Mean</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Average oral score</td>
<td>10</td>
<td>5.00</td>
<td>8.20</td>
<td>6.54</td>
<td>1.103</td>
</tr>
<tr>
<td>Time 1</td>
<td></td>
<td>(50.0%)</td>
<td>(82.0%)</td>
<td>(65.4%)</td>
<td>(11.04%)</td>
</tr>
<tr>
<td>Average oral score</td>
<td>10</td>
<td>5.60</td>
<td>9.00</td>
<td>7.31</td>
<td>1.172</td>
</tr>
<tr>
<td>Time 2</td>
<td></td>
<td>(56.0%)</td>
<td>(90.0%)</td>
<td>(73.1%)</td>
<td>(11.7%)</td>
</tr>
</tbody>
</table>

The aggregated speaking scores from the examiner show clear though not statistically significant changes over time. Mean scores at Time 1 was around 65%, although the highest score was above 80%. By Time 2 the mean had increased to 73%. This also reflected a wider individual range than at Time 1 - the highest scoring participant achieved 90%, while the lowest scored less than 60%. It is interesting to note that this shifted the comparison between writing and speaking as measured purely in examination scores – participants were rated overall lower in speaking than writing at Time 1, but higher in speaking than writing at Time 2.

These differences seem to support claims that speaking is the language skill most assisted by study abroad (e.g. Freed et al. 2004). However, closer analysis of the two oral sub-measures combined here – accuracy and pronunciation – showed differences between the two sub-measures. Pronunciation improved from a mean of 7.15 to 7.5, while accuracy remained similar with a mean 6.4 at Time 1, and 6.3 at Time 2.

We then transcribed and analysed the oral tests to look further at oral proficiency, using specific linguistic measures of fluency and lexical development. The transcripts were analysed using CLAN (MacWhinney 2000) to measure total number of utterances, mean length of utterance, lexical diversity (D), mean speech rate and disfluency. These measures were calculated using characters for consistency. Speech rate scores for Time 1 and Time 2 are shown in Table 5 below.

Table 5: Oral scores by sub-measure

<table>
<thead>
<tr>
<th></th>
<th>N</th>
<th>Minimum</th>
<th>Maximum</th>
<th>Mean</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total utterances Time 1</td>
<td>10</td>
<td>44</td>
<td>112</td>
<td>67.80</td>
<td>21.75</td>
</tr>
<tr>
<td>Total utterances Time 2</td>
<td>10</td>
<td>71</td>
<td>216</td>
<td>113.20**</td>
<td>39.61</td>
</tr>
<tr>
<td>Mean length of utterance</td>
<td>10</td>
<td>8.01</td>
<td>12.46</td>
<td>9.91</td>
<td>1.45</td>
</tr>
<tr>
<td>Time 1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean length of utterance</td>
<td>10</td>
<td>6.61</td>
<td>13.67</td>
<td>9.30</td>
<td>2.03</td>
</tr>
<tr>
<td>Time 2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>D score Time 1</td>
<td>10</td>
<td>22.91</td>
<td>47.03</td>
<td>34.57</td>
<td>7.40</td>
</tr>
<tr>
<td>D score Time 2</td>
<td>10</td>
<td>32.51</td>
<td>60.02</td>
<td>48.76**</td>
<td>9.46</td>
</tr>
<tr>
<td>Speech rate Time 1</td>
<td>10</td>
<td>25.5</td>
<td>41.75</td>
<td>32.55</td>
<td>5.31</td>
</tr>
<tr>
<td>Speech rate Time 2</td>
<td>10</td>
<td>37.75</td>
<td>55.25</td>
<td>46.83**</td>
<td>6.24</td>
</tr>
<tr>
<td>Disfluency Time 1</td>
<td>10</td>
<td>.19</td>
<td>.46</td>
<td>.31</td>
<td>.097</td>
</tr>
<tr>
<td>Disfluency Time 2</td>
<td>10</td>
<td>.10</td>
<td>.56</td>
<td>.26</td>
<td>.142</td>
</tr>
</tbody>
</table>

** significant difference between Time 1 and Time 2, p < .01
There were highly significant improvements in total utterances, lexical diversity (D) and speech rate (p<.01); but small and non-significant reductions in mean length of utterance and disfluency. In other words, they produced more overall, with wider vocabulary choice and at a faster rate. This supports the snapshot view presented earlier that oral scores evidently improved, but present interesting differentiations as to how that oral improvement is actually assessed. Particularly of note is the discrepancy between the evidence of more total spoken output, produced at a faster rate, but with no significant reduction in disfluency or increase in length of individual utterance. Taken in conjunction with the finding earlier of no significant improvement in grammatical accuracy (see Table 1 above), this suggests that improvement in oral fluency could be ascribed to easier lexical retrieval due to wider vocabulary, and faster articulation whether accurate or not rather than a more developed grammatical repertoire. However, more in-depth analyses of the lexical and grammatical range, as well as more fine-grained temporal analysis of the speech data, which are beyond the scope of this paper, are needed to substantiate this claim.

In this cohort, overall improvements were found more evidently in speaking than in writing. In order to see if these differences were due to individual variation in quantity of language exposure, we cross-checked reports of individual language usage during the SA period with the pre and post test results. The quantitative language reports (following Wright 2013) were simple forms, so as to be very quick to complete, and designed to capture hours of language use across the four language skills over each day of a typical week. Dividing the eventual total by 7 provided a daily average, presented here as overall average and average for speaking. However, we were limited in how we could use these reports due to low numbers of returns. Seven participants provided data at Time 1 (within six weeks of arrival), summarised in Table 6 below.

Table 6: Average daily hours of language use, and speaking at Time 1

<table>
<thead>
<tr>
<th></th>
<th>N</th>
<th>Minimum</th>
<th>Maximum</th>
<th>Mean</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Average usage</td>
<td>7</td>
<td>.32</td>
<td>2.81</td>
<td>1.45</td>
<td>.84</td>
</tr>
<tr>
<td>Average speaking</td>
<td>7</td>
<td>.75</td>
<td>4.86</td>
<td>2.24</td>
<td>1.46</td>
</tr>
</tbody>
</table>

Only three participants provided further reports for a mid-point and end-point of SA, so the reports cannot be used for statistical association with the linguistic measures. However, in view of the increase in oral scores discussed above, it is interesting to note the participant with the highest average hours of speaking during the whole period of study (over 4.5 hours both at the beginning and end of SA) overall did improve markedly in her oral scores (7.0 out of 10 at Time 1, 8.8 at Time 2). The participant who had the lowest average hours of speaking (below 3.5 across the SA period) did not markedly improve on her oral scores (5.0 out of 10 at Time 1, 5.6 at Time 2). The third participant who had the most marked increase in average hours of speaking across the year (2.71 at the beginning to 3.7 at the end) also consistently scored amongst the highest on all measures at Time 1 in both oral and written exam scores. These patterns suggest there may be a complex linguistic-affective “threshold” effect (Wright & Schartner 2013), where those who are already feeling proficient and/or confident before SA can then engage more fully with, and benefit more, from the target language environment, while those who do not feel proficient may not feel able to engage as fully, and may not make noticeable gains over time. Further qualitative investigations using a more substantial report method (such as Regan et al. 2009) would allow this claim to be substantiated.
4. Discussion and implications for future study

This study focused on identifying changes in written and oral production comparing pre and post Study Abroad, with mixed results as seen above. We identify three key strands emerging in evaluating these results in terms of linguistic outcomes, methodological limitations, and pedagogic issues, all of which have implications for future research in L2 Mandarin learning and teaching.

Analysing the linguistic data, we saw that oral scores improved generally, and often significantly, in line with expectations, and confirms the assumption in SA literature that oral proficiency is most clearly boosted by exposure to the target language in the target country. This exploratory study was not designed in experimental terms, so we cannot speculate here how different the findings would have been with an intensive stay at home group, as examined by Freed et al. (2004). However, we found marked individual variation between aspects of the oral measures, so more detailed analysis is needed on the specific changes in speech rate and other fluency measures found here to understand the scope and effects of these individual variations. Further analysis is also required of the improved lexical range noted here, to test whether the change represents greater use of existing taught formulaic chunks, or represents novel lexical development, although this analysis needs greater development of reliable methodological tools for investigating L2 Mandarin speech.

Evidence of development from the written data is less clear; however, within the data discussed here, there is some evidence of expected increases in proficiency in terms of overall length of writing produced, and in grammatical complexity, indicating improvements had occurred despite lack of change in the summative accuracy ratings as used for examined writing.

Some of the tasks showed wide ranges and evident individual variation in responses, which may reflect differential effects of SA on individuals, and we argue there could be a linguistic-affective ‘threshold’ where those with less confidence or language proficiency found it harder to engage, which then affected the rate of development. Quantitative data designed to tap individual experiences of interaction were too few to use for statistical association. Further qualitative research into the nature of language interaction and experiences in the target country would also provide a rich source of insights into individual differences, to identify some of the complex interaction of factors that would be expected to impact on language learning and individuals’ rate of development (Kinginger 2011).

Methodologically, in analysing the written data, clear markers of grammatical development were not as evident as hoped, as the tasks did not yield sufficient numbers of target morphemes, and not all tasks were completed by the whole participant pool. A wider range of target morphemes, using specific elicitation tasks, designed to tap developing complexity, together with ensuring more consistent task completion, would minimise these problems in future research.

Analysing the oral data, we found that due to the nature of the spoken examination conditions, there was some lack of consistency in task performance between Time 1 and Time 2, and evidence of individual variation in how the tester conducted the test. For example two of the participants ran out of time in completing all four sub-tasks.

Further difficulties arose when looking for ways to transcribe and analyse the data reliably for fluency and accuracy. There seems to be a serious gap in existing methodological practice for analysing L2 Mandarin oral proficiency which needs addressing urgently, including establishing reliable norms for transcribing characters as single or bi-morphemes, to make comparisons with other L2 research using syllables, morphemes or lexemes, to overcome the problem of what constitutes a “word” in cross-linguistic analysis. We also could not find any existing guidance on how to transcribe L1-influenced Mandarin or other
Chinese interlanguage forms in tone and pronunciation of certain difficult phonemes. However, by establishing our own norms (co-referenced between 3 linguistically-trained speakers of Mandarin), we were able to construct and analyse the data as shown above. But it was not possible in the scope of this study to establish, for example, clear evidence of how to assess lexical frequency of content or functional elements, or evidence of formulaic chunks, which are commonly examined in other L2 speech research (Wray 2000). We call for future research into how to normalise analysis of L2 Mandarin speech to investigate such areas.

Another issue in this study is the reliability and validity of our pre/post-test design. Data collected under exam conditions (Time 1) and in non-exam conditions (Time 2) may have affected participants’ attitudes to completing the tasks; failure of several participants to complete the timed written data at Time 2 suggests this may have been the case\(^3\). We also noted a degree of individual variation in how the interviewer and interviewee conducted the oral tasks, leading to some imbalance between the lengths of focused time on each task across the sample.

Nevertheless, we argue that the ecological validity of using the same examination format to collect data pre and post-SA gives us a useful insight into the language learning process from a pedagogic point of view; in post-hoc discussions, the teachers and students themselves expressed a clear preference to finding out how their authentic learning was progressing using the examination format, and suggested that they would have been less keen to participate if additional data elicitation methods had been required. So, in future research design, there needs to be a judicious elicitation balance between recruiting participants and constructing adequately reliable data collection instruments.

Pedagogical implications of how study abroad can measurably aid language change, as highlighted in this study, is also an area to pursue further. While we acknowledge the limited scope of this study, based on a single institution’s cohort of Mandarin learners, we argue that the deeper insights into language change found in the linguistic analyses of the oral and written data compared to the examination outcomes means that it is vital to build up discussions between teachers and linguistic researchers to understand better whether and how to include linguistically-principled measures within assessments of proficiency.

This also gives impetus to much needed research to benchmark examinations in L2 Mandarin, given its rapid expansion throughout the UK and beyond, to ensure how teachers and researchers can work together to see how linguistically-informed elements can be included, in a graded, sequential way, within summative measures of L2 Mandarin proficiency, as is well established for assessments of L2 English such as the Oxford Online Placement Test (Oxford English Testing 2013, Purpura 2004).

We therefore believe it is very timely to establish a broad research agenda into teaching methods, and classroom input/output in L2 Mandarin classes. It has been known for many years that instructed input may not always lead to effective learning of language knowledge (intake, or linguistic competence) or capacity to produce (output, or communicative competence) in L2 English or European languages (Ellis et al. 2009; VanPatten 2003). We also know for many European languages there are clear predicted stages of development (Pienemann 1998; Vainikka & Young-Šcholten 2005). It is assumed that instruction may speed up the rate of development, but is unlikely to alter the route of development, whether in more formal or more communicative classrooms (Norris & Ortega 2000). However, these claims have not yet, to the authors’ knowledge, been validated for L2 Mandarin.

\(^{33}\) As with many YA studies, comparable pre-post data collection issues are compounded by participant drop-out: out of 22 participants originally recruited at Time 1 in this study, complete data sets were only available for eight participants on the written data, and ten on the oral data.
Tracking such stages may be difficult, in addition, depending on teaching practices. Given an assumed traditional value in Chinese pedagogy placed on drilling and recitation, e.g. in L2 English (Jin & Cortazzi 2006), it could be hypothesised that current expectations of L2 Mandarin development may to some extent consist of building up greater skills in producing memorised chunks of language, tied to the content of a standard syllabus, particularly as delivered by Chinese-trained native Mandarin speakers. But we need to test this hypothesis empirically, to see how far L2 Mandarin language development reflects instructed input, or rather follows a more linguistically-driven developmental trajectory.

It is worth also exploring the range of current language pedagogies being used: it would be useful to see how far learners and teachers have similar or different attitudes to traditional teacher-fronted versus more communicative learner-centred practices, and to see if there are differences in teaching practice in different settings or between learners of Mandarin from different educational contexts. So would the Chinese learners of English referred to by Jin and Cortazzi (2006) be similar or different to the English learners of Mandarin in this current study? And how would they compare to other Asian or African learners of Mandarin, in China or in local L1 contexts? The rapid rise in Mandarin classes at universities and schools in all these different settings could provide a good opportunity to bring theory and practice together to clearly assess what constitutes most effective instruction across different global contexts when aiming to build both language knowledge (linguistic competence) and language use (communicative competence) for this new area of language learning.

5. Conclusion

This study tracked the progress of ten UK university students of L2 Mandarin in written and speaking performance before and after eight months’ Study Abroad (SA) in China. Despite some methodological challenges, clear results were seen in both writing and speaking. Some development in written proficiency was partially found, especially in writing length and improvement in one measure of grammatical complexity (use of the de-relative morpheme), though improvement was not seen in broad-brush examination scores of accuracy. Expected significant improvement in oral proficiency were found using broad-brush examination scores for accuracy, and in linguistic analyses of speech rate, number of utterances and lexical diversity, supporting findings from SA research in other L2s. However, not all measures of oral proficiency showed consistent significant improvement; this suggests that SA effect on oral proficiency was not as robust as predicted, and that more fine-grained analysis is needed to illustrate the complex nature of oral proficiency.

These differences between proficiency as measured by examination scores, or by analyses of linguistic development, raised questions of how SLA research questions and language pedagogy intersect – how to assess what is learned through input in and out of the classroom and how it should be tested, and how to benchmark best practice in L2 Mandarin pedagogy and assessment.

Collecting and analysing the data revealed many gaps in existing L2 models and methodological conventions of linguistic knowledge as applied to L2 Mandarin; one crucial gap lies in transcribing and analysing oral data, such as how to annotate interlanguage forms, and how to equate characters to the more usual terms of syllable, morpheme or word.

Indeed, we have found very few research studies published in English relating to longitudinal research in L2 Mandarin, especially combining a linguistic and pedagogic perspective as this study sought to do. Therefore, despite its limitations, we believe the innovative nature of our study retains validity for providing a starting point in analysing the development of L2 Mandarin, and stimulating some suggestions for urgently needed lines of research.
References


Yuan, B. (2013). Personal communication with lead author (March 21, 2013).


Wright & Zhang

Examining the Effects of Study Abroad on Mandarin Chinese Language Development among UK University Learners

(Corresponding author)
Clare Wright
Department of English Language and Applied Linguistics
University of Reading
Whiteknights
PO Box 218
Reading, RG6 6AA
United Kingdom

c.e.m.wright@reading.ac.uk