[ð]
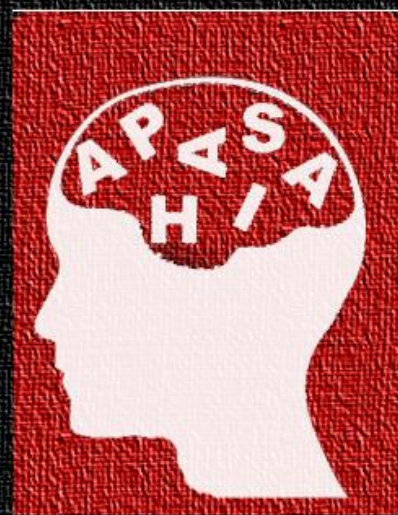
[æ]

[ə]

Newcastle and Northumbria Working Papers in Linguistics

Volume 21.2 (2015)

NEWCASTLE AND NORTHUMBRIA WORKING PAPERS IN LINGUISTICS

VOLUME 21.2 (2015)

CONTENTS

EDITORIAL NOTE

Newcastle Working Papers in Linguistics aims to publicise the research of the Centre for Research in Linguistics and Language Sciences (CRiLLS) as well as external contributions. This volume exhibits a focus on Asian languages (Arabic, Chinese and Japanese) as well as signed language and gesture – considered from a variety of theoretical perspectives. The papers were contributed by postgraduate students at Newcastle University and members of other institutions.

The editors would like to thank the anonymous reviewers for working with us to support this edition with their expertise.

*PERFECTIVE MARKING IN L2 MANDARIN: AGREEMENT WITH INHERENT LEXICAL ASPECT OR THE DERIVED SENTENTIAL ASPECTUAL CLASS?**

DANIEL BELL AND CLARE WRIGHT
(Newcastle University / University of Reading)

Abstract

This study evaluates the differing claims of the Aspect Hypothesis (Anderson & Shirai 1996) and the Sentential Aspect Hypothesis (Sharma & Deo 2009) for perfective marking by L1 English learners of Mandarin. The AH predicts a narrow focus on inherent lexical aspect (the verb and predicate) in determining the use of the perfective marker *le*, whilst the SAH suggests that – subject to L1 influence – perfective marking agrees with the final derived aspectual class of the sentence. To test these claims data were collected using a controlled *le*-insertion task, combined with oral corpus data. The results show that learners' perfective marking patterns with the sentential aspectual class and not inherent lexical aspect (where these differ), and that overall the sentential aspectual class better predicts learners' assignment of perfective marking than lexical aspect.

1. Introduction
1.1. Models of aspect

Following Vendler's (1967) categorisation of verb types, treatments of aspect began to consider the compositional nature of aspect (notably explored by Verkuyl 1972) and the fact that elements beyond the verb can crucially contribute to aspectual derivation. This fed usefully into a number of second language (L2) acquisition studies in the subsequent decades, which initially considered only the role of verbal arguments (e.g. Slabakova 1999), with other elements – like adverbials – not coming into focus until more recently (e.g. Baker and Quesada 2009). In fact, because the vast majority of research into the acquisition of aspect has investigated the Aspect Hypothesis (AH) (Andersen & Shirai 1996), which is characterised by an exclusive focus on inherent lexical aspect (the verb and predicate), the fully compositional nature of aspect represented in theoretical models has yet to be seriously unpacked in terms of its implications for acquisition. An important advancement, though, came from the Sentential Aspect Hypothesis (SAH) (Sharma & Deo 2009), which proposes that learners' aspect marking patterns with the sentential aspectual class (the final aspectual derivation) rather than with the verbal predicate alone, and that the AH's findings concerning lexical aspect are an epiphenomenon of sentential agreement.

The term aspect denotes the 'internal temporal constituency of a situation' (Comrie 1976: 3), and describes how eventualities unfold in time, whilst tense is deictic, locating a situation in relation to the speech time as past, present or future. Treatments of aspect can be divided into two main camps: two-component models that distinguish lexical from grammatical aspect, and unitary models, which do not. The AH is rooted in Vendler's (1967) verb classification and in a two-component model of aspect, whilst the SAH emerges from the

---

unitary approach of the semantics literature (e.g. de Swart 1998). In the former approach, grammatical aspect morphology provides ways of viewing a situation (e.g. as having internal duration or as a completed event), whilst lexical aspect (sometimes called *aktionsart,* 'kinds of action') pertains to the inherent properties of situations themselves, defined using binary features (e.g. telic/atelic, stative/dynamic, durative/punctual). Lexical aspect is derived from verbs and predicates, as in Vendler's (1967) verb categorization (discussed in Section 2.2). Representative of the two-component approach is Smith's influential model (1997) and that of Xiao & McEnery (2004). In such models, grammatical perfective/imperfective viewpoint markers express the dichotomy between viewing situations as a single whole (with endpoints) as opposed to referencing their internal duration (and excluding endpoints). For example, *He walked to school* and *He was walking to school* respectively express the perfective and imperfective viewpoint on the same telic situation (*walk to school*).

Of direct relevance to this study (and specifically to the Sentential Aspect Hypothesis) is the fact that these models account for aspectual shifts performed by elements beyond the verbal predicate. Thus in Smith's model, Vendler's fourway ontological categorisation is applied to the sentence-level. Similarly, Xiao & McEnery (2004) show how sentential aspect is derived through mappings between three levels: the predicate, the predicate and its arguments, and the sentence. In these models, universally invariant situation types are derived lexically and subjected to fine-grained aspectual viewpoints encoded by language-specific grammatical markers. However, alternative treatments of aspect have emerged which accomplish this without utilizing a distinction between lexical situation types and viewpoint markers.

The semantics literature (e.g. Mourelatos 1978, Dowty 1986, Krifka 1989) has given rise to unitary models in which both lexical and grammatical devices can act as aspectual operators, modifying the derived aspectual class of the sentence through their basic mereological properties (see de Swart 1998). Three ontological eventuality types are distinguished − states, processes and events − which differ according to whether they are homogenous or quantized, and stative or dynamic. In this framework, states and processes are homogeneous, denoting divisive reference (subintervals of *being happy* equal *being happy*) and cumulative reference (*reading* plus *reading* is simply *reading*), whilst events are non-homogeneous and quantized (subintervals of *eat an apple* are not equal to the whole, and unlike states and processes, events can be counted).

On this approach atelic verbs (e.g. *know),* habitual adverbials (e.g. *everyday*) and traditional aspect markers like the English progressive are all homogeneous. Correspondingly, cardinally quantified adverbials (e.g. *for five years*), telic verbs (e.g. *recognise*) and traditional perfective markers are all quantized. Crucially, then, both two-component and unitary approaches emphasise the compositionality of aspect and the role played by elements beyond the verbal predicate (e.g. Dowty 1986: 43, Xiao & McEnery 2004: 80). Moreover, both approaches can be applied to acquisition studies, as the AH and the SAH respectively demonstrate, although to date relatively little research has been done on non-European languages, leaving a gap in our understanding of how languages which are typologically different may show effects of transfer within either approach. Hence, this paper addresses the issue of the acquisition of aspect by English learners of Mandarin, given that the two languages are often held to differ in being a tense vs. aspect language respectively.

Having now summarised the main approaches to aspect, a more detailed discussion of the key concepts is provided in the following sections, which are organized as follows. The next section defines telicity and reviews Vendler's (1967) verb classification. Section 3 considers differences in aspectual derivation in English and Mandarin, as well as Mandarin perfective *le*. Section 4 examines the AH and the SAH, evaluating problems with the claim − which underpins the research questions of this study − that inherent lexical aspect exclusively conditions learners' aspect marking. The following sections then provide details of our

2

methodology (Section 5), results (Section 6), discussion (Section 7) and conclusions (Section 8).

2. Composition of aspect

This section explores the concept of telicity and reviews Vendler's (1967) verb classification, which underlies the Aspect Hypothesis.

2.1. Telicity

Telicity is central to the discussion of aspect, and in the broadest sense denotes boundedness (cf. Jackendoff 1996). However, Garey (1957: 105), who introduced the term to the field of aspect, posited an entailment test to identify inherent verbal telicity: if one was *verbing*, and was interrupted while *verbing*, has one *verbed*? The answer for telic verbs, such as *drown*, is no, whilst for atelic verbs (e.g. *swim*), the answer is yes. This definition of telicity thus concerns eventuality structure, with telicity involving the notion of completion and the attainment of a *telos* (goal/endpoint). But the semantics literature has also given rise to mereological definitions, which pertain to the relation between parts and the whole.

Thus telicity has been defined as the lack of the subinterval property (i.e. Krifka's (1998) notion of quantization) such that a telic sentence (e.g. *He walked to the school*) is not true at any subinterval of the interval for which the proposition holds; atelic sentences (e.g. *He walked*), on the other hand, have the subinterval property, being equally true at any subinterval of the interval for which they are asserted. However, not all bounded predicates are quantized or even telic in the sense of having an inherent endpoint in eventuality structure (a *telos*, in Garey's original sense). For example, some (e.g. *study a quantity of books*) have a *telos* (i.e. when the books have been studied) and are bounded but yet are not quantized because they possess the subinterval property (i.e. a subpart of *a quantity of books* is still *a quantity of books*) and cumulative reference (cf. Krifka 1998: 220-4). Likewise, other expressions are quantized, and hence yield sentences which are 'perfective' in the SAH account but not telic in the sense of Vendler (1967) or the AH (e.g. *wait 2 hours*). Both, however, are bounded in a general sense. This shows that the definitions of aspect adopted in the AH and the SAH do not fully overlap and moreover certain boundedness phenomena are not captured in either account.

Additionally, in contrast to the mereological approach which can be applied across syntactic levels, telicity has also been treated as a fundamentally spatial property (e.g. Xiao & McEnery 2004), such that its primary application is to nominals (only secondarily creating boundedness in eventuality structure). This fact prompted Xiao & McEnery to observe that 'spatial delimitedness [i.e. of an NP argument] always implies temporal boundedness [a 'telic' situation type] but the reverse is not true' (Ibid. 188). This has been illustrated for incremental theme predicates like *eat a cake* (Dowty 1991), in which an endpoint in eventuality structure and in the temporal dimension are both derived from spatial delimitedness (i.e. the endpoint of the activity arises from the bounded nature of the argument). Having now considered telicity, Vendler's aspectual classification based on the semantics of verb types will be briefly outlined.

2.2. Verb-based aspectual classes

Inherent lexical aspect, central to the Aspect Hypothesis, has as its basis Vendler's (1967) classification of verbs into states, activities, accomplishments and achievements, according to their telicity/atelicity, whether they denote a point-in-time or a duration (i.e. are punctual/durative) and whether they are dynamic or stative (i.e. whether or not they involve change). The four verb types can be summarized as follows (cf. Croft 2009: 6):

States: stative, durative and atelic (*be American, love)*
Activities: dynamic, durative and atelic (*sing, dance)*
Achievements: dynamic, punctual and telic (*shatter, discover, reach [the summit]*)
Accomplishments: dynamic, durative and telic (*cross [the street], read [the book]*)

Thus *love* (stative) and *sing* (activity) are atelic, with no terminative point, whilst *read the book* (accomplishment) and *discover* (achievement) are telic, both involving endpoints – although differing in that accomplishments are intrinsically durative with a point of completion whilst achievements are instantaneous. The classification is limited to verbs, incorporating the predicate level mainly to define accomplishments (e.g. *run to school*, in which the object provides an endpoint). This thus assumes that verbs come pre-specified with aspectual features, but does not adequately account for the role of arguments and other sentential elements. In the next section, we will consider how perfective marking is used in Mandarin against the backdrop of differences in aspectual derivation in English and Mandarin.

3. Boundedness marking in English and Mandarin

Aspectual derivation in English and Mandarin will now be contrasted to clarify the influence of learners' L1 upon Mandarin *le* acquisition, transfer being predicted by the SAH but not by the AH (these hypotheses are set out in full in Section 4).

3.1. Parametrized differences in telic eventualities

We focus here on a crucial parameterized difference in the means of deriving telic eventualities, namely that English relies mainly upon nominal devices, whilst Mandarin makes greater use of verbal devices (cf. the cross-linguistic generalization, stated in Kabakciev 2000: 156 and discussion of the distinction in Slabakova 1999). It should be pointed out that despite Mandarin sometimes being considered an 'aspect language' and English a 'tense language' (i.e. without grammatical perfective marking) (e.g. Xiao & McEnery 2004: 2), the aspectual function of bounding associated with perfective marking is accomplished nominally in English by corresponding grammatical devices (e.g. the definite article). [1] In fact, definite/indefinite marking and perfective/imperfective marking respectively can be regarded as 'equivalent techniques of quantification in the nominal and verbal domain' (Leiss 2007: 1), with perfective marking creating definiteness and imperfective marking creating indefiniteness. This suggests that the task of learners of Mandarin with a nominal-marking L1 (e.g. English) is not the acquisition of a new grammatical category, but rather to reset the locus of boundedness marking to the verbal domain (something which the learners in this study appeared to achieve; see Section 6 and 7). Next, the nature of *le* as a boundedness marker will be considered.

3.2. Perfective marking in Mandarin

In Mandarin, verbal-*le*[2] is a perfective marker that occurs after the verb and indicates an event viewed as a bounded whole, whether 'temporally, spatially or conceptually' (Li & Thompson 1989: 185). On this view, *le* is therefore distinct from (past) tense. Li & Thompson mention

---

[1] This is connected to the typological generalization that 'aspect languages [e.g. among others, Mandarin, Russian, and the Slavic languages] avoid article systems, and article languages [e.g. English] avoid aspect systems' (Leiss 2007: 87).

[2] Hereafter, simply *le*, since sentence-final *le*, a marker that denotes a 'currently relevant state' (Li & Thompson, 1989) is not examined in this study.

four types of boundedness that typically lead to *le* usage: quantified events, definite/specific events, events in which the verb's semantics contain a bound (e.g. *fall asleep*) and events that are first in a sequence. Example (1) shows *le* being triggered by quantification, namely through the boundedness arising from the duration adverbial *sanshi fenzhong* ('thirty minutes') which bounds the verb in the second instance (whilst the first use of the verb is aspectually neutral):

(1)   Zhangsan zai bowuguan men kou   **deng** Lisi, **deng le**   sanshi   fenzhong
        Zhangsan at museum   door mouth wait Lisi, wait PFV thirty   minute
        'Zhangsan waited for Lisu at the museum entrance for 30 minutes.' (Ibid.: 189,
        emphasis added)

        Also, the quantification triggering *le* can be spatial (i.e. material) rather than temporal, as (2) below shows:

(2)   Ta   chi   le   san   ge   pingguo
         he   ate   PFV three CL   apples
        'He ate 3 apples.'

In (2) the boundedness arises through a cardinally quantified object NP. This technique for deriving telicity (i.e. through object-marking) is used extensively in English, with VP telicity mostly being derived in accomplishment and activity predicates through nominal quantizing devices that delimit the object and thereby provide an endpoint to the eventuality (cf. Slabakova 1999).
        Examples (3), (4) and (5) below illustrate the other types of boundedness that Li and Thompson say lead to *le* usage: a definite or specific object, an inherently bounded verb and the first event in a sequence (examples from Ibid. 192, 197, 199):

(3)   Wo   pengdao   le   Lin Hui
        I   encounter   PFV   Lin Hui
        'I ran into Lin Hui.'

(4)   Gaizi diao   le
        lid   fall   PFV
        'The lid fell off.'

(5)   Ta   kai   le   men, ni   jiu   jin qu
        he   open PFV door,   you   then in go
        'When he opens the door, you go in.'

Next, the acquisition of aspect morphology as predicted by the Aspect Hypothesis and the Sentential Aspect Hypothesis will be discussed.

4. The acquisition of aspect

In this section the Aspect Hypothesis and the Sentential Aspect Hypothesis will be examined more closely, and limitations of the AH discussed to justify the need for a fully compositional account that does not stop short at the predicate level. The findings of Sharma & Deo's (2009) study of sentential aspect effects among Hindi learners of English – the methodology of which is reflected in the *le*-insertion task in the present study – will then be discussed. After that, we

will consider the role of temporal adverbials, which are excluded from the AH's lexical aspect account but included in the SAH (and central to the present study).

4.1. The Aspect Hypothesis

The claims of the Aspect Hypothesis are shown below (quoted from Andersen & Shirai 1996: 533), and this study is concerned with claim 1. They are formulated for L1 acquisition, but are also argued to hold in L2 acquisition:

> 1. Children first use past marking (e.g. English) or perfective marking (Chinese, Spanish etc.) on achievement and accomplishment verbs, eventually extending its use to activity and stative verbs.
> 2. In languages that encode the perfective-imperfective distinction, imperfective past appears later than perfective past, and imperfective past marking begins with stative verbs and activity verbs, then extending to accomplishment and achievement verbs.
> 3. In languages that have progressive aspect, progressive marking begins with activity verbs, then extends to accomplishment or achievement verbs.
> 4. Progressive markers are not incorrectly overextended to stative verbs.

The AH regards the aspectual class of verbs (classified according to Vendler's (1967) four-way categorization) as universally determining the emergence of grammatical aspect markers, such that morphological marking appears in a fixed order determined by verb type: first on achievement verbs, then on accomplishments, and only thereafter on activity verbs and (finally) on stative verbs (cf. Li & Shirai 2000: 50). Grammatical aspect marking is thus posited as patterning with lexical aspect (verbs and predicates, represented by Vendler's four idealized situation types). The fact that the AH predicts agreement with lexical aspect, so defined, means that other elements that perform aspectual shifts (such as temporal adverbials) are not predicted to affect learners' assignment of aspect markers and do not influence coding. The claims above are hypothesized to hold cross-linguistically, precluding L1 transfer and the influence of properties of the L2.

A significant number of SLA studies broadly support the above claims (e.g. for English, Bardovi-Harlig et al. 1998; for Spanish, Cadierno 2000), although claim 4 is the most tenuous (contradicted, for example, by the occurrence of *-ing* on stative verbs in Robison 1990). Studies on non-European languages are relatively few, but corroborating findings have been found, for example, for Japanese (Shirai 1995; Shibata 1999) and Mandarin (e.g. Jin & Hendriks 2005, Duff & Li 2002). However, at the same time, other studies have yielded contradictory findings (e.g. Dietrich, Klein & Noyau 1995, Salaberry 1999, Rohde 1996, Rocca 2002), whilst others provide evidence of transfer (e.g. Slabakova 1999, Laleko 2008, Gabriele 2009) or highlight the role of temporal adverbials (e.g. Salaberry 2013), which constrains the universality of the AH's claims. Therefore to move forward it is necessary to more precisely determine how L1/L2 factors and other sentential elements (e.g. adverbials) delimit the patterning of aspect marking with inherent lexical aspect.

4.2. Critiques of the Aspect Hypothesis

This section will briefly discuss three factors that delimit the AH's universalist claims. These are L1 transfer, the internal incoherence of the Vendlerian view of aspect, and its exclusion of temporal adverbials (further critiques can be found in Sharma & Deo 2009: 5-6).

Firstly, the AH's claims are tempered by findings showing the effects of the L1 upon the acquisition of tense/aspect morphology. This has many manifestations, but one example is

that because languages differ in where aspectual information is located, where learners initially look for aspectual information has been shown to be conditioned by their L1. Thus Slabakova (1999) observes that in English, telicity is encoded by the cardinality of nominal arguments (e.g. *eat cakes* is atelic; *eat two cakes* is telic), but in Slavic a preverb is used (cf. Slabakova 2005). She found that beginning Slavic learners tended to interpret English telic sentences as atelic because they are not sensitive to the object's cardinality, and so aspect marking reflects the absence of a verbal telicity marker. Thus in such cases beginning learners are not sensitive to the predicate's aspectual class (as predicted by the AH) because of L1 transfer.

Secondly, the AH conflates the inherent aspect of verbs with VP aspect, a confusion which stems from Vendler's original classification. Thus the AH states that 'learners will initially be influenced by the inherent semantic aspect of *verbs or predicates* in the acquisition of tense and aspect markers' (Andersen & Shirai 1994: 1, emphasis added). But this is problematic because these do not always correspond. That is, this prediction does not elucidate familiar cases such as the activity verb *run* which can yield an activity predicate or an accomplishment predicate depending on its arguments: *run miles* (activity; atelic); *run a mile* (accomplishment; telic). A similar situation arises with Mandarin resultative verb compounds (RVCs), which combine an activity verb with a resultative complement (a further verb or adjective) to yield achievement predicates. In these cases the predicate's aspectual class differs from that of the main verb (e.g. *shuo-wan* 'speak-finish'). This approach effectively works backwards from the predicate's aspectual class, which is compositionally derived (e.g. *run a mile* = accomplishment VP and *run miles* = activity VP), in order to assign individual verbs to an aspectual class (or sometimes more than one) and hypothesize that aspect marking agrees with these verb types. But this approach is flawed because it misses the compositional nature of the aspectual derivation.

Thirdly, with direct relevance to the present study (and in particular to the *le*-insertion task), the aspectual function of temporal adverbials in acquisition studies calls for closer attention (cf. Salaberry 2013: 207), as it appears that they can have a strong influence on learners' selection of aspect markers (cf. Slabakova & Montrul 2008). They emerge before both tense and aspect morphology and initially substitute for morphological marking (e.g. Noyau 2002: 107, Starren 2006). Thereafter, they play a role in aiding the assignment of tense/aspect morphology, although tense markers have been the focus of most studies (e.g. Musumeci 1989), and relatively few have considered the impact of temporal adverbials on aspect marking.

Temporal adverbials alter the aspectual class of the base predicate, and it has been found that learners are sensitive to these effects. For instance, Salaberry (2013: 210) notes the role of duration adverbials in triggering perfective marking in Spanish. Moreover, the learners of Spanish in Slabakova & Montrul's (2008) study demonstrated a strong sensitivity to the aspectual effects of completive and duration adverbials. Additionally, in Baker and Quesada's (2009) study, temporal adverbials conditioned learners' use of preterit and imperfective marking, particularly helping intermediate learners to use these aspect markers accurately. However, problematically, aspect shifts produced by temporal adverbials are excluded from coding in AH studies. Hence Shirai (2013: 298) notes that in both of the following examples *hid in the attic* is an activity predicate i.e. ignoring the fact that at the sentential level (6) is temporally bounded whilst (7) can be construed as an achievement because of the punctual adverbial:

(6)   He hid in the attic for an hour.
(7)   He hid in the attic when the sheriff arrived.

The above findings (which corroborate those of the present study) suggest an important role for temporal adverbials during the acquisition of aspect and call for a unified account of

aspect that incorporates their effects as well as those of the verb and predicate. To that end, the next section examines Sharma & Deo's Sentential Aspect Hypothesis.

4.3. The Sentential Aspect Hypothesis

The Sentential Aspect Hypothesis (2009) attempts to offer a more complete account of the acquisition of aspect that incorporates the effects of lexical aspect by positing that learners are sensitive to sentential aspectual class (to which the verb and predicate contribute) and that the L1 and the L2 also play a role in determining the nature of the emergent system. Thus, the SAH claims that, conditioned by the L1,

> Learners hypothesize that morphological marking is a form of agreement with the aspectual class of the sentential predication (not narrowly with lexical aspect alone) (Sharma & Deo 2009: 7).

Sharma & Deo (2009) tested Hindi speakers to see whether they retained sensitivity to sentence-level imperfectivity, present in their L1, or whether their acquisition of English was guided solely by the aspectual class of the L2 English verbs/predicates (as the AH predicts). In Hindi (but not English), all past eventualities must be marked morphologically as perfective or imperfective; whilst in English, the past tense marker *-ed* is compatible with both aspects: *I lived in Bombay three years ago* can denote a bounded, completed eventuality (perfective) or one that still holds (unbounded, imperfective) (Ibid. 8). They coded production data for predicate and sentential aspect, and it was found that sentential perfectivity was a significantly stronger trigger of past tense marking than VP-level telicity.

Additionally, cases of misalignment between inherent lexical aspect and sentential aspect were considered in order to clarify the nature of learners' aspectual sensitivity. Sharma & Deo found that perfective sentences containing atelic predicates overwhelmingly triggered past tense marking (84.6%), contrary to the predictions of the AH that aspect marking always follows inherent lexical aspect. The AH predicts that past tense marking should be low frequency with atelic verbs and predicates, irrespective of subsequent aspectual operations. This generally held true, because as (8) below illustrates, sentential aspect is normally the same as predicate aspect. However, (9) and (10) show how sentential operators can shift the aspectual class from that of the base predicate, here causing a dramatic increase in learners' use of past tense morphology and revealing a sensitivity to sentential aspect rather than VP aspect alone (examples all have past time reference and are from Ibid. 18):

(8)   Lexical aspect: *Activity* Sentential aspect: *Imperfective*
        a. I work with the French people, no?
        b. I study in Punjab also, I study in Delhi also. Because of moving.

(9)   Lexical aspect: *State*; Sentential aspect: *Imperfective → Perfective*
        a. **For first 12 year** I was there because my father was posted there.
        b. **Six months** I was there in the kitchen.

(10)   Lexical aspect: *Activity*; Sentential aspect: *Imperfective → Perfective*
        a. I worked **for 14 years**. That's enough.
        b. So we did the schooling over there **and then moved**.

The duration adverbials in (9) and (10) (shown in bold) (as well as the second clause in (10b)) bound the atelic eventualities, and derive sentences that are quantized (perfective); the

increase in past tense marking here is predicted by the SAH but not the AH. Likewise, in sentences containing telic predicates, the presence of an imperfectivizing sentential operator dramatically reduced past tense marking (to only 13.3%), whereas when the telicity of the VP was unmodified by higher operators, past tense marking remained high frequency (75.9%).

Given the competing claims of the AH and SAH, this study therefore examined how L1 English learners of Mandarin used the perfective marker *le*. Our research questions were:

1)   Does learners' perfective marking pattern with inherent lexical aspect (i.e. to confirm the AH) or with sentential aspect (i.e. to confirm the SAH)?
2)   How does perfective marking show evidence of development over time, measured before and after a period spent studying abroad in China?

5. Methodology

A controlled experimental task was combined with longitudinal corpus data to examine how L1 English learners of Mandarin used the perfective marker *le*. The experimental task honed in on how two types of temporal adverbials affect perfective marking and the latter free speech data provided a window on developmental changes for a specific group of individual learners.

5.1. *Le*-insertion task
5.1.1. Participants

Ten English learners of Mandarin (the majority of whom were Chinese language students at Newcastle University), completed the *le*-insertion task, which contained sentences prepared in collaboration with a native Chinese speaker. A control group of eight native Mandarin speakers provided a baseline for *le* usage. The learners self-reported their proficiency from beginner to low intermediate (with two advanced) and in most cases it was possible to second-check this via informal discussion prior to completing the questionnaire.

5.1.2. The task

The aim of both Part One and Two of this task was to test whether the AH or the SAH more accurately predicted learners' perfective marking – namely whether learners' perfective marking patterned with inherent lexical aspect or the final sentential aspectual class. Therefore, aspectually mismatched sentences were presented (in a randomized order) to test the effects of habitual frequency and duration adverbials respectively. Learners were instructed to decide whether the sentences were OK as presented, or whether they should insert *le*, and tokens of the perfective marker were totalled and compared across type (1) and type (2) sentences.

All sentence pairs presented had past time reference, indicated by a locative time adverbial (e.g. *shang ge yue* 'last month') and/or by the simple past tense translation beneath each sentence. This eliminated the confounding factor of tense and enabled comparison of aspectual effects, because learners may transfer the past tense value to the perfective marker, and so be biased against marking non-past tense sentences perfectively (cf. Wen 1995).

Tokens of *le* were counted verb-finally only, discounting incorrectly placed tokens (i.e. if *le* was placed sentence-finally rather than verb-finally, this was not counted, because sentence-final *le* is often regarded as a distinct morpheme denoting a 'currently relevant state' (Li & Thompson, 1989: 296) and is not studied in this paper. Where *le* was simultaneously verb-final and sentence-final (due to the verb being sentence-final), following Li and Thompson's criteria these tokens were counted because in this case they unambiguously denoted a perfective event (due to the simple past tense English translation provided beneath

the test sentences). Examples of the sentences used are included below for both parts of the *le* insertion test (and the full list is provided in appendix B).[3]

5.1.3. Part one: imperfectivizing adverbials

In Part One, sentence pairs containing telic verbs were presented, with and without a habitual frequency adverbial (e.g. *meitian* 'everyday') that modifies the sentential aspectual class from that of the telic verb and yields an imperfective sentence.[4] This is a stative mapping (deriving sentences with divisive and cumulative reference from quantized VPs). In each pair, sentence (1) was telic at the level of the lexical verb and the sentence level (because no further operators apply), and so these sentences are predicted to trigger perfective marking according to both the Aspect Hypothesis (AH) and the Sentential Aspect Hypothesis (SAH). These cases provide a baseline value for learners' use of the perfective marker with these verbs. Sentence (2) contains the same telic verb but is imperfective at the sentence level through the addition of a habitual temporal adverbial, which leads to perfective marking being predicted by the AH but not by the SAH in these cases. The AH predicts no difference in the use of perfective marking between these sentence types, because only lexical aspect is considered relevant, whilst the SAH predicts a difference based on their differing sentential aspectual classes. The examples below illustrate the different sentence types:

(1)   *Lexical aspect (telic) = sentential aspect (perfective)*
        Shang   ge   yue   wo   **dapo**   yi   ge   beizi
        last   CL   month   I   **break** one CL   glass
        'Last month I broke a glass.'
(2)   *Lexical aspect (telic) → sentential aspect (imperfective)*
        Shang   ge   yue   wo   **meitian**   **dapo**   yi   ge   beizi
        last   CL   month   I   **everyday** **break** one CL   glass
        'Last month I broke a glass everyday.'

(1)   *Lexical aspect (telic) = sentential aspect (perfective)*
        Ta   **dasi**   yi   tiao   chongzi
        he   **kill**   one CL   insect
        'He killed an insect.'

---

[3] Since completing this study, it has come to our attention that some native speakers do not find some of the sentences used in the *le*-insertion task to be fully natural. This is partly due to the existence of real variation in native speakers' grammaticality judgements, but can also be attributed to the fact that the sentences were presented in isolation without the context of a following clause (e.g. the sentence, *ta shangci deng ni* 'Last time he waited for you', would be more natural with a following clause i.e. *ta shangci deng ni, zheci ye hui deng ni* 'Last time he waited for you, and this time he will also wait for you'). However, by presenting the sentences in isolation, we removed the confounding factor of the aspectual effect of a following clause, which is crucial because in Sharma & Deo's (2009) account these are regarded as aspectual operators cf. Section 4.3 example (10b). We also note that the learners who completed the task did not show any awareness of this unnaturalness (the majority being beginners/low intermediate level), and so this is unlikely to have influenced their *le* usage. The results therefore remain valid concerning the conditioning effect of sentential aspectual class on *le* marking in learners' interlanguage.

[4] The terms *perfective* and *imperfective* are used here to describe the sentential aspectual class, whilst *telic* and *atelic* are reserved for the lexical aspectual class, following Sharma & Deo (2009). Both *telic* and *perfective* denote quantized predications, whilst *atelic* and *imperfective* denote homogeneous predications (cf. discussion of these terms in Section 1.1).

(2)   *Lexical aspect (telic) → sentential aspect (imperfective)*
Ta **meizhou   dasi** yi   tiao   chongzi
he **every week kill**   one   CL   insect
'Every week he killed an insect.'

5.1.4. Part two: perfectivizing adverbials

In Part Two, the type (1) sentences are atelic at the lexical level and this is unaltered by higher aspectual operators, whilst the type (2) sentences are atelic at the lexical level but perfective at the sentential level through the addition of a duration adverbial, which acts as a perfectivizing operator by providing a temporal bound (cf. Moens 1987, de Swart 1998: 357).[5] As in Part One, the type (1) sentences provide a baseline value for perfective marking, this time with stative verbs. Neither the AH nor the SAH predict perfective marking in these cases because they are atelic at the lexical level and imperfective at the sentential level. However, the type (2) sentences yield contrasting predictions. The AH predicts perfective marking with neither sentence type, because both contain atelic verbs and perfective marking is hypothesized to be conditioned solely by the lexical aspectual class. However, the SAH predicts that perfective marking will occur in the type (2) sentences because they are sententially perfective due to the temporal bound.

(1)   *Lexical aspect (atelic) = sentential aspect (imperfective)*
Qunian   ta   zai   nali   zhu
last year he   at   there   live
'Last year he lived there.'

(2)   *Lexical aspect (atelic) → sentential aspect (perfective)*
Qunian   ta   zai   nali   zhu   liang   ge   yue
last year he   at   there   live   two   CL   months
'Last year he lived there for 2 months.'

(1)   *Lexical aspect (atelic) = sentential aspect (imperfective)*
Shangci   ta   zhan   zai   nar
last time   he   stand   at   there
'Last time he stood there.'

(2)   *Lexical aspect (atelic) → sentential aspect (perfective)*
Shangci   ta   zai   nar   zhan   ji   ge   xiaoshi
last time   he   at   there sat   several CL   hour
'He stood there for several hours.'

The next section presents the methodology for the corpus component of the study.

---

[5] These sentences can be seen to contain atelic verbs because the lexical aspect approach uses compatibility with duration adverbials as a test for atelic verbs (Vendler 1967: 101); the subsequent quantizing mapping performed by the duration adverbial is itself ignored by the AH (Andersen & Shirai 1996: 530-1; also see Section 4), but is incorporated into the SAH account.

5.2. Corpus data
5.2.1. Participants

Longitudinal oral corpus data was collected from eight ab initio English learners of Mandarin on a full-time Chinese language course at Newcastle University; at the first time of testing (Time One) they had completed two years of their course and were intermediate level or below (approximating to A2-B1 level in the CEFR).

5.2.2. Data collection

Learners took a ten-minute oral exam at the end of their second year, repeated at the start of their fourth year of study, involving four tasks, conducted one-to-one with their Chinese teacher; the exam data were collected before and after their Study Abroad year in China, providing a longitudinal set of data to enable developmental comparison between Time One and Time Two following immersion in the target language country. The tasks included a planned talk about everyday activities, an unplanned picture description task, a planned role-play and an open discussion about life in China. These conversations were recorded and transcribed by a native speaker using the CLAN (Computerized Language Analysis) software program.

5.2.3. Data analysis

Utterances containing the perfective marker *le* were coded for inherent lexical aspect (achievement, accomplishment, activity and state) using standard diagnostic tests adapted for Mandarin by Chen & Shirai (2010) (reproduced in Appendix A) and as sententially perfective or imperfective (quantized or non-quantized) (following Sharma & Deo, 2009). Uses of *le* not encoding perfectivity were excluded (47 tokens); for example, learners made extensive use of *le* as a mood marker (e.g. *tai gui le* 'too expensive!', *hao le* 'good!', *cuo le* 'wrong!', *hao duo le* 'a lot better!', *jiu hao le* 'then it's better!') and in formulaic chunks with a frozen meaning (e.g. *wei le* 'for', *chu le* 'apart from', *zenme le* 'what's wrong?').

Next, the results will be presented for each part of the study outlined above, beginning with the *le*-insertion task.

6. Results
6.1. *Le*-insertion task
6.1.1. Part one: imperfectivizing adverbials

From the results of Part One of the *le*-insertion task (and in Part Two below), we can see the difference in the rate of perfective marking between type (1) sentences, which contain no temporal adverbial, and type (2) sentences, which contain a temporal adverbial (see Sections 5.1.3-4 above).

Table 1. Learner perfective marking with and without habitual frequency adverbials

Context	Rate of *le* marking
(1) No habitual frequency adverbial	57.5%
(2) Habitual frequency adverbial	17.5%

Table 2. Control group perfective marking with and without habitual frequency adverbials

Context	Rate of *le* marking
(1) No habitual frequency adverbial	75.0%
(2) Habitual frequency adverbial	6.0%

The AH predicts no difference in the rate of perfective marking between these sentence types because the lexical aspectual class, predicted to determine learners' assignment of grammatical aspect markers, remains telic in each case. However, as Table 1 shows, learners demonstrated a strong sensitivity to the stative mapping effected by the habitual frequency adverbials *meitian* ('everyday'), *tiantian* ('everyday') and *meizhou* ('every week'). The presence of these operators strongly precluded perfective marking, despite the telicity of the verb, in line with the predictions of the SAH but contrary to the predictions of the AH.

Sentences of type (1), which are telic at the lexical level and perfective at the sentence-level (i.e. lexical aspect was not modified by a higher operator) triggered perfective marking, as predicted by both accounts: *le* was used in 57.5% of cases. However, when a habitual adverbial was added in the type (2) sentences, perfective marking fell to only 17.5%. This demonstrates the incompleteness of the AH account because it shows that learners' aspect marking is strongly influenced by operators beyond inherent lexical aspect. That is, in the type (2) sentences, where VP aspect differs from sentential aspect, learners' perfective marking reflects the final sentential aspectual class and not that of the verbal predicate. Learners show sensitivity to the aspectual mapping performed by habitual sentence adverbials, with the low rate of perfective marking reflecting the imperfectivity of the sentence despite the telicity of the verbal predicate.

As Table 2 shows, native Mandarin speakers almost did not use *le* at all (a rate of 6%) in imperfective sentences, despite the fact that these sentences contain telic verbs (e.g. resultative verb compounds). Learners' rate of perfective marking here is also low (only 17.5%), revealing an almost native-like knowledge of the impermissibility of perfective marking in imperfective sentences, despite their L1 (English) lacking grammatical perfective marking (cf. discussion in Section 3.1).

6.1.2. Part two: perfectivizing adverbials

Table 3. Learner perfective marking with and without duration adverbials

Context	Rate of *le* marking
(1) No duration adverbial	28%
(2) Duration adverbial	48%

Table 4. Control group perfective marking with and without duration adverbials

Context	Rate of *le* marking
(1) No duration adverbial	10.0%
(2) Duration adverbial	92.5%

The results for the aspectually mismatched sentences in Part Two corroborate the above findings, showing this time that learners are sensitive to the perfectivizing mapping performed by duration adverbials. In the baseline type (1) sentences (atelic at the lexical level and sententially imperfective due to the absence of higher operators), learners used the perfective marker in a minority of cases (28%). Both hypotheses predict a low rate of perfective marking in these cases. However, learners' rate of perfective marking increased to 48% in the type (2) sentences which contain a duration adverbial (i.e. atelic at the level of inherent lexical aspect but perfective at the sentence-level), showing that perfective marking is substantially influenced by sentential aspect.

Overall, learners' sensitivity to the aspectual mappings performed by temporal adverbials resemble, to a fair extent, that of native speakers. Their relatively greater use of perfective marking with duration adverbials (48%) in Part Two (compared to 28% without), irrespective of the unchanged atelic (stative) lexical aspectual class, corresponds to the even more decisive increase in the use of *le* by the control group when a duration adverbial is present (from 10% to 92.5%).

6.2. Corpus data

The results from the corpus data at Time One – namely that perfective marking occurs more with activity verbs than achievement/accomplishment verbs (a rate of 54.5% compared to 36.3% and 9% respectively) – are contrary to the predictions of the AH, which predicts the occurrence of perfective marking first on achievement and accomplishment verbs, and only afterwards on activity and stative verbs.

Table 5. Patterning of perfective marking with inherent lexical aspect

Lexical aspect	Time 1	Time 2	Total
ACH	36.3% (4)	83.3% (15)	65.5% (19)
ACC	9% (1)	0%	3.4% (1)
ACT	54.5% (6)	11.1% (2)	27.6% (8)
STA	0%	5.6% (1)	3.4% (1)
Total	11	18	29

Note: The parentheses show the token frequency. ACH = achievement; ACC = accomplishment; ACT = activity; STA = state.

Table 6. Patterning of perfective marking with sentential aspectual class

Sentential aspect	Time 1	Time 2	Total
Perfective	100% (11)	94.4% (17)	96.6% (28)
Imperfective	0%	5.6% (1)	3.4% (1)

However, these instances of perfective marking on activity verbs are predicted by the SAH account: despite their inherent lexical aspect being atelic, the derived aspectual class of these sentences is perfective because of the quantizing bound provided by duration adverbials.

The tokens of *le* on activity verbs at Time One all occurred in sentences with the following structure (either with the verb *xue* 'study' or *zuo* 'do/make'):

(17)   wo   xue   zhongwen   xue   le   **san   nian**
        I   study Chinese   study PFV   **three year**
        'I studied Chinese for three years.'

Learners used the verb copying construction, in which the verb occurs again before the duration adverbial (shown in bold). Here the inherent lexical aspect of *xue zhongwen* ('study Chinese') is atelic, and so perfective marking is not predicted by the AH. However, the bound provided by the duration adverbial directly triggers the perfective marker in these instances (cf. Li & Thompson 1989: 189) in the same way that it does in sentences without verb copying (e.g. *wo xue le san nian* 'I studied for three years'). Therefore in these instances learners' perfective marking, like that of native speakers, is directly conditioned by aspectual operators beyond lexical aspect.

Likewise, at Time Two, the patterning of perfective marking with the final (sentential) aspectual derivation rather than inherent lexical aspect in mismatched cases is also apparent. Thus, the occurrence of *le* with the stative predicate *zhu nabian* 'live there' is not predicted by the AH, but it is predicted by the SAH because the duration adverbial *shi ge yue* '10 months' provides a bound that perfectivizes the sentence:

(18)   Wo   zhu   le   nabian   shi   ge   yue[6]
        I   live PFV   there   ten   CL   month
        'I lived there for 10 months.'

Hence, across Time One and Two, the corpus data uniformly supports the results of the *le*-insertion task in showing that in mismatched cases, learners assign the perfective marker according to the final aspectual derivation of the sentence, rather than inherent lexical aspect. Overall, considering all tokens, *le* marking patterns with sentential perfectivity at a rate of 96.6%, compared to patterning with lexical telicity (achievement and accomplishment predicates) at a rate of 68.9%.[7] This supports the proposal of the SAH that – whilst inherent lexical aspect does substantially condition perfective marking – nevertheless agreement with inherent lexical aspect is subsumed within sentential agreement. That is, all cases of agreement with inherent lexical aspect also entail agreement with sentential aspect, but the converse is not true. In a significant minority of cases (7/29 tokens in this data, or 24.1%), inherent lexical aspect differs from sentential aspect because of subsequent mappings performed by temporal adverbials, and in these cases perfective marking always patterns with the final (sentential) aspectual derivation.

Additionally, certain semantically conditioned patternings are also evident. For instance, by Time Two, the perfective marker particularly patterned with a subtype of achievement verbs, resultative verb compounds (RVCs) – an association that has been found in a number of previous studies (e.g. Wen 1995, Fan 2005, Ma 2006). Every occurrence of a RVC (in which a verb or adjective follows an atelic verb to indicate the resultant state) triggered perfective marking, constituting one third (5/15) of achievement verbs marked with *le* at Time Two. Learners used a range of resultative complements to derive telic predicates from atelic

---

[6] Note that, as pointed out by a reviewer, this example from a learner contains an error. The correct sentence would be *wo zai nabian zhu le shi ge yue* (i.e. with the location adverbial *nabian* preceding the verb, and introduced by the preposition *zai* 'at').

[7] With regard to the small drop in *le* usage across Time One and Two from 100% to 94.4%, this is assumed to be statistically insignificant.

lexical verbs (e.g. *shuo-wan* 'speak-finish', *kan-dao* 'look-see', *chi-huai* 'eat-sick'), showing the productivity of such complements as perfectivizing operators (performing a process → event mapping). This supports Wen's findings (1995, 1997) which showed that the resultative verb complement *wan* 'finished' consistently triggered perfective marking even where only optional, as well as Ma's results (2006: 103), also for intermediate learners, in which RVCs substantially triggered perfective *le* (along with cardinally quantified eventualities). In the following discussion it will be suggested that the use of *le* with all RVCs (combined with the overuse found in previous studies) suggests that learners' perfective marking is conditioned by the prototypicality of the derived *telos*, which encodes a stronger notion of result/completion with RVCs than other achievement verbs.

7. Discussion

Our research questions examined evidence in two types of data (experimental and longitudinal oral corpora) to see if *le* usage by learners of Mandarin showed the patterning predicted by the AH or by the SAH, and if any changes were evident over time. The results of Part One and Two of the *le*-insertion task, in which sentences were presented containing aspectual mismatches between inherent lexical aspect and sentential aspect, show that learners were not constrained by lexical aspect alone when assigning perfective marking (as predicted by the AH), but consistently attend to the derived sentential aspectual class (following the SAH). This was shown using a range of sentence-level temporal adverbials that perform aspectual mappings in both directions: deriving imperfective sentences from telic predicates and perfective sentences from atelic predicates.

In particular, the strong aspectual effect of habitual frequency adverbials was found to condition perfective marking despite a very high degree of telicity at the predicate-level. In Part One, the telic VPs presented were composed of achievement verbs – predicted by the lexical aspect account to be the verb type that most strongly triggers perfective marking. Moreover resultative verb compounds were included in the test sentences, which as well as being achievement 'verbs' (cf. Chang 2013: 13), particularly trigger perfective marking among L2 learners (e.g. Wen 1995, 1997). However, despite the strength of verbal telicity, learners still overwhelmingly ignored inherent lexical aspect and their non-use of perfective marking was conditioned by the derived atelic sentential aspectual class.

These results (the rate of perfective marking being 40% lower with habitual adverbials and 20% higher with duration adverbials) confirm the the sentential aspect hypothesis for an L1 without verbal perfective marking, pointing towards a sensitivity to sentential aspect even in the absence of direct L1 transfer. Sharma & Deo (2009) suggested that sensitivity to sentential aspect may be constrained to learners with positive L1 transfer, concluding that verbal morphology in Indian English is sensitive to sentence-level perfectivity and imperfectivity because this is marked overtly in Hindi. However, in this study, speakers of English – which lacks verbal perfective morphology – nevertheless assigned L2 perfective markers according to sentential aspect. This sensitivity may be partly due to the interlingual equivalence of temporal adverbials, which perform the same perfective → imperfective mappings (or vice versa) in both languages, and which learners can therefore use to accurately trigger perfective marking. However, it is important to remember that although English lacks verbal perfective marking, it is not the case that English is without devices that mark boundedness. As was noted in Section 3, these are simply concentrated in the nominal rather than the verbal domain, and so learners may benefit from indirect L1 transfer when acquiring perfective marking.

Perfective marking in the corpus data also patterned strongly with specific (cardinally quantified) bounds. Overall, 38% of tokens of *le* (11/29) occurred with quantified bounds, and

90.1% (10/11) of these were specific (i.e. cardinally quantified) bounds; only one was vague (*hen duo shijian* 'a long time').

Moreover, this patterning was particularly clear at Time One (when 72.7% of perfective marking (8/11 tokens) occurred with quantified objects/duration adverbials, all but one of which was cardinally as opposed to vaguely quantified). Accordingly, a development is evident from nominal means of deriving telic eventualities at Time One to verbal means (e.g. resultative verb compounds) at Time Two.

Over one third (34.5%) of learners' total *le* usage was triggered by cardinal quantification, either of duration adverbials or objects, and this same triggering role is discernable in other Mandarin studies. Thus Duff & Li (2002: 443) found that cardinally quantified bounds were the dominating context for accurate suppliance of *le* in their written editing task, whilst in Ma's (2006: 104) production data, cardinal quantification of frequency adverbials was one of the main triggers of perfective *le* from intermediate level. As noted above, cardinal quantification triggers perfective marking because English does not mark the perfective aspect verbally, so boundedness is often derived nominally and learners initially transfer their L1 technique for deriving telic VPs to the L2 (cf. Slabakova 1999). That is, in English most activity/accomplishment predicates are made telic through quantization of the object, by which the eventuality is delimited (Slabakoba 2005: 64). Thus it is unsurprising that learners make extensive use of this means of telicizing predicates in the L2.

The other notable trigger of perfective marking, resultative verb compounds, can be explained using a prototype-based model of telicity (cf. Andersen & Shirai 1994: 146) in terms of their greater prototypicality compared with other telic verbs. As well as simply containing a temporal endpoint, RVCs entail the successful completion of the activity, such that a subsequent (resultative) state/phase is initiated – indicated by the resultative complement (e.g. *zuo wan* 'do-finish') (cf. Xiao & McEnery 2004: 61).

8. Conclusion

In conclusion, this study has demonstrated the incompleteness of the AH account in that learners' perfective marking patterns with the final derived aspectual class of the sentence and not inherent lexical aspect, where mismatches occur through the aspectual shifts caused by temporal adverbials. Moreover, the experimental and corpus data converge to show that from the early stages of the emergence of perfective *le*, its use or non-use is directly determined by duration and habitual frequency adverbials, which override inherent lexical aspect and yield perfective and imperfective sentences respectively. These results support the SAH's claim that learners' aspect marking agrees with the final sentential aspectual derivation rather than exclusively with lexical aspect (which underdetermines the eventuality's aspectual class).

Additionally, the patterning of interlanguage perfective marking was subject to L1 influence, a possibility excluded from the AH's universal template of morphology to verb type mappings. Transfer from English was identified in learners' progression from, at Time One in the corpus data, an extensive reliance on object-marking to derive telic eventualities from activity verbs – a strategy used in their L1 – to the L2 strategy of using verb-marking at Time Two (supporting similar findings by Slabakova 1999, 2005).

Here, learners who did not directly benefit from positive transfer (i.e. their L1, English, does not contain grammatical perfective marking in the verbal domain) nevertheless adopted the sentential aspectual class as the determining factor in their assignment and non-assignment of the L2 perfective marker. This reliance on sentential aspect is likely to be substantially due to cognitive and linguistic universals, including the cross-linguistic (semantic) equivalence of temporal adverbials. It remains to be seen, though, how far the manipulation of temporal adverbials as a bootstrapping device from the early stages of the acquisition process may be a

learning strategy characteristic of L2 rather than L1 acquisition, as well as how far devices beyond the verbal predicate condition L2 aspectual marking for other aspect markers across other language pairings.

Appendix A: Operational Tests
(quoted from Chen and Shirai 2010)

Predicates are indicated in bold italics.

Step 1: State or nonstate (nondynamic vs. dynamic)

The verb (or verb phrase) cannot have a habitual interpretation without any aspect marker attached, can it?
If it cannot → state (e.g., *Wo **ai** ni* 'I love you'→ no habitual reading)
If it can → nonstate (e.g., *Wo tiantian **chi** mifan* 'I every day eat rice' (I eat rice every day) → habitual reading possible) → Go to Step 2

Step 2: Punctual or durative

[If test (a) does not apply, apply test (b)]
(a) Can you say 'X kaishi VP' (= 'X begin to VP') without an iterative interpretation?
— If you cannot → Achievement (e.g. #*Ta kaishi **si*** 'he begins to die') → Go to step 4.
— If you can → Accomplishment (e.g. *Ta kaishi **xie yi feng xin*** 'he begins to write a letter') or Activity (e.g., *Ta kaishi **paobu*** 'he begins to run') → Go to Step 3
(b) Can you say 'X will VP at Y o'clock (e.g. 2 o'clock) sharp'?
— If you can → Achievement (e.g. *Huiyi hui zai 2 dian zheng **kaishi*** 'Lit: meeting will at 2 o'clock sharp begin' (The meeting will begin at 2 o'clock sharp) → Go to step 4
If you cannot → Accomplishment or Activity → Go to Step 3

Step 3: Accomplishment or Activity/semelfactive (Telic vs. atelic)

[If test (a) does not apply, apply test (b)]
(a) Can 'X *chadianr* VP *le*' (= 'X almost VP le') mean 'X started V but did not complete it'?
— If it can → Accomplishment (e.g. *Ta chadianr **pao dao xuexiao** le* 'Lit: he almost run arrive school *le*' (he almost ran to the school) can mean that he started running but he didn't reach the school).
— If it cannot → Activity or semelfactive (e.g. *Ta chadianr **pao le bu*** 'he almost ran *le*' (he almost ran) can only be interpreted as he almost started running) → Go to Step 4.
 (b) Can you say 'X will VP for Y time' (e.g., 10 min)?
— If you can → Activity (e.g., *Ta hui zuo 10 fenzhong* 'he will sit for 10 minutes') or semelfactive (*Ta kesou le 10 fenzhong* 'he coughed for 10 minutes').
— If you cannot → Accomplishment (e.g. #*Ta pao dao xuexiao 10 fenzhong* 'he run arrive school 10 minutes' (#He ran to school for 10 minutes) → Go to Step 4.

Step 4: Achievement or Semelfactive

Can you say 'X zai VP' with iterative/repetitive (i.e. iteration on one occasion. Not habitual) interpretation?
— If you can → Semelfactive (e.g. *Ta zai kesou* 'he *zai* cough' [he is coughing])
If you cannot → Achievement (e.g. #*Ta zai si* 'he *zai* die')

Appendix B: Test Sentences

Habitual frequency adverbials

(1)   Shang   ge   yue   wo   dapo   yi   ge   beizi
        last   CL   month   I   broke   one CL   glass
        'Last month I broke a glass.'

(2)   Shang   ge   yue   wo   meitian   dapo   yi   ge   beizi
        last   CL   month   I   everyday   broke   one CL   glass
        'Last month I broke a glass everyday.'

(1)   Shang   ge   xingqi wo   dao   chaoshi
        last   CL   week   I   went   supermarket
        'I went to the supermarket last week'

(2)   Tiantian   wo   dou   dao   chaoshi
        everyday   I   always   went   supermarket
        'I went to the supermarket everyday.'

(1)   Ta   dasi   yi   tiao   chongzi
        he   kill   one   CL   insect
        'He killed an insect.'

(2)   Ta   meizhou   dasi yi   tiao   chongzi
        he   every week   kill one   CL   insect
        'Every week he killed an insect.'

(1)   Shang   ge   xingqi wo   diu   wode   qianbao
        last   CL   week   I   lost   my   wallet
        'Last week I lost my wallet.'

(2)   Shang   ge   yue   wode   haizi   meizhou   diu   tade   wanju
        last   CL   month   my   child   every week   lost his   toy
        'Last month my child lost his toy every week'

Duration adverbials

(1)   Ta   zai   nar   zuo
        he   at   there   sit
        'He sat there.'

(2)   Ta   zai   nar   zuo   ji   ge   xiaoshi
        he   at   there   sit   several   CL   hour
        'He sat there for several hours.'

(1)   Shangci   ta   zhan   zai   nar
        last time   he   stand   at   there
        'Last time he stood there.'

(2)   Shangci   ta   zai   nar   zhan   ji   ge   xiaoshi
        last time he   at   there   stood   several   CL   hour
        'Last time he stood there for several hours.'

(1)   Qunian   ta   zai   nali   zhu
        last year he   at   there   live
        'Last year he lived there.'

(2)   Qunian   ta   zai   nali   zhu   liang   ge   yue
        last year he   at   there   live   two   CL   month
        'Last year he lived there for 2 months'

(1)   Ta   shangci   deng   ni
        he   last time   wait   you
        'Last time he waited for you.'

(2)   Shangci   ta   deng   ni   yi   huir
        last time he   wait   you   a   while
        'Last time he waited for you for a while.'

(1)   Ta   zai   chuang   shang   tang
        he   at   bed   on   lie
        'He lay on the bed.'

(2)   Ta   zai   chuang   shang   tang   yi   huir
        he   at   bed   on   lie   a   while
        'He lay on the bed for a while.'

References

Andersen, R. & Shirai, Y. (1994). Discourse motivations for some cognitive acquisition principles. *Studies in Second Language Acquisition* 16, 133-156.

Andersen, R. & Shirai, Y. (1996). Primacy of aspect in first and second language acquisition: the pidgin/creole connection. In Ritchie, W. C. & Bhatia, T. K. (eds.), *Handbook of Second Language Acquisition*, 527-570. San Diego, CA: Academic Press.

Baker, J. & Quesada, M. (2009). The effect of temporal adverbials in the selection of preterit and imperfect by learners of Spanish L2. Paper presented at *Second Language Research Forum*, Michigan State University, Lansing, Michigan.

Bardovi-Harlig, K. (1998). Narrative structure and lexical aspect: conspiring factors in second language acquisition of tense-aspect morphology. *Studies in Second Language Acquisition* 20, 471-508.

Cadierno, T. (2000). The acquisition of Spanish grammatical aspect by Danish advanced language learners. *Spanish Applied Linguistics* 4, 1-53.

Chang, H-H. (2013). The aspect marker -le in early child Mandarin. *Journal of Universal Language* 14 (1), 7-54.

Chen, J. & Shirai, Y. (2010). The development of aspectual marking in child Mandarin Chinese. *Applied Psycholinguistics* 31, 1-28.

Comrie, B. (1976). *Aspect*. Cambridge: Cambridge University Press.

Croft, W. (2009). Aspectual and causal structure in event representations. In Gathercole, V. (ed.), *Routes to Language Development: In Honor of Melissa Bowerman*, 139–66.

Mahwah, NJ: Lawrence Erlbaum Associates.

de Swart, H. (1998). Aspect shift and coercion. *Natural Language and Linguistic Theory* 16, 347–385.

de Swart, H. (2000). Tense, aspect and coercion in a cross-linguistic perspective. In Butt, M. & King, T. H. (eds.), *Proceedings of the Berkeley Formal Grammar conference*. University of California, Berkeley: Center for the Study of Language and Information (CSLI) Publications.

Dietrich, R., Wolfgang K., Colette N., & Josée C. (1995). *The acquisition of temporality in a second language*. Amsterdam: John Benjamins.

Dowty, D. (1986). The effect of aspectual class on the temporal structure of discourse: semantics or pragmatics? *Linguistics and Philosophy* 9, 37-61.

Dowty, D. (1991). Thematic proto-roles and argument selection, *Language* 67, 547-619.

Duff, P., & Duanduan L. (2002). The acquisition and use of perfective aspect in Mandarin. In Salaberry, R. & Shirai Y. (eds.), *The L2 Acquisition of Tense-aspect Morphology*, 417-453. Amsterdam: John Benjamins.

Fan, H. (2005). *Acquisition of tense-aspect morphology by English learners of French and Chinese*. Ph.D dissertation, University of Florida.

Gabriele, A. (2009). Transfer and transition in the SLA of aspect. *Studies in Second Language Acquisition* 31, 371-402.

Garey, H. (1957). Verbal aspects in French. *Language* 33, 91-110.

Jackendoff, R. (1996). The proper treatment of measuring out, telicity, and perhaps even quantification in English. *Natural Language and Linguistic Theory* 14, 305-354.

Jin, L. & Hendriks, H. (2005). The development of aspect marking in L1 and L2 Chinese. *Working Papers in English and Applied Linguistics* 9, 69–99.

Kabakciev, K. (2000). *Aspect in English: a 'common-sense' view of the interplay between verbal and nominal referents*. Dordrecht: Kluwer.

Krifka, M. (1989). Nominal reference, temporal constitution and quantification in event semantics. *Semantics and Contextual Expression 75*, 115.

Krifka, M. (1998). The origins of telicity. In Rothstein, S. (ed.), *Events and Grammar*, 197-236. Dordrecht: Kluwer.

Leiss, E. (2007). Covert patterns of definiteness/indefiniteness and aspectuality in Old Icelandic, Gothic, and Old High German. In Stark, E., Leiss, E. & Werner, A. (eds.), *Nominal Determination Typology, Context Constraints, and Historical Emergence*. Amsterdam: John Benjamins.

Laleko, O. (2008). Compositional telicity and heritage Russian aspect. In Grosvald, M. & Soares, D. (eds.), *Proceedings of the Thirty-Eighth Western Conference on Linguistics (WECOL)* 19, 150-160.

Li, C. & Thompson, S. (1989). *Mandarin Chinese: a functional reference grammar*. Berkeley: University of California Press.

Li, P. & Shirai, Y. (2000). *The acquisition of lexical and grammatical aspect*. Berlin: Mouton de Gruyter.

Ma, L. (2006). *Acquisition of the perfective aspect marker le of Mandarin Chinese in discourse by American college learners*. Ph.D dissertation, University of Iowa.

Moens, M. (1987). *Tense, aspect and temporal reference*. Ph.D dissertation, Centre for Cognitive Science, University of Edinburgh.

Musumeci, D. (1989). *The ability of second language learners to assign tense at the sentence level: a crosslinguistic study*. Ph.D dissertation, University of Illinois.

Mourelatos, A. (1978). Events, processes and states. *Linguistics and Philosophy* 1, 199-220.

Noyau, C. (2002). Temporal relations in learner varieties: grammaticalization and discourse construction. In Salaberry, R. & Shirai, Y. (eds.), *The L2 Acquisition of Tense-aspect Morphology*, 107-128. Philadelphia: Benjamins.

Robison, R. (1990). The primacy of aspect: aspectual marking in English interlanguage. *Studies in Second Language Acquisition* 12, 315-330.

Rocca, S. (2002). Lexical aspect in child second language acquisition of temporal morphology. In Salaberry, R. & Shirai, Y. (eds.), *The L2 acquisition of tense-aspect morphology.* Amsterdam: John Benjamins.

Rohde, A. (1996). The aspect hypothesis and the emergence of tense distinctions in naturalistic L2 acquisition. *Linguistics* 34, 1115-1137.

Salaberry, R. (1999). The development of past tense verbal morphology in classroom L2 Spanish. *Applied Linguistics* 20 (2), 151-178.

Salaberry, R. (2013). Research design: operationalizing and testing hypotheses. In Salaberry, R. & Comajoan, L. (eds.), *Research Design and Methodology in Studies on L2 Tense and Aspect*. 187-218, Berlin: Walter de Gruyter.

Sharma, D. & Deo, A. (2009). Contact-based aspectual restructuring: a critique of the lexical aspect hypothesis. *Queen Mary's Occasional Papers Advancing Linguistics* 12, 1-29.

Shibata, M. (1999). The use of Japanese tense-aspect morphology in L2 discourse narratives. *Acquisition of Japanese as a Second Language* 2, 68-102.

Shirai, Y. (1995). Tense-aspect marking by L2 learners of Japanese. In MacLaughlin, D. & McEwen, S. (eds.), *Proceedings of the 19th Annual Boston University Conference on Language Development* 2, 575-586. Somerville, MA: Cascadilla Press.

Shirai, Y. (2013). Defining and coding data: lexical aspect in L2 studies. In Salaberry, R. & Comajoan, L. (eds.), *Research Design and Methodology in Studies on L2 Tense and Aspect*, 271-308. Berlin: Walter de Gruyter.

Slabakova, R. (1999). The parameter of aspect in second language acquisition. *Second Language Research* 15 (3), 283-317.

Slabakova, R. (2005). What is so difficult about telicity marking in L2 Russian? *Bilingualism: Language and Cognition* 8 (1), 63-77.

Slabakova, R. & Montrul, S. (2008). Aspectual shifts: grammatical and pragmatic knowledge in L2 acquisition. In Liceras, J., Zobl, H. & Goodluck, H. (eds.), *The Role of Formal Features in Second Language Acquisition*, 455-83. New York: Lawrence Erlbaum Associates.

Smith, C. (1997). *The parameter of aspect*. Dordrecht: Kluwer.

Starren, M. (2006). Temporal adverbials and early tense and aspect markers in the acquisition of Dutch. In Van Geenhoven, V. (ed.), *Semantics in Acquisition*, 219-244. Dordrecht: Springer.

Vendler, Z. (1967). *Linguistics in philosophy*. Ithaca, New York: Cornell University Press.

Verkuyl, H. (1972). *On the compositional nature of the aspects*. Dordrecht: D. Reidel Publishing Company.

Verkuyl, H. (2005). Aspectual composition: surveying the ingredients. In Verkuyl, H. J., de Swart, H. & van Hout, A. (eds.), 19-39. *Perspectives on aspect*. Dordrecht: Springer.

Wang, L. (2012). *Second language acquisition of Mandarin aspect markers by native Swedish adults*. Ph.D dissertation, Uppsala University.

Wen, X. (1995). Second language acquisition of the Chinese particle le. *International Journal of Applied Linguistics* 5, 45-62.

Wen, X. (1997). Acquisition of Chinese aspect: an analysis of the interlanguage of learners of Chinese as a second language. *ITL: Review of Applied Linguistics* 117/118, 1-26.

Xiao, R. & McEnery, T. (2004). *Aspect in Mandarin Chinese: a corpus-based study*. Amsterdam and Philadelphia: John Benjamins Publishing Company.

Daniel Bell
School of English Literature, Language and Linguistics
Newcastle University
Newcastle upon Tyne
Tyne and Wear, NE1 7RU
United Kingdom

d.m.bell4@ncl.ac.uk
dannn.bell@gmail.com

Clare Wright
Centre for Literacy and Multilingualism / Department of English Language and Applied Linguistics
University of Reading
Whiteknights
Reading, RG6 6AA
United Kingdom

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

CROSS-LINGUISTIC DIFFERENCES BETWEEN ENGLISH AND JAPANESE IN MONOLOGUES: IMPLICATIONS FROM A MICRO-LEVEL BIDIRECTIONAL ANALYSIS OF L1-L2 DESCRIPTIONS IN SPEECH AND GESTURE-IN-TALK[*]

KAZUKI HATA
(Newcastle University)

Abstract

The present study conducts a cross-linguistic bidirectional analysis of gesture-in-talk with English and Japanese speech as both L1 and L2. The discussion is based on a micro-level qualitative analysis without recourse to statistical perspectives, which is rare in recent bidirectional cross-linguistic gesture studies. In this study, 24 speakers participated in a story-narrative task: 6 monolinguals as L1 speakers of each language and 12 bilinguals. As consistently reported in recent work on gesture studies, the results highlight the high degree of semantic synchrony of gestures with accompanying speech units that encode the same type of information simultaneously. These patterns were recognised in all analysed language groups, regardless of the language. The data considered here, however, also demonstrates that the speakers possess different semantic conceptions regarding the meaning of the motion verb *climb* and its Japanese equivalent *noboru* 'climb (up)', making it difficult to conduct an equivalent comparative analysis of speech production in L1 and L2. Thus although the present case study supports the recent findings in the literature regarding the semantic synchrony and uniformity of gestural shapes, it also reveals that the effect of inequivalent semantic variation should be taken into account during cross-linguistic investigation of the speech-gesture relationship in speakers' descriptions of motion events.

1. Introduction

The present study investigates the use of speech and the use of iconic gestures (henceforth, iconics; McNeill 1992) in two target languages, English and Japanese, occurring in story-narrative tasks. The primary research question is whether idiosyncratic variation and preferences affect the use of verbal and gestural expressions by monolingual and bilingual speakers in each target language, and if so, to what extent. Here, it is worth emphasising that the present study considers the expression of manner or path information in speech and gesture, in connection with the effect of semantic variation of specific word forms, via a micro-level analysis with no recourse to statistical perspectives. This micro-level analytic investigation suggests that it is arguably necessary to focus on the semantic concepts in speakers' minds as some words in one language represent semantic information in a different way from their translation equivalents in other languages (Brown & Gullberg 2011: 91). The study design is fully experimental, video-recorded and task-oriented, and therefore, results in the present study may not be comparable to those from natural speech contexts (see Garfinkel 1996: 11, ten Have 1999: 48-49).

2. Background

The term 'gesture-in-talk' (i.e. 'gesticulation'; henceforth, 'gesture') is defined as embodied movements with communicative roles, which are systematically intertwined with spoken words by encoding the semantically and/or pragmatically parallel message (Kendon 1980, 1996, 2004, McNeill 1992). This suggests that gestural expressions are neither just co-productions of speech nor simple translations of utterances (McNeill 1992: 105). They can be seen as tightly synchronised symbolic signals co-expressed with speech to relay messages that are not ready-made for spoken forms of language (McNeill 1985: 360, Clark & Krych 2004: 78, Holler and Beattie 2004: 67, Wilcox 2004: 525). Maintaining this, the major difference is that utterances in spoken language generate a meaning by linguistic items as a 'constructed unit', but gesture depicts a complete picture without a strictly obligatory construction process, thus working as a 'presentational unit' (McNeill 1992: 21, Kendon 1994: 49).

Recent findings about gestural functions in L1 and L2 shed light on cross-linguistic variation regarding encoded messages in speech and gesture and patterns of use. In studies based on story-narrative tasks, a key consideration is often the speaker's pattern of encoding information about manner and path in speech and gesture to describe the target motion events in the stimulus. According to Talmy (1985), expressions of manner vary among different languages depending on how path information for motions is encoded. Consider the following examples (1)-(2):

(1) a. I twisted the cork into the bottle.
     b. I twisted the cork out of the bottle.   (Talmy 1985: 72)

(2)   *booru-ga saka-wo korogatte iku*
         ball   hill   roll   go
         'The ball goes rolling on the hill'   (Brown & Gullberg 2008: 228)

English is classified as a satellite-framed language, which usually encodes trajectory in a combination of satellites, that is, 'certain immediate constituents of a verb root' (Talmy 1985: 102, 1991). Thus, the main verb might not be affected by the expression of path information, as in (1a) and (1b). On the other hand, verb-framed languages like Japanese express path information in the main verb, as in (2). In the latter case, manner information could be expressed in either a subordinated verb or adverbial, or it is omitted altogether (Talmy 1985, Slobin 2004a, 2004b).

Although the literature shows cross-linguistic variation between satellite-framed and verb-framed languages, one source reports that iconics, a type of representational gesture performed in talk, demonstrate a high degree of uniformity with regard to gestural shapes, despite considerable differences between spoken languages (McNeill 1992: 221-222, see the following section for the semiotic features of iconics). This claim can be explained by the Growth Point theory (GP) following Vygotsky (1986), which holds that speakers' minimal idea units can develop from an image into 'a complex structure in which both the gestures and the linguistic structure are integral parts' (McNeill 1992: 29-30). If the GP is a starting point of thought coming from a presupposed background, it should only exhibit slight variation from language to language. However, if different languages encourage different ways of thinking (Slobin 1991: 23), there must be room to suggest that gesture will also vary across languages. In fact, later studies report cross-linguistic variation regarding the use of iconics across different languages (Kita 2000, McNeill & Duncan 2000). Specifically, Kita and Özyürek (2003) found that gestures depicted spatial details that might not be expressed in the accompanying speech of some languages due to inequivalent lexical entries to describe the specific motion event. For

example, there is a lack of linguistic resources to express the swing-across motion in Japanese, which leads to difficulty in articulating arc movement. English uses one clause to express manner and path, whereas Japanese and Turkish tend to use two or more. Given this, Kita and Özyürek propose 'the Interface Hypothesis', based on GP theory, claiming that gestures arise from an interface representation between speaking and spatial thinking, which is information tailored for speech, that is, thinking-for-speaking (McNeill & Duncan 2000). That is to say, gestures are considered to be formed in a different way across languages similarly to the different patterns in linguistic representation, and therefore, linguistic differences influence how spatial information arises in the process of speech production and is expressed by gesture in story narratives.

However, bidirectional cross-linguistic differences between L1 and L2 speakers are arguably still unclear, as there have been relatively few attempts to collect data on Japanese as L2, compared to an accumulation of English samples (see McCafferty 1998, Kita 2000, Kita & Özyürek 2003, Brown & Gullberg 2008, 2011). Given this, there is a research gap in the study of cross-linguistic variation in the expression of manner in speech and gesture between English and Japanese. Consequently, the present study conducts a bidirectional analysis involving the two target languages: both L1 and L2 use of English and Japanese. Considering that gestures might demonstrate idiosyncratic (i.e. inter- or intra-speaker) variation between speakers and/or a single speaker (Kendon 1996, Kipp, Neff & Albrecht 2007), statistical comparison is not a valid and/or reliable procedure. Thus, this study undertakes a micro-level qualitative analysis with no recourse to statistical perspectives in order to investigate the expression of manner and path information in speech and gesture.

3. Methodology
3.1. Participants

The present study recruited 24 people aged 22 to 35 at the time of recording: 12 British English L1 speakers and 12 Japanese L1 speakers. These speakers fell into four different sub-groups depending on their language background: English monolinguals (EM), English bilinguals (EB), Japanese monolinguals (JM) and Japanese bilinguals (JB). In order to standardise their background, the study defines monolingual speakers as people who were born and raised in their home country, and have not had long-term experiences of studying/working abroad and are not acclimatised to other cultures as well as or more than their first culture. This requirement was set to minimise the effect of the speaker's culture on gesture-in-use with the diversity and complexity of meaning (see Birdwhistell 1970, Morris 1979, Richmond & McCroskey 1999). Bilingual speakers in this study, on the other hand, are simply defined as people who use or had used the L2 in their daily activities.

Regarding the categorisation of the participants as monolingual or bilingual, their proficiency level and language experience in L2 were considered, supported by their score in official language proficiency tests for L2 (i.e. IELTS and JLPT; see below) and their experience of studying abroad; all the bilinguals had experience studying or working abroad. The general information on language usage and background, collected by using a pre-test questionnaire, which was based on the language background questionnaire developed by the Max Planck Institute (Gullberg & Indefrey 2003), is summarised in Table 1.

Table 1. Summary of participants, based on the language background questionnaire

Language background	English monolingual (n=6)	English bilingual (n=6)	Japanese monolingual (n=6)	Japanese bilingual (n=6)
Mean age	36	25	37	27
Mean JLPT / IELTS score	-	N2 (JLPT)	-	6.5 (IELTS)
Mean AoE to L2	0 yrs	5.3 yrs	8 yrs	12.8 yrs
Mean usage of L2 (per day)	-	2.1 hrs	-	5.1 hrs
Mean self-rating in L2 use	-	4.1	-	3.6
Study abroad experience	-	1-3 yrs	-	1-2 yrs

Key: AoE = Age of Exposure; IELTS = International English Language Testing System (Score range: 0.0-9.0 as highest); JLPT = Japanese-Language Proficiency Test (Grade range: N5-N1 as highest)

3.2. Stimulus

The study used an animated colour cartoon *Canary Row* (Freleng 1950) as the stimulus due to its potential not only to evoke a variety of gestures (McNeill & Levy 1982) but also to be used with a variety of speakers, including L2 learners (McNeill 1992: 365). For research purposes, the focus of this study lies on the participants' descriptions of the two climb-up motions shown in the stimulus, as both languages potentially have enough lexical entries to describe these motion events (see Kita & Özyürek 2003 for descriptions of 'swing across' in the same video stimulus). The study used an edited version of the cartoon, which was shortened to approximately two minutes and included the two target motion events, in order to allow it to be easily retold within a single attempt.

3.3. Procedures

Each participant was involved in the experiment for one day in the case of monolinguals or two days in the case of bilinguals; all the bilingual speakers were gathered again and saw the same stimulus for the same task in a different language. In order to counterbalance their L1 and L2 performances, half of the group of participants were randomly required to tell the story in the L1-L2 order, and the other half did so in the L2-L1 order. This procedure was necessary to control the unwelcome effect stemming from their background information on either L1 or L2 descriptions (Gullberg 2003). A second meeting took place one week following the first recordings.

Prior to the recordings, each participant met with one of the participating researchers for a short informal conversation about general topics for two or three minutes. This served as a warm-up task in the language to be used, with two aims: (a) to relax participants and (b) to put them into the 'monolingual mode' (Grosjean 1998, Brown & Gullberg 2008). Following the warm-up section, each participant was given instructions. They were told that they were participating in a story-retelling task and informed that their retellings would be videotaped, and the tapes would be shown to other participants who do not know the story. The latter is a pseudo-instruction, strategically placed to enforce participant output accuracy during production. At that point, no specific information about the stimulus was given to the participants except its length. During the activities, participants were strongly advised to speak not to the video camera but to the researcher in front of them, as some studies showed that

absence of listeners might lead to a reduction in the rate of gesture frequency (Cohen & Harrison 1973, Bavelas, Chovil, Lawrie & Wade 1992, Alibali et al. 2001).

After the participants were given sufficient instruction, each of them signed the written consent form. The focus of this study was not mentioned to them at any point prior to execution of the task, in order to record natural and spontaneous behaviour and to avoid participant attention being directed at the focus (McNeill 1992: 374). After the completion of the experiment, they were informed about the true aim of the study, and were again asked for permission to use their data after the recordings. A post-hoc interview took place immediately following the task. This was recorded and it was confirmed that no participants identified the focus of this study while executing the task.

After the recording phase, a follow-up questionnaire with a semantic judgement task was carried out in order to check their semantic conceptualization of the English word *climb* and its Japanese equivalent *noboru* 'climb (up)'. As initially suggested in Brown and Gullberg (2011: 91), it was uncertain whether Japanese speakers used these words as manner or path verbs. This post-recording interview clarified that bilinguals were seen to have different semantic conceptions of the verb *climb* and its Japanese translation compared to monolinguals.

3.4. Annotation process

The present study identified that all gestures which accompanied speech referring to the climb-up motion event are iconics. Iconics, a semiotic type of gesture-in-talk, are seen as more semantically tied to the accompanying speech than other types of gestures, showing a close synchronicity between speech and gesture (McNeill 1992, see also Beattie & Shovelton 2002, Holler & Beattie 2002, 2004). This type of gesture represents the same feature of the action or event as its verbal description and is thus semantically parallel to the verbal expression (McNeill 1985, 1992). Iconics are also used to encode manner or path information which is not directly expressed in accompanying speech (Brown & Gullberg 2008). As they encode manner or path information, iconics can represent the viewpoint of the narrator: the character's viewpoint (C-VPT) and the observer's viewpoint (O-VPT) (McNeill 1992: 123-133). For example, a speaker depicts the event from the C-VPT when moving his/her arms up and down, demonstrating the character's motion going up the ladder from the character's viewpoint. Alternatively, he/she can refer to the same event from the O-VPT when raising up his/her hand, depicting the path information of the character rising upward (McNeill 1992: 118-119).

The study used transcriptions and annotations of video data created through ELAN annotation software, a free multi-modal annotation tool for audio and video data originally developed at the Max Plank Institute for Psycholinguistics (Wittenburg et al. 2006). The software was utilised for spatio-temporal synchronisation between the different types of data with separate annotation tiers for two different types of communicational signals, speech and gesture. All spoken data were represented in orthographic transcription and synchronised with the gesture annotations and the corresponding audio and video data in the software. As the study investigates non-linguistic communicative signals (i.e. gesture) and non-English language in use, transcription and its translation, as well as gesture annotations, were double-checked by three annotators through discussion, in order to establish 'inter-rater reliability'; this is arguably a mandatory process to avoid too much selectivity arising from a single annotator in any transcription/annotation processes (McNeill 1992: 375, Chawla & Krauss 1994: 588, Kendon 1996: 5). The study recruited three Japanese speaking L2 English speakers to translate Japanese utterances into their English equivalents, and an English-speaking annotator with experience in annotating gestures. In each case, a consensus was reached through discussions among the annotators. The list of motion verbs uttered by different groups of participants is shown in Table 2.

Within the annotation process, gestures were annotated utilising McNeill's semiotic gesture categorisation. The present study analysed only the stroke phase of each gesture, as this segment is considered obligatory and thus the most significant part of a gestural movement (McNeill 1992, Kita, van Gijn & van der Hulst 1998, Kita 2000). Each stroke phase was identified through annotation of the gesture's internal temporal structure, which identifies the start and end point of the stroke or hold phase (Kipp et al. 2007: 330). In the present study, each gesture annotation is described with co-expressed accompanying speech by stating the stroke phase in boldface and bracketing the entire gesture phrase, followed by the description of the actual movement, as in example (3).

(3)   The cat is [**climbing** on the wall]
        Manner: wiggling movement of hands depicting the manner information for the climb-up motion of the character

Table 2. List of motion verbs linguistically expressed in monologues

	English monolingual	English bilingual	Japanese monolingual	Japanese bilingual
English	climb (+up) crawl get (+into / up) go (+up / through)	climb (+up) get (+into/up)		climb (+up) approach (+to) enter go (+up)
Japanese	*noboru* (climb) *agaru* (go up) *iku* (go)	*noboru* (climb up) *iku* (go) *chikazuku* (approach) *haiagaru* (climb up)	*noboru* (climb up) *kuru* (come) *iku* (go) *chikazuku* (approach) *haiagaru* (climb up)

Key: 1. (+up) = the word 'climb' was in some cases followed by the adverbial 'up', but not always.
2. Semantically, '*noboru*', '*agaru*' and '*haiagaru*' are always only associated with the description of upwards motion; however, it should be noted that this semantic value was violated by English-Japanese bilinguals.

4. Analysis of speech-gesture synchrony

Remarkably, all the manner-encoding gestures were strictly tied with the motion verbs referring to the climb-up event, regardless of the speaker's language background. The Japanese speaking participants who used manner-encoding gestures always synchronised the stroke phase with the motion verb rather than adverbials or noun phrases, thus directly expressing the climb-up motion. Consider the following examples (4)-(6).

(4)   *isshou-kenmei [**nobotte   iku**]*
        with his best   climb up   go
        '(he) goes **climbing up** with his best'
        Manner: wiggling movement of hands depicting the manner information for the climb-up motion of the character

(5)   *haisuikan-wo   [**nobori-masu**]*
        the drainpipe   climb up
        '(he) **climbs up** the drainpipe'
        Manner: wiggling movement of hands depicting the manner information for the climb-up motion of the character

(6)   he tries to [**climb up** on the wall]
        Manner: wiggling movement of hands depicting the manner information for the climb-up motion of the character

        Example (4) shows that the stroke of the manner-encoding gesture starts by saying *nobotte* 'climbing (up)'. The sentence pattern seen in (4), packaging a manner verb and a path-verb within the same unit (Ohara 2002, Sugiyama 2005), is called a quasi-lexical unit (Slobin 2006: 69). Interestingly, this synchrony is also seen in descriptions by the English bilinguals in their L2 narratives, such as in (5); although a quasi-lexical unit was not seen in this example. There is inter-speaker variation in the length of the stroke; some of the speakers hold their hand movement right after *climb*, but the others continue their moving gestures until they complete the sentence. Nevertheless, the starting point of the stroke is similar and always tied with accompanying spoken words encoding the same type of information (i.e. manner). Interestingly, this pattern is found not only in the Japanese-speaking narratives but also in the English-speaking attempts, as in (6). Whatever the circumstances, the speakers always executed the stroke phase when uttering the main verbs as a reference to the climb-up motion, when they encoded the manner information in gestures.
        Significantly, such a strict speech-gesture synchrony, regardless of language selection, is also seen in path-encoding gestures. For all speakers, the stroke of path-encoding gestures is tied to the adverbials but not to the main verbs, regardless of the language spoken; compare (7)-(9).

(7)   *amadoi-wo   [**totte   chikazuku**]*
        the drainpipe   through   come close
        '(he) comes close (by going) **through** the drainpipe'
        Path: upward movement of one hand depicting the trajectory information of the climb-up motion of the character

(8)   *nakanihai-tte [**ue-made   itte**]*
        going inside   into up   go
        '(he) gets into the inside (of the pipe) then goes **up**'
        Path: upward movement of one hand depicting the trajectory information of the climb-up motion of the character

(9)   he climbs up [**through** the drainpipe]
        Path: a hand movement rising up depicting the trajectory information of the climb-up motion of the character

The results show that the stroke of the path-encoding gestures is synchronised with an expression of path information in speech. Some speakers hold gestural shapes until they complete the sentence in speech, but others simply tie path-encoding gestures to the accompanying adverbials.

Note that these findings do not generate new perspectives, but rather support the claim of speech-gesture synchrony in the literature (e.g. McNeill 1985, 1992, Kendon 1997, Gullberg 2003), by claiming that gestures tend to be performed at the same time as semantically parallel linguistic items. Thus, gestures co-expressed with speech typically encode the same information derived from the same 'idea unit'; the meanings conveyed by the speech and the gesture complement each other in depicting this identical information (McNeill 1992: 27). In fact, the data in this study show that each dynamic stroke with a core message is seen as semantically synchronised with the content of the accompanying speech unit. The data also demonstrate that the stroke never crosses clause boundaries in speech although gestures are held until the sentence is completed, similarly to McNeill's discussion of the optional hold phase (1985: 361).

Another notable point is that semantic synchrony is seen in all samples regardless of the speaker's language background and the language to be used in the narratives, which casts doubt on the idea of cross-linguistic variation in gesture-in-use. As McNeill claims, the synchrony rule will not always be clear since some iconics are so vague that it is difficult to show that the gestures clearly encode the same idea unit as the speech does (1992: 27). Nevertheless, the study provides evidence of the semantic synchrony between gestures and accompanying speech units, which holds for all the speakers recruited in the present study. Participants in the present study demonstrated uniformity of gestural shape usage. Regarding this, the inter-speaker (or intra-speaker) variation of gestures is arguably still in debate; it has been on the one hand reported that there is consistency of gestural shapes between different language speakers unless they describe the same event (McNeill 1992), but on the other hand it is claimed that there might be significant variation (Kendon 1996, Kipp et al. 2007). Maintaining that different language speakers would demonstrate different ways of thinking (Slobin 1991), it is expected that gestural expressions have significant inter-speaker variation. On the other hand, this paper will demonstrate that all the speakers, regardless of their language background, exercised manner- or path-encoding gestures with similar shapes to describe the same motion event.

There are two patterns of gestures used to depict either the manner or path information of the climb-up motion. Figure 1 represents a typical example of manner-encoding gestures showing the climb-up motion event resulting from the stimulus. The speakers wiggled both hands one after the other, which depicted the motion of the cat from the C-VPT. This form of the gesture resembles the climb-up motion of the cat wiggling both paws when climbing up the drainpipe in the stimulus. Regardless of the use of words or phrases spoken, the accompanying manner-encoding gestures of different speakers are similar in terms of their shapes. Another example of path-encoding gestures, with one hand moving up, depicted the path of the character from the O-VPT as shown in Figure 2.

Figure 1. The screenshots at the point when speakers wiggled both hands one after the other to depict manner information of the character

Figure 2. The screenshots at the point when speakers moved the right hand with right finger raised up to depict the path of the character

Considering the high degree of similarity between gestural shapes among all the speakers, the effect of specific language preferences is questioned. Based on the results, it is reasonable to suppose that gestures are derived from the GP as a starting point of thought coming from the information in the stimulus rather than from the effect of linguistic preferences regarding information encoding (McNeill 1992, McNeill & Duncan 2000). As speech is expressed in completely or slightly different ways in different languages (Talmy 1985, Slobin 1991), it is natural to expect that, if linguistic preferences affect the means of encoding information in gestures, the gestures used by speakers with different language backgrounds would demonstrate significant variation according to their L1. However, this assumption does not explain the high degree of uniformity of gesture shapes with semantic synchrony across all the speakers observed in this study. Whereas it is acknowledged that there is a slight difference in terms of gesture space, in which kinesic forms and movements are performed (see McNeill 1992, Kendon 1996), similar shapes are seen in the gestures used in the present study. However, note that the result is tentative with only six speakers for each group, and thus further investigation is required to generalise this point.

5. Discussion: an implication of the impact of translation equivalents in L1-L2

As described so far, no Japanese L1 speakers in L1 narratives depicted the event with manner expressions in their speech, whereas the English L1 speakers in L1 narratives tended to use manner information encoded in their speech. This pattern is actually also seen in their gestural expressions. Namely, a greater number of English speakers used manner-encoding gestures than Japanese L1 speakers. This point is supported by recent findings claiming that languages

stimulate different forms of thinking even when speakers express the same motion events (Kita 2000, McNeill & Duncan 2000, Kita & Özyürek 2003). More specifically, this cross-linguistic difference in speech and gesture patterns between Japanese and English conforms to the suggestion by Brown & Gullberg (2008) that more manner expressions in gestures are found for English speakers than Japanese speakers. This result suggests that linguistic preference has an effect on gestural preferences in encoding information.

Nevertheless, it should also be acknowledged that not every EM, EB, and JB speaker when speaking English used manner-encoding gestures. Rather, even speakers of English as L1 decided to encode the path information rather than the manner information in more than half of the total cases; this does not seem so significant when compared to the data from the Japanese L1 speakers. Arguably, this tendency contradicts the idea that due to linguistic preference the speakers paid more attention to manner information of the motion and hence used more manner-encoding gestures. Unfortunately, the study could not conduct a statistical analysis due to the limited number of samples, even though it has offered a micro-analytic perspective, and thus leaves this question for future studies.

Rather, what the micro-level analysis highlights is that very few studies to date have considered the difference in the semantic information contained in specific words between the original word and its translation equivalent. For example, the Japanese word *noboru* 'climb (up)' is claimed to be a path-encoding verb as it comprises only the meaning of the up direction, although the word is in most cases translated as *climb* in English, a manner-encoding verb, which can be co-expressed with adverbials encoding the meaning of both 'up' and 'down' directions (Brown & Gullberg 2011: 91). As for speech descriptions, the emphasis should be firstly on the classification of verbs as encoding either manner or path information. The participants recruited in the present study used a variety of verbs encoding the climb-up motion event (refer back to Table 2). Here, the novel finding from the post-recording questionnaire is that bilingual speakers judged semantic concepts of motion verbs differently from monolingual speakers when speaking their L2. Interestingly, the study reveals that JB speakers treated *climb* in English as a path verb since they associated the word with only the meaning of the up direction, saying that *climb down* is semantically unacceptable. This wrong semantic judgement is also seen in EB speakers, as they treated *noboru* 'climb (up)' as a manner verb because they thought that the Japanese verb could be used to refer to either the up or the down direction, although Japanese L1 speakers do not use the verb in the latter case.

With this classification in mind, it is observed that there were no cases of manner expressions in speech by JM and JB speakers in L1 narratives. This is consistent with the suggestion that Japanese, a verb-framed language, tends to encode manner in adverbials rather than verbs (Talmy 1985, Sugiyama 2005, Slobin 2006). This is not the case for the EM speakers and EB speakers when speaking Japanese as their L2. Regarding the latter case, their Japanese descriptions include the use of '*noboru*' as a manner verb. Although it is true that some EM and EB speakers encode the path information verbally with path verbs, such as 'go', all Japanese L1 speakers encoded only the path information. This supports the claim that speakers of a verb-framed language do not encode manner as much as those of a satellite-framed language (Hohenstein, Eisenberg & Naigles 2006, Brown & Gullberg 2008). In the present study, L1 speakers of Japanese used various verbs to depict the motion event. However, all of them are semantically classified as path verbs encoding the up direction. From this perspective, it is indeed arguable that no cases of manner-encoding verbs in the descriptions by Japanese speakers could be simply explained by limited linguistic entries in Japanese; since EB speakers also encoded no manner information in Japanese speech when strictly considering *noboru* 'climb (up)' as a verb containing the meaning of 'only' an upward direction in a strict sense.

Given these points, it should be acknowledged that linguistic resources for depiction of a climb-up motion are not equivalent between the two languages due to the variety of semantic

conceptions of words and limited lexical entries in Japanese compared to English. Although the data samples evidence L1 speakers of English focusing on manner, such language qualities arguably do not allow an equivalent comparison of different patterns in speech between these two languages. Therefore, the study does not take into account the different patterns in speech depicting manner or path information of the motion event.

6. Conclusion

The present study shows the high degree of semantic synchrony between gestures and their accompanying speech units. Furthermore, it is notable that there is a uniformity of gesture-in-use not only in the gestural phrase but also in the actual shapes encoding information. These patterns are highly recognisable in both languages in this study, regardless of speakers' linguistic background. Although these findings support the findings of previous empirical gesture studies, this study goes further in that it highlights significant challenges present when investigating cross-linguistic differences in terms of monolingual and bilingual speakers. This micro-level analysis of speech and gesture production by monolinguals and bilinguals has implications regarding the inequivalent semantic concepts of *climb* and its Japanese translation. Nonetheless, the findings presented here require qualification, as there were only six speakers recruited for each group.

References

Alibali, M. W., Heath, D. C. & Myers, H. J. (2001). Effects of visibility between speaker and listener on gesture production: Some gestures are meant to be seen. *Journal of Memory and Language* 44, 169-188.

Bavelas, J. B., Chovil, N., Lawrie, D. A. & Wade, A. (1992). Interactive gestures. *Discourse processes* 15(4), 469-489.

Beattie, G. & Shovelton, H. (2002). An experimental investigation of some properties of individual iconics that mediate their communicative power. *British Journal of Psychology* 93(2), 179-192.

Birdwhistell, R. L. (1970). *Kinesics and context: Essays on body motion communication*. Philadelphia: University of Pennsylvania Press.

Brown, A. & Gullberg, M. (2008). Bidirectional crosslinguistic influence in L1-L2 encoding of manner in speech and gesture: A study of Japanese speakers of English. *Studies in Second Language Acquisition* 30(2), 225-251.

Brown, A. & Gullberg, M. (2011). Bidirectional cross-linguistic influence in event conceptualization? Expressions of path among Japanese learners of English. *Bilingualism: Language and Cognition* 14(1), 79-94.

Chawla, P. & Krauss, R. (1994). Gesture and speech in spontaneous and rehearsed narratives. *Journal of Experimental Social Psychology* 30(6), 580-601.

Clark, H. H. & Krych, M. A. (2004). Speaking while monitoring addressees for understanding. *Journal of Memory and Language* 50(1), 62-81.

Cohen, A. A. & Harrison, R. P. (1973). Intentionality in the use of hand illustrators in face-to-face communication situations. *Journal of Personality and Social Psychology* 28(2), 276-279.

Freleng, I. (Director). (1950). *Canary Row* [Animated film]. New York: Time Warner.

Garfinkel, H. (1996). Ethnomethodology's program. *Social Psychology Quarterly* 59 (1), 5-21.

Grosjean, F. (1998). Studying bilinguals: Methodological and conceptual issues. *Bilingualism: Language and Cognition* 1(2), 131-149.

Gullberg, M. (2003). Gestures, referents, and anaphoric linkage in learner varieties. In Dimroth, C. & Starren, M. (eds.), *Information structure and the dynamics of language acquisition*, 311-328. Amsterdam: John Benjamins.

Gullberg, M. & Indefrey, P. (2003). *Language history questionnaire*. http://www.mpi.nl/departments/other-research/research-projects/the-dynamics-of-multilingual-processing/tools/ [Retrieved on 06 September, 2014]

Hohenstein, J., Eisenberg, A. & Naigles, L. (2006). Is he floating across or crossing afloat? Cross-influence of L1 and L2 in Spanish-English bilingual adults. *Bilingualism: Language and Cognition* 9(3), 249-261.

Holler, J. & Beattie, G. (2002). A micro-analytic investigation of how iconics and speech represent core semantic features in talk. *Semiotica* 142, 31-70.

Holler, J. & Beattie, G. (2004). The interaction of iconic gesture and speech in talk. In *Gesture-Based Communication in Human-Computer Interaction*, 63-69. Berlin: Springer.

Kendon, A. (1980). Gesticulation and speech: Two aspects of the process of utterance. In Key, M. R. (ed.), *The relationship of verbal and nonverbal communication*, 207-228. New York: Mouton Publishers.

Kendon, A. (1994). Human gesture. In Gibson, K. R. & Ingold, T. (eds.), *Tools, language and cognition in human evolution*, 43-62. Cambridge: Cambridge University Press.

Kendon, A. (1996). An agenda for gesture studies. *Semiotic review of books* 7(3), 8-12.

Kendon, A. (1997). Gesture. *Annual Review of Anthropology* 26, 109-128.

Kendon, A. (2004). *Gesture: Visible action as utterance*. Cambridge: Cambridge University Press.

Kita, S. (2000). How representational gestures help speaking. In McNeill D. (ed.), *Language and gesture,* 162-185. Cambridge: Cambridge University Press.

Kita, S. & Özyürek, A. (2003). What does cross-linguistic variation in semantic coordination of speech and gesture reveal? Evidence for an interface representation of spatial thinking and speaking. *Journal of Memory and Language* 48(1), 16-32.

Kita, S., van Gijn, I. & van der Hulst, H. (1998). Movement phases in signs and co-speech gestures, and their transcription by human coders. In Wachsmuth, I. & Fröhlich, M. (eds.), *Gesture and sign language in human-computer interaction*, 23-35. Berlin: Springer.

Kipp, M., Neff, M. & Albrecht, I. (2007). An annotation scheme for conversational gestures: how to economically capture timing and form. *Language Resources and Evaluation* 41(3), 325-339.

McCafferty, S. G. (1998). Nonverbal expression and L2 private speech. *Applied Linguistics* 19(1), 73-96.

McNeill, D. (1985). So you think gestures are nonverbal?. *Psychological Review* 92(3), 350-371.

McNeill, D. (1992). *Hand and mind: What gestures reveal about thought*. London: University of Chicago Press.

McNeill, D., & Duncan, S. D. (2000). Growth points in thinking-for-speaking. In McNeill, D. (ed.), *Language and gesture,* 141-161. Cambridge: Cambridge University Press.

McNeill, D. & Levy, E. T. (1982). Conceptual representations in language activity and gesture. In Jarvella, R. J. & Klein, W. (eds.), *Speech, place, and action: Studies of deixis and related topics,* 271-295. New York: Wiley.

Morris, D. (1979). *Gestures, their origins and distribution*. London: Jonathan Cape.

Ohara, K. H. (2002). Linguistic encodings of motion events in Japanese and English: A preliminary look. *Hiyoshi Review of English Studies* 41, 122-153.

Richmond, V. P. & McCroskey, J. C. (1999). *Nonverbal behavior in interpersonal relations*. 4th ed, Boston: Allyn and Bacon.

Slobin, D. I. (1991). Learning to think for speaking: Native language, cognition, and rhetorical style. *Pragmatics* 1(1), 7-25.

Slobin, D. I. (2004a). How people move: Discourse effects of linguistic typology. In Moder, C. L. & Martinovic-Zic, A. (eds.), *Discourse across languages and cultures,* 195-210. Amsterdam: John Benjamins.

Slobin, D. I. (2004b). The many ways to search for frog: linguistic typology and the expression of motion events. In Stromqvist, S. & Verhoeven, L. (eds.), *Relating events in narrative, Volume 2: Typological and contextual perspectives,* 219-257. Mahwah: Lawrence Erlbaum Associates.

Slobin, D. I. (2006). What makes manner of motion salient?. In Hickmann, M. & Stephane, R. (eds.), *Space in languages: Linguistic systems and cognitive categories,* 59-81. Amsterdam: John Benjamins.

Sugiyama, Y. (2005). Not all verb-framed languages are created equal: The case of Japanese. *Proceedings of the Annual Meeting of the Berkeley Linguistics Society* 31(1), 299-310.

Talmy, L. (1985). Lexicalization patterns: Semantic structure in lexical forms. In Shopen, T. (ed.), *Language typology and syntactic description, Volume. 3: Grammatical categories and the lexicon*, 57-149. Cambridge: Cambridge University Press.

ten Have, P. (1999). *Doing conversation analysis: A practical guide*. London: Sage.

Vygotsky, L. S. (1986). *Thought and language* (A. Kozulin, Trans.). Cambridge: The Massachusetts Institute of Technology Press.

Wilcox, S. (2004). Language from gesture. *Behavioral & Brain Sciences* 27(4), 525-526.

Wittenburg, P., Brugman, H., Russel, A., Klassmann, A. & Sloetjes, H. (2006). ELAN: A professional framework for multimodality research. In *Proceedings of International Conference on Language Resources and Evaluation 2006*, 1556-1559.

Kazuki Hata
School of Education, Communication and Language Sciences
King George Building VI
NE1 7RU
United Kingdom

k.hata@newcastle.ac.uk

Kazuki Hata
School of Education, Communication and Language Sciences

TREATMENT OF PHONOLOGICAL ANOMIA IN APHASIA: SOME SUGGESTIONS FOR USERS OF SIGNED LANGUAGE*

VANESSA MCINTOSH
(Newcastle University)

Abstract

Aphasiologists agree that with damage to the left hemisphere of the brain, both those who use vocal and those who use signed forms of communication are in jeopardy of developing aphasia, an impaired ability to produce and/or perceive language. Of the many symptoms that are associated with aphasia, anomia (a word retrieval impairment of language) is the most common, but while treatments for phonological anomia (a type of anomia caused by defective phonological processes) are available for users of spoken language, the same cannot be said for users of signed language. This paper aims to fill that gap in the literature by making some suggestions for the treatment of phonological anomia in signed language users. In spoken languages, phonological word finding deficits can be remediated by phonemic cueing tasks in which the initial phonemes of target words are provided as cues to phonologically prime targets and facilitate their retrieval. In this paper I argue that by simply substituting phonemes as cues for cheremes (the signed language equivalent of phonemes), we could adapt the phonemic cueing approach to treatment of phonological anomia in spoken languages to make it suitable for use with signed language users.

1. Introduction

The study of signed language linguistics has long established that signed languages are natural human languages in the way that spoken ones are (Stokoe 1960, Klima & Bellugi 1979, Senghas et al. 2004) and that while the two modalities differ in their methods of production and perception, their underlying linguistic structures are the same (Sandler & Lillo-Martin 2006). We also know from brain imaging studies that the left cerebral hemisphere of the brain is responsible for processing all modalities of language whether those be spoken or otherwise (Bavelier et al. 1998, Hickok et al. 2002), so one would expect that when deaf signers endure neurological damage to their left cerebral hemisphere, the consequences for language for them would be very much the same for them as they are for hearing speakers.

This expectation seems to be supported: aphasiologists agree that with left cerebral hemisphere damage those who use vocal forms of communication are in jeopardy of developing aphasia, an impaired ability to produce and/or perceive language (Naeser & Hayward 1978, Kertesz et al. 1979, Damasio 1989, Damasio 1992, Fitzgerald 1996, Corina 1998, Kolb & Whishaw 2003, Kearns 2005, Keller et al. 2009) and a number of studies have confirmed this to be the case for signed language users too (Leischner 1943, Tureen et al. 1951, Douglas & Richardson 1959, Poizner et al. 1987, Corina et al. 1992, Hickok et al. 1998, Corina et al. 1999, Hickok et al. 2011). One of the symptoms most commonly shared among people with aphasia is anomia, a word retrieval impairment of language (Goodglass & Wingfield 1997). Two different types of anomia have been described in the aphasiology literature: semantic anomia, which is caused by defective semantic processes and phonological anomia, which is caused by problems with the phonological processing system (Nettleton & Lesser 1991, Raymer et al.

1993, Hillis & Caramazza 1994, Boyle & Coelho 1995, Miceli et al. 1996, Nickels & Best 1996, Drew & Thompson 1999, Kiran & Thompson 2003), but while treatment for phonological anomia is available for users of spoken language, the same cannot be said for users of signed language. This paper aims to fill that gap in the literature by making some suggestions for the treatment of phonological anomia in signed language aphasia from the perspective of theoretical linguistics.

The body of discussion is categorised into five main parts (Sections 2, 3, 4, 5 and 6). The first part (Section 2), presents an introductory account of the major characteristics or symptoms of aphasia observed in spoken and signed languages; while Section 2.1 addresses aphasia in spoken languages, Section 2.2 addresses it in signed ones. Section 3 hones in on anomia, an impairment of language that is common in aphasia and characterised by an inability to access words in and retrieve them from the mental lexicon. It describes the differences between semantic anomia and phonological anomia, and points out the need for treatment for phonological anomia in signed languages. Sections 4 and 5 give an overview of the theory underpinning phonologically targeted treatment for phonological anomia. While Section 4 concerns itself with how the mental lexicon is organised and how words are accessed in it, Section 5 describes how phonological priming, an implicit memory effect that is triggered in phonological intervention for phonological anomia, can influence lexical access. Section 6 describes a particular type of treatment for phonological anomia for hearing speakers with aphasia (phonemic cueing treatment) before making some suggestions for its adaptation for use with signed language users. Section 7 summarises and concludes.

2. Aphasia
2.1. Spoken language aphasia

Aphasia is an acquired disorder of language that presents as a result of damage to the brain areas responsible for the production and/or comprehension of language and its components (Helm-Estabrooks et al. 2014). Depending on the site or sites of damage and the relative impairment or sparing of the language components, different subtypes of aphasia can emerge (Helm-Estabrooks et al. 2014). There are two major subtypes of aphasia: Broca's aphasia and Wernicke's aphasia (which are caused by damage to the lateral frontal, suprasylvian, pre-Rolandic area or the adjacent subcortical periventricular white matter, and the posterior third of the superior temporal gyrus, respectively). Broca's aphasia is otherwise known as expressive aphasia and is so named because it is characterised by impaired language production with unimpaired language comprehension (Kerterz 1982). Wernicke's aphasia, on the other hand, (which can also go by the name of receptive aphasia) causes problems of a similar vein but in the opposite direction: while a Broca's patient experiences problems with language production but not its comprehension, a Wernicke's patient has them with its comprehension, but not its production (Albert et al. 1981). Other more minor subtypes of aphasia include conduction aphasia, transcortical motor aphasia, transcortical sensory aphasia and global aphasia (Helm-Estabrooks et al. 2014), and with all of the different subtypes, different symptoms are associated. Some level of anomia, meanwhile, is seen in all of the aphasias (Manasco 2014).

2.2. Signed language aphasia

It has long been established that the signed languages of deaf communities are fully-fledged human languages in the way that the spoken languages of hearing communities are (Stokoe 1960, Klima & Bellugi 1979, Senghas et al. 2004) in that there are a number of autonomous mutually unintelligible systems of communication used in deaf communities around the world (Woll et al. 2001, MacSweeney et al. 2008). It is also generally accepted by linguists that while

spoken languages and signed languages differ in their methods of production and perception (in that spoken languages are produced by the mouth and perceived by the ears whereas signed languages are produced by the hands and perceived by the eyes), their underlying linguistic structures are the same (Sandler & Lillo-Martin 2006). This is primarily because key linguistic features (or language universals) in the theoretical domains of phonology, syntax and semantics identified in spoken languages have been identified in signed languages as well, but similarities observed between the neural bases of spoken and signed languages have also been taken to reflect that all linguistic functioning is underpinned by the same system. Neuroimaging studies have found evidence to suggest that like spoken languages, signed languages are processed by the left hemisphere of the brain (see Hickok et al. 1998), that both speech and sign production is processed by the left inferior frontal gyrus (Braun et al. 2001, Corina et al. 2003, Emmorey et al. 2003, Emmorey et al. 2007) while speech and sign comprehension rely on the left superior temporal gyrus and sulcus for their processing (Neville et al. 1998, MacSweeney et al. 2002, Newman et al. 2002, Meyer & Damien 2007, Corina et al. 2007, Capek et al. 2008).

If signed language processing like spoken language processing is left dominant, this should be reflected by the patterns of impairment that follow brain damage (Marshall et al. 2004), should it not? That is, aphasia in signed languages should follow left- but not right-hemisphere damage in the way that it does in spoken ones. A number of lesion studies in signers have confirmed that this is indeed the case (e.g. Poizner et al. 1987, Corina 1998, Marshall et al. 2004 and Atkinson et al. 2005) and there is even evidence to suggest that the symptoms found in signed language impairments are consistent with those found in spoken language impairments.

In 1987, for example, Poizner et al. reported six cases of unilateral stroke in deaf users of American Sign Language (ASL). While those with left hemisphere damage presented with language problems in their case studies, those with right hemisphere damage did not. What's more they found that damage to the left frontal regions resulted in production difficulties (as we see in spoken language Broca's patients) while damage to the left temporal lobe caused problems with comprehension (just like in spoken language Wernicke's aphasia). These findings have been supported by a number of other signed language researchers (see Corina et al. 1992 for just one representative example).

People with signed language aphasia have also been said to experience the word retrieval difficulties that people with spoken language aphasia are all too familiar with. Marshall et al. (2004), for instance, described a case of aphasia in a user of British Sign Language (BSL) called Charles which arose as a result of a left cerebrovascular accident. A computer tomography (CT) scan revealed damage to the left posterior frontal and parietal lobes extending into the corona radiate and temporal lobe, and as one would predict, Charles' resultant language problems were with production. He struggled to retrieve the correct forms of signs and frequently was unable to access signs at all, but in contrast to his production problems, his comprehension of signs was unaffected. In an informal comprehension test Charles scored 10/10, indicating that he had a Broca's-like aphasia.

Convinced that Charles had anomia as a symptom of his aphasia, Marshall et al. (2004) administered a number of assessments to see whether their predictions could be proven. In a picture naming task which required him to name pictures on confrontation Charles scored only 27/37, leading Marshall et al. (2004) to conclude that he did indeed have a problem with the retrieval of words. In 1982 Chiarello et al. witnessed a similar situation in their investigation of a prelingually deaf woman who had suffered a left parietal infarct involving the left supramarginal and angular gyri. She too developed a sign retrieval impairment that mirrored the word retrieval impairments of oral aphasic patients.

Now although there is evidence to suggest that anomia is common to both the spoken and the signed modalities of language, there has been little to no research into its treatment in

signed language aphasia. In this paper I aim to inform the neglected area of research by suggesting that we adapt current treatment for phonological anomia in spoken languages for use with signed language users, but before any suggestions can be made, however, we must first look at anomia in more detail.

3. Anomia

As we observed above, anomia is an impaired ability at accessing words in and retrieving them from the mental lexicon (Goodglass & Wingfield 1997) which can range from a mild difficulty in producing desired words during conversational discourse to a virtual inability to produce them under any conditions at all (Helm-Estabrooks et al. 2014). This section will provide a more thorough, comprehensive account of anomia. It will begin with an overview of its causes before describing its symptoms, diagnosis and treatment.

3.1. Its causes

As a symptom of aphasia, anomia is thought to result from damage to regions in the left hemisphere of the brain (Woollams et al. 2008). This damage can either be traumatic or acquired (Damasio 1992) with causes of traumatic brain injury including the likes of falls, accidents or violence involving a blow to the head and causes of acquired brain injury poisoning, infection, strangulation, choking, drowning, stroke, heart attacks, brain tumours, aneurysms and even neurodegenerative diseases (such as Parkinson's disease, Alzheimer's disease and Huntington's disease) (Budd et al. 2010). It was initially thought that damage to Broca's area (the speech production centre of the brain) or Wernicke's area (the speech comprehension centre) could be held responsible for the onset of anomia (see Fridriksson et al. 2009 and Hamilton et al. 2010, for example), but more recent studies have suggested that anomia is caused by damage to the arcuate fasciculus (the nerve tract that connects the caudal temporal cortex and inferior parietal cortex to locations in the frontal lobe) more specifically (Catani & Thiebaut de Schotten 2008, Carlson 2012).

3.2. Its symptoms

The most accurate way of determining whether anomia has developed in an individual is through tests, though there are a number of symptoms that can point to its presence. Unsuccessful attempts at word retrieval tend to result in speech errors that resemble the target in meaning (semantic speech errors) or in sound (phonological speech errors), which are therefore taken to be symptomatic of anomia.

Paraphasias are a particular type of speech error that are commonly observed in people with aphasia. These can be divided into two broad classes: semantic paraphasias in which an irretrievable target word is substituted with a word that is semantically related to it such as *refraction* for *reflection* (which both are to do with the dispersion of light) (Damasio 1992, Fitzgerald 1996, Marshall et al. 1998) and phonological paraphasias which substitute irretrievable target words with words or nonwords that resemble them phonologically like *viscosity* ([vɪskɒsɪti]) for *velocity* ([vɪlɒsɪti]) (a real word, formal error) or *chromosone* ([krəʊməsəʊn]) for *chromosome* ([krəʊməsəʊm]) (a nonword phonemic error) (Brookshire 1997).
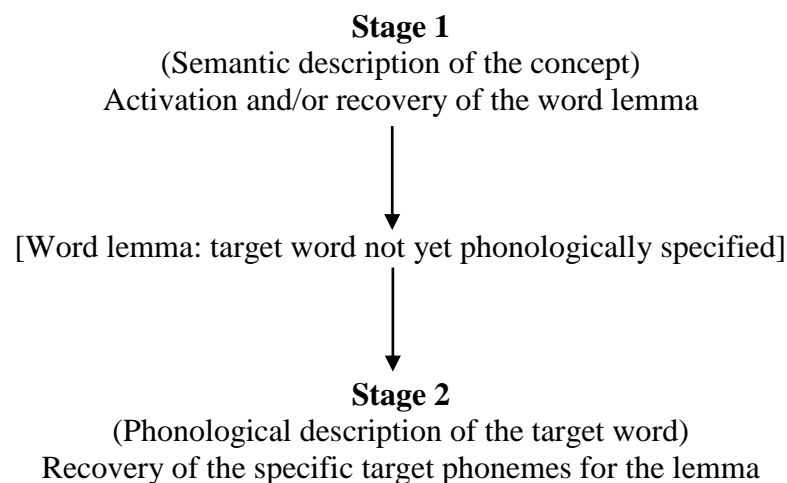
Semantic paraphasias suggest that while phonological information about a target word can be successfully retrieved by the patient, not all of its semantic information can be. Phonological paraphasias, on the other hand, in particular phonemic errors in which only partial phonological information about a target is retrieved and the rest is deformed by the likes of

addition, deletion, substitution or transposition (Lecours & Lhermitte 1969, Blumstein 1973) suggest that while semantic information about a target is available for retrieval, there are problems with a patient's access to the sounds constituting its spoken form.

3.3. Its diagnosis

In 1979 Benson drew a distinction between two types of anomia: semantic anomia and phonological anomia (Devinsky & D'Esposito 2004) and one way that researchers have tried to understand the difference between the two and how to diagnose them is by studying word retrieval errors like the paraphasias described above. An approach that many aphasiologists find useful for diagnosis of the anomias is one that characterises a patient's word retrieval problems with reference to a cognitive neuropsychological model of the word retrieval process outlining the different information processing stages involved in it (Ellis et al. 1992) as exemplified in Figure 1. Armed with a model like this, aphasiologists are able to relate the speech errors an aphasic patient produces to breakdown of one or more of the model's components (and ergo of one or more of a patient's cognitive operations):

Figure 1. A two-stage model of word retrieval (adapted from Helm-Estabrooks et al. 2014: 232)

Stage 1
(Semantic description of the concept)
Activation and/or recovery of the word lemma

↓

[Word lemma: target word not yet phonologically specified]

↓

Stage 2
(Phonological description of the target word)
Recovery of the specific target phonemes for the lemma

Aphasiologists generally agree that word retrieval begins with the recovery of semantic information about a target word and ends with the recovery of its phonological information (Allport 1985), so two-stage models of word retrieval like the one in Figure 1 are helpful in that they draw a distinction between those two stages of the word retrieval process: a first stage in which an object is perceived and comprehended and a second stage in which the name for the object is retrieved (Helm-Estabrooks et al. 2014). It is thought that stages 1 and 2 of the model constitute separable cognitive components (the semantic system and the phonological system) that can be selectively impaired giving rise to two different types of anomia (semantic anomia and phonological anomia, respectively). It is thought that semantic paraphasias are caused by defective semantic processes and are therefore symptomatic of semantic anomia, while phonological paraphasias (caused by problems with the phonological processing system) are a symptom of phonological anomia, that failure to retrieve enough semantic information about a target to activate its appropriate word lemma at stage 1 leads to semantic speech errors, and failure to adequately activate stage 2's phonological description of a target causes phonological errors (Helm-Estabrooks et al. 2014).

Now although we can make a preliminary diagnosis of a patient's anomia by looking at the types of error they typically produce, all provisional diagnoses must be confirmed by a proper evaluation. A patient's comprehension of the words he or she is unable to retrieve might be evaluated using the Pyramids and Palm Trees test (Howard & Orchard-Lisle 1984), for example. Object recognition and comprehension can be assessed using the all picture version of the test which requires a patient to indicate which of two pictured objects belongs with a third pictured object while recognition and comprehension of their names can be assessed by asking a patient which of two pictured objects matches the name of a third object.

The Psycholinguistic Assessments of Language Processing in Aphasia (PALPA) Battery (Kay et al. 1992) can also be used to test for semantic word retrieval impairments. Comprehension of object names is assessed with a 40-item test in which a patient must indicate which of five pictures (including the target word itself, two semantic distractors, a visual distractor that is related in appearance but not in function to the target word and a word that is related in function to the visual distractor but not in its appearance) matches a heard word. Patients with semantic impairments will tend to make errors involving the semantic distracters, but if there is no obvious pattern in a patient's speech errors, their deficit will probably not be semantic in nature Kay et al. (1992) claim.

The possibility of a phonological impairment might then be explored using a set of pictures that was published by Snodgrass and Vanderwart in 1980. Patients who show an effect of word frequency when matching words to these pictures are often described as having phonological anomia, as are patients who have problems with the articulation of the longer words but not the shorter ones or who can demonstrate their understanding of the words but are unable to produce them (Ellis et al. 1992).

3.4. Its treatment

A number of different approaches have been shown to be effective in the remediation of word retrieval impairments in aphasia (Howard et al. 1985, Pring et al. 1990, Hillis & Caramazza 1994, Fick et al. 2002, Hickin et al. 2002, Raymer & Ellsworth 2002). Some aphasiologists argue that treatment should be targeted to the damaged level of language processing, that word retrieval difficulties caused by impaired word meaning respond best to semantically targeted treatments while phonologically targeted treatments are best used to treat word retrieval difficulties caused by damage to the phonological processing system (Nettleton & Lesser 1991, Miceli et al. 1996). Semantically targeted tasks include the likes of matching words to pictures, answering yes/no questions about pictures and picture categorisation (Marshall et al. 1990, Davis & Pring 1991, Byng 1995, Nickels & Best 1996) and phonologically targeted tasks phonemic cueing (whereby the word-initial phonemes of target words are provided to patients for processing) (Crofts et al. 2004).

While it has indeed been demonstrated that semantically targeted tasks are effective in the treatment of semantic anomia (e.g. Marshall et al. 1990, Pring et al. 1990, Nickels & Best 1996) there is also evidence to suggest that they can improve word retrieval for those who do not have an impairment at the semantic level (Nickels & Best 1996). This is also true for phonologically targeted tasks. It was also originally thought that phonological interventions for anomia should be used exclusively for those with phonological impairments (Nettleton & Lesser 1991, Hillis & Caramazza 1994, Micili et al. 1996) but they have in actual fact been found to be successful at treating word retrieval impairments at the semantic level as well (Raymer et al. 1993, Nickels & Best 1996). The success of both types of treatment at improving word retrieval in each of the anomias could be attributed to the fact that they are not actually entirely different from one another after all, however. Howard (2000) points out that in semantically targeted treatments, the spoken form of a target word is provided evoking

phonological processing and in phonologically targeted treatments, a picture of a target is presented to the patient evoking semantic processing so in each of the treatments, both semantic and phonological information is available to the patient meaning they have their effects in the exact same way: by strengthening the connections between word meaning and word form.

Since Howard's (2000) observation predicts equivalent effects for individuals with semantic and phonological anomia with each of the different types of treatment, it may well be the case that adapting phonologically targeted treatments for use with signed language users could have consequences for individuals with semantic anomia as well as those with phonological anomia, but given that few other studies have examined the issue I think it would be presumptuous to make such a claim at such an early stage. In this paper, I will be making suggestions for the treatment of phonological anomia in signed languages only and my proposal is as follows. If it is true that the same phonological system that underlies spoken languages underlies signed ones, phonological processing should occur in each modality in the same way, meaning treatments that target phonological processing to improve word retrieval in hearing people with phonological anomia could have equivalent effects for deaf phonological anomics provided they were made suitable for their methods of production and perception. But before we begin to think too deeply about the ways in which word retrieval may be influenced, we must first explore how words are stored in the mental lexicon, how they are organised with respect to one another and how they can be accessed.

4. The mental lexicon

Psycholinguists argue that the underlying representations of words (and thus our knowledge of them) are stored in permanent memory in a mental lexicon (Collins & Quillian 1969, Caron 1989) from which they must be accessed. The study of lexical access sets out to answer two types of question: the first concerns how units are represented and organised while the second bears on the processes involved in accessing items in the lexicon. This section begins by addressing the first question before addressing the second.

4.1. Organisation of the mental lexicon
4.1.1. Hierarchical network models

Collins & Quillian's (1969) hierarchical network model was the first systematic model of semantic memory. It suggested that semantic memory is organised into a series of hierarchical networks consisting of nodes and their properties where nodes represent concepts such as *animal, bird* and *canary* and properties, rather self explanatorily, the characteristics of those concepts. The model is arranged with broader concepts on the higher levels and narrower concepts on the lower levels as exemplified in Figure 2:

Figure 2. A hierarchical network model of semantic information related to animals (Collins & Quillian 1969: 241)
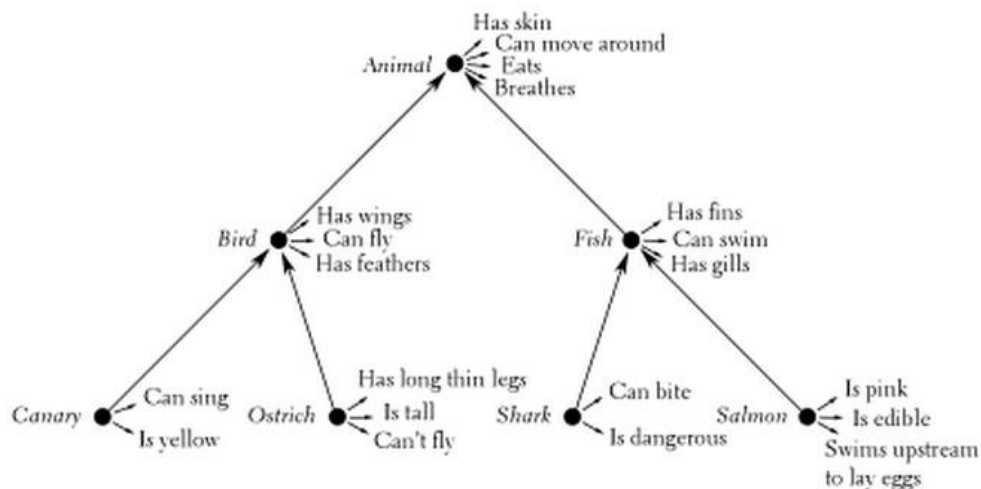

In Figure 2, the concept ANIMAL dominates the concepts BIRD and FISH as both birds and fish are animals. Likewise, CANARY and OSTRICH are dominated by BIRD as they are both types of bird and SHARK and SALMON by FISH as they are both types of fish, but although this sort of organisation in which broader categories are organised higher in the hierarchy than narrower categories seems to be logical, it also has its problems. Collins & Quillian's (1969) hierarchical network model assumes that all items on a given level of a hierarchy are more or less equal. Since SHARK and SALMON are both subordinates of fish and are organised on the same level as each other, they should, in theory, take the same amount of time to access in the mental lexicon. In actual fact, they do not, and this generally seems to be the case (Carroll 2008). Smith et al. (1974) found that items more typical of a given subordinate take less time to access than atypical items. For example, in everyday conversation, the word *salmon* is more likely to come up than the word *shark* is as it is a foodstuff. Being a higher frequency word its concept is accessed in the mental lexicon more quickly (Smith et al. 1974) and this 'typicality effect' suggests that a hierarchical network might not necessarily be the best way to model the mental lexicon after all.

### 4.1.2. Spreading activation models

A second class of models known as spreading activation models retain Collins & Quillian's (1969) idea of a network but discard their hierarchical assumption. Collins & Loftus (1975) propose that the organisation of words is more like a web of interconnecting nodes than a hierarchy of them. They argue that retrieval occurs by a process of spreading activation which begins at a single node and spreads to its neighbouring nodes in the network attenuating over distance, ensuring that closely related concepts are more likely to be activated than distant concepts are as in Figure 3:

Figure 3. A schematic representation of a spreading activation model of semantic knowledge (Collins & Loftus 1975: 412)
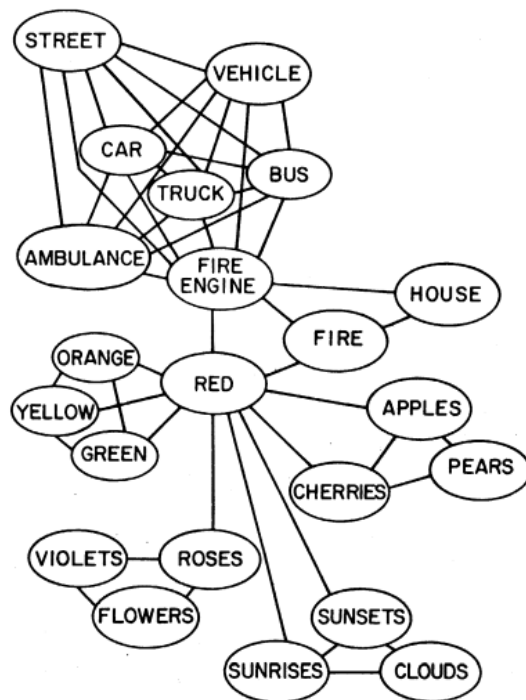
According to Collins and Loftus' (1975) model, if the concept CAR was activated, the STREET, VEHICLE, BUS, TRUCK, AMBULANCE and FIRE ENGINE concepts are more likely to be activated than the SUNSETS, SUNRISES and CLOUDS concepts. This makes sense given that the first group of concepts are more similar in meaning to the concept CAR than the second group of concepts, but although Collins & Loftus' (1975) model is a step forward from Collins and Quillians' (1969) model, it too has its limitations. Very little information is paid to the phonological and syntactic aspects of words so in a sense it is more a model of concepts than a model of words.

A more recent spreading activation model that incorporates lexical as well as conceptual aspects was established by Bock & Levelt (1994):

Figure 4. Bock & Levelt's (1994) spreading activation model (Bock & Levelt 1994: 951)
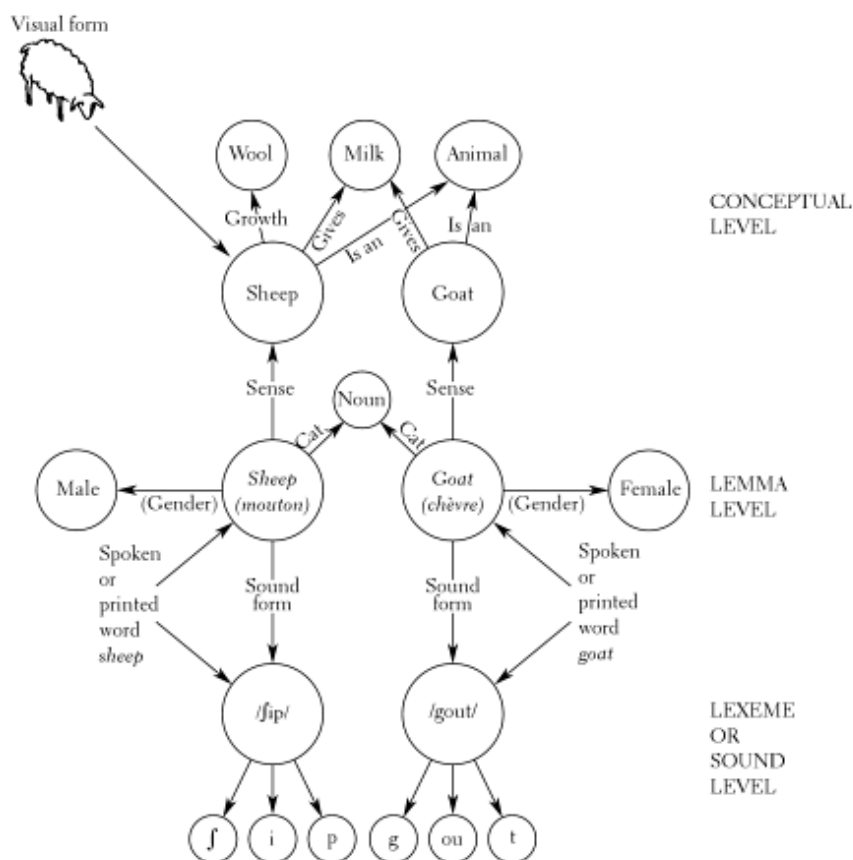
Bock & Levelt (1994) assume that our knowledge of words exists at three different levels: a conceptual level, a lemma level and a lexeme level where the conceptual level is comprised of concepts, the lemma level of the syntactic aspects of word knowledge and the lexeme level a word's phonological properties (Levelt 1989). Spreading activation models like this one which incorporate conceptual, syntactic and phonological knowledge are thought to offer the most realistic picture of the mental lexicon (Posner & Snyder 1975, Neely 1977, 1991, Marcel 1983), but if this is how words are organised in the mental lexicon, how are they then accessed in it? This question is addressed in Section 4.2.

4.2. Lexical access

Theories of lexical access generally agree that word retrieval occurs in two stages: a first stage known as lemma access (which is the mapping from a conceptual representation to a lemma (an abstract conceptual form of a word that has been mentally selected for utterance in the early stages of speech production but has not yet assumed its phonological form)) and a second stage called phonological access (the mapping from the lemma to its phonological representation) (Fromkin 1971, Garrett 1975, Potter & Faulconer 1975, Fay & Cutler 1977, Kempen & Huijbers 1983, Dell 1986, Butterworth 1989, Glaser & Glaser 1989, Levelt 1989, La Heij 1990, Levelt et al. 1991, Roelofs 1992, Bock & Levelt 1994). The relationship between lemma selection and phonological encoding, however, has been subject to serious debate. While some models of lexical access assume that lemma selection is completed before the activation of any phonological information (see Levelt 1989 and Roelofs 1997, for example), other researchers have found indications that this may not be so (e.g. Peterson & Savoy 1998, Cutting & Ferreira

1999). In light of that newfound evidence, Dell et al. (1997) went on to develop a new model of lexical access which preserved earlier models' distinctions between lemma selection and phonological encoding but denied that those stages were modular.

For Dell et al. (1997), lexical knowledge is embedded in a network of three layers: a semantic layer, a word (or lemma) layer and a phonological layer as in Figure 5 where each unit in the word layer is connected bidirectionally to their semantic and phonological features allowing activation to spread in a top-down fashion from semantic units to word units to phonemic units and bottom-up from phonemic to word to semantic units. Being nonmodular, Dell et al.'s (1997) later stages may begin processing before the earlier ones have finished (cascading) and processing at the later levels may influence that at earlier ones (feedback):

Figure 5. An illustration of a lexical network in the two-step lexical access model of word retrieval (Dell et al. 1997: 805)
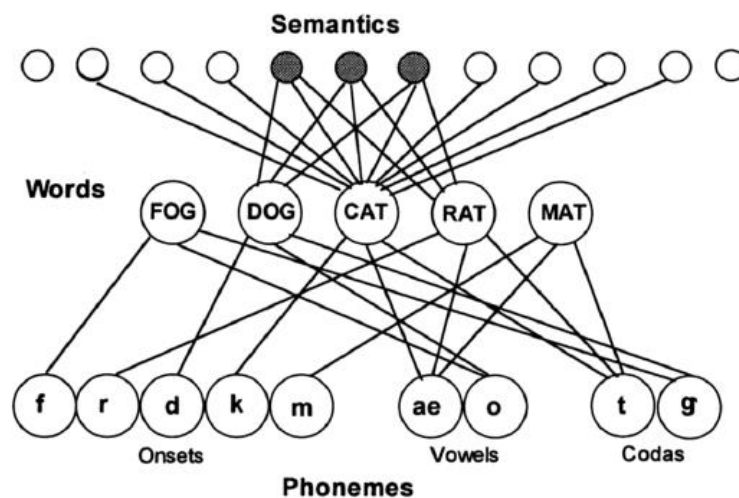
Dell et al. (1997) have actually been known to use their model to explain the error patterns one might observe in aphasic people with anomia. They argue that two steps are involved in the word retrieval process, lemma selection and phonological encoding, and that breakdowns at each can give rise to the speech errors we see in anomia.

In a scenario where a picture of a cat is presented to a person, they say, visual processes outside of the model will immediately identify that the concept illustrated by the picture is a cat and activate the semantic features of that concept. This activation then spreads through bidirectional excitatory connections for a fixed number of time steps according to a linear activation function until all three network levels are active, and in addition to the target word unit CAT, semantic neighbours such as DOG can become activated through shared semantic nodes, as can units such as MAT, CAT and CAN which, as phonological neighbours, can receive cascading activation from phonemes shared with the target. This first stage of the word retrieval process concludes with the most highly activated word being selected for retrieval, but this process is not always as straightforward as one might think, they point out. A semantic or phonological neighbour of the target word may well be selected if their activation levels exceed that of the target. This is an all too common occurrence in people with anomia.

This ties in with an explanation of anomia with the transmission deficit (TD) model in terms of the strength of the network connections that transmit activation to the representations involved in the retrieval of words (MacKay & Burke 1990, Burke et al. 1991). According to MacKay & Burke (1990) and Burke et al. (1991), semantic anomia occurs when the strength of connections among semantic representations become too weak through damage to transmit enough activation to the target word's semantic representation in the mental lexicon for it to be

retrieved, while damage to connections among phonological representations can result in phonological anomia by making them too weak to transmit sufficient activation to target words' phonological representations for their retrieval.

Phonological errors can also occur in the second stage of the word retrieval process, Dell et al. (1997) argue. In this stage, a chosen word is activated and activation spreads for another fixed number of time steps. As is the case for lemma selection, activation spreads both upwards and downwards during the phonological encoding process meaning that nodes other than the ones directly connected to CAT can be activated. The most highly activated phonemes are then selected for retrieval and errors occur when one or more phonemes are more active than those of the target word. This can result in real word phonological errors such as *mat*, *sat*, and *can* for *cat* and even nonword phonological errors such as *lat*.

Some variables have been said to influence how easily words are accessed in the mental lexicon (Carroll 2008). Among these are lexical factors such as typicality and frequency but perhaps the biggest influence on lexical access is a contextual factor known as priming. It is widely agreed that the processing of lexical items is affected by their preceding context; that people will tend to process a word more quickly and accurately when they have previously been exposed to a similar word (Goldinger et al. 1989, Marslen-Wilson & Zwitserlood 1989, Slowiaczek & Hamburger 1992). Word retrieval experiments for example have shown that people will tend to retrieve a word like *boy* more quickly if they have previously been exposed to a word like *girl* as opposed to something unrelated due to their similarity in meaning (semantic priming) and a word like *orgasm* more quickly if they have been previously exposed to a word like *organism* than something unrelated due to their similarity in sound (phonological priming) (McDonough & Trofimovich 2008).

In fact, it is precisely because phonological priming is known to affect lexical access that phonemic cueing tasks (which make use of it) are used in the treatment of phonological anomia, but how exactly does phonological priming occur and how can it improve word retrieval in phonological anomics, one might wonder? Sections 5 and 6 address those very questions.

5. Phonological priming and its influence on lexical access
5.1. Phonological priming in spoken languages

Phonological priming of spoken words refers to the improved recognition or retrieval of targets when they are preceded by primes that resemble them phonologically by sharing one or more of their constituent phonemes (James & Burke 2000). In phonological priming experiments, participants are typically presented with two words (a prime and a target) and asked to perform a task on the second to determine the effect hearing a prime has on the processing of a target (Pitt & Shoaf 2002).

In 1987, Slowiaczek et al. reported the results of three phonological priming experiments. In each of their three experiments, subjects were tasked with accurately identifying the target words before and after the provision of a prime. In the first experiment, the primes and targets were all real words and their shared phonemes appeared word-initially. Facilitatory effects were found for all of the primes. In the second experiment, nonwords were used as primes and similar effects were observed. Finally, in the third experiment, real word primes and targets were used once more, but rather than them being in word-initial positions, the shared phonemes were in word-final positions. Facilitatory effects were observed in the third experiment as well.

Slowiaczek et al. (1987) explained their results in terms of McClelland & Rumelhart's (1981) interactive-activation model of word recognition which like Dell et al.'s (1997) two-stage model of word retrieval assumes that phonemes and words exist on separate levels of

representation with excitatory activation passing between them. They argued that when a phonological prime is activated in the lexicon, the nodes for its phonemes are too and spread activation upwards through phoneme-to-word links to all other words that share those phonemes. These words remain partially active for a brief period, they claimed, allowing successful retrieval of the target word to be accomplished more easily than it would have been had a number of candidates for it to be selected from not been activated.

The results of other studies in the phonological priming literature would suggest, however, that effects largely depend on the extent to which primes' and targets' segments overlap. When primes overlap with targets by a small number of segments as was the case in Slowiaczek et al.'s (1987) study, facilitatory effects tend to be reported (e.g. Goldinger et al. 1992, Slowiaczek & Hamburger 1992, Hamburger & Slowiaczek 1996, 1999, Goldinger 1999, Spinelli et al. 2001) but when too many segments overlap inhibition of processing can be observed (Slowiaczek & Hamburger 1992, Hamburger & Slowiaczek 1996, 1999, Dufour & Peereman 2003).

Slowiaczek & Hamburger (1992), for example, conducted a priming experiment on targets preceded by primes that had no phonemes in common and therefore did not bear any phonological similarity to each other (e.g. clump (prime) and green (target)), primes that shared one phoneme with the target (e.g. goals (prime) and green (target)), primes that shared two phonemes (e.g. grope (prime) and green (target)) and primes that shared three (e.g. grief (prime) and green (target)) (Dufour & Peereman 2003). They found that response times were faster when primes shared one phoneme with the target than when there was no phonological similarity, but that the facilitatory effect decreased as the degree of phonemic overlap increased. This pattern of results led the researchers to conclude that too much similarity between a prime and a target can cause competition between the two during the word retrieval process and inhibit the target word's retrieval (Dufour & Peereman 2003). They argued that if there is a large phonemic overlap between a prime and a target, when the target is processed its prime is easily reactivated and acts as a strong competitor for retrieval, slowing down (and occasionally preventing) retrieval of the target word. When there is only a single phoneme overlap, however, the prime is not as easily reactivated and so does not interfere with the processing of the target (Dufour & Peereman 2003).

This can explain why priming with word-initial phonemes of targets in phonemic cueing treatments for phonological anomia can improve word retrieval but are not known to inhibit it. Because the primes used in phonemic cueing tasks are phonemes, not words, they cannot act as competitors for retrieval (except perhaps in cases where a single phoneme may be analysed as a word in its own right like /aɪ/ as the personal pronoun I). In Section 6 we will take a look at these phonemic cueing techniques in more detail before making some suggestions for their adaption for use with signed language users.

5.2. Phonological priming in signed languages

Most theories of phonological priming have been based on studies of spoken languages with relatively few examining how phonological priming takes place in the mental lexicon of signed language users, but this is not very surprising, really, given that until fairly recently most phonological research has concentrated on spoken languages rather than on signed ones.

There is strong evidence to suggest, however, that the same phonological system that underlies spoken languages underlies signed languages as well (Sandler & Lillo-Martin 2006), and if it is the case that their phonological systems are the very same, it should follow that phonological processes occur in each modality in the same way. Parallels were first drawn between the phonologies of spoken and signed languages by Stokoe in 1960 who classified signs according to the shape of the hand, the location of the hand in relation to the body and the

movement of the hand through the signing space. He proposed that signs are constructed from a limited set of formational elements drawn from three articulatory parameters: handshape, movement and location (Stokoe 1960, Stokoe et al. 1965) which are formally analogous to the three phonological parameters of speech production: voicing, place of articulation and manner of articulation (Chomsky & Halle 1968). Just as the voicing, place of articulation and manner of articulation parameters of spoken languages combine to produce phonemes, he said, the handshapes, movements and locations of signed languages come together to create the smaller phonological units of signed language phonology, cheremes (Stokoe 1960).

Some of the most commonly used handshapes in signed languages are FLAT HAND, FIST HAND and CUPPED HAND (Ottenheimer 2012). Rather self explanatorily, in a FLAT handshape the hand is flat, in a FIST handshape the hand is rolled into a fist and in a CUPPED handshape the hand is curved as in Figure 6 in which a FIST handshape, FLAT handshape and a CUPPED handshape are labelled a, b and c:

Figure 6. Some handshapes of a signed language (adapted from Ottenheimer 2012: 123)

a   b   c
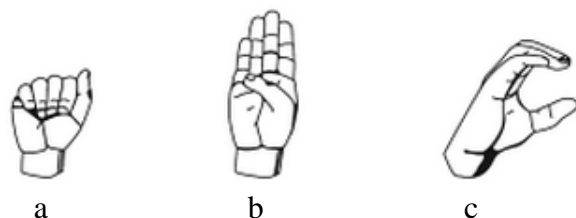
Any one of these handshapes can then be combined with a movement (such as UP, DOWN, TOWARD THE BODY, AWAY FROM THE BODY or ACROSS THE BODY) and a location (such as FACE, HEAD or UPPER BODY) to form a chereme, which can together combine with other cheremes to make up a sign.

Psycholinguistic studies in more recent years have gone on to explore whether the phonological priming that occurs in spoken languages can occur in signed languages as well (Carreiras et al. 2007). Corina (2000), for instance, conducted a series of experiments to investigate whether phonological parameter values of signs could be used for priming. Since phonological priming would be indicated by faster response times when a sign was preceded by a phonologically related sign, compared to an unrelated sign, Corina (2000) used a lexical decision task where deaf signers would view two sign stimuli in succession and be asked to decide whether the second sign was a true sign or a nonsign. The sign pairs either shared the same movement, the same location, or were phonologically unrelated (i.e. shared no phonological parameter values). Corina (2000) found no evidence for priming with either movement or location, which contrasted with results from an earlier experiment she conducted with Emmorey in 1993 in which lexical decision tasks revealed facilitatory effects when signs shared the same movement as the stimulus and inhibitory effects for lexical access when they shared the same location. When pairs of signs and stimuli shared a common handshape neither facilitation nor inhibition was found (Corina & Emmorey 1993).

Corina & Hildebrandt (2002) investigated priming in signed languages with the movement and location parameters, but found no evidence to suggest that priming occurred with either of the two. Dye & Shih (2006), on the other hand, have reported data from users of BSL which showed evidence of facilitatory effects for both movement *and* location.

Now, although the results of psycholinguistic studies examining phonological priming with handshape, movement and location in signed languages have been decidedly mixed, I would like to argue that this inconsistency may well be because these studies have only made

use of the loose phonological parameters identified by Stokoe in 1960 and by doing so have not actually accounted for phonological variation within those parameters.

We know from studies such as those of Sandler (1989), van der Hulst (1993), Brentari (1998) and van der Kooij (2002) of the variation within the phonological parameters of signed language. Sandler (1989), for example, pointed out that while handshapes may be flat, curved or fisted, the presence or absence of the fingers and thumb and their individual configurations when they are present can give rise to a number of handshapes that are flat, curved or fisted. Either one, some or all of the fingers may be present in a handshape and the thumb may be either present or absent. By the same token, the extremities that are present may be open (i.e. extended), closed (bent at the joint so that the tip of the finger or thumb makes contact with the palm), curved (extended but bent at the joint nonadjacent to the palm), bent (bent at the base joint but not so much so that the tip of the finger or thumb touches the palm) or even spread (fanned).

It would be naïve to think that signers would be insensitive to such variation. We certainly would not assume that speakers are unaffected by it, would we? Consider the following example from spoken English. The English word *dog* is monosyllabic, made up of the onset /d/ and the rime /ɒ g/. The word-initial phoneme /d/ in the onset is a voiced alveolar stop, so has the voicing parameter value [+ voice], the place of articulation parameter values [+ anterior] and [+ coronal] and the manner of articulation value [- continuant]. If the word *dog* was to act as a target in a phonological priming task, it would make more sense for a word (or non-word) that shares its word-initial phoneme to act as its prime, rather than a word that shares its voicing parameter value. This is because there is only one voiced alveolar stop in the English language, but there are a number of other voiced sounds, including the voiced obstruents [b], [g], [v], [z], [ʒ], [ð] and [dʒ] as well as the sonorants [m], [n], [ŋ], [l], [ɹ], [j], [w], [i], [e], [ɛ], [a], [ɑ], [ɔ], [o] and [u], to name but a few. A methodology that primes with phonemes would be likely to yield more reliable results than one that primes with parameter values as it could control better for variation.

Recall from Section 5.1 that in analogy to the phonemes of spoken words, Stokoe (1960) argued for the existence of cheremes, a discrete set of meaningless units that combine to produce a potentially infinite set of signs (Stokoe 1960). I would like to argue that if psycholinguistic studies on sign processing had been more like the studies on word processing and controlled for variation by selecting signs that share a chereme with the target as prime stimuli rather than signs that share a handshape, movement or location, they would have evinced more positive priming results. We would not expect two people with a bacterial infection to recover at the same rate if one was administered antibiotics and the other placebo pills, so why would we expect two different types of prime to have the same effects? It makes sense to predict that phonological priming will occur equivalently in each modality if the primes that are used in one modality are equivalent to the primes that are used in the other, but surely not if they are not. It would be interesting to see then whether signs that are selected as primes by virtue of them sharing a chereme with the target could facilitate word retrieval (or even whether cheremes alone could), as this could have serious consequences for the treatment of phonological anomia in aphasia.

6. Some suggestions for the treatment of phonological anomia in signed languages

The stimuli used in phonemic cueing tasks for the treatment of phonological anomia in spoken languages tends to be a set of line drawings that are presented to patients to be named on confrontation (Snodgrass & Vanderwart 1980). Patients are given a time limit within which they must name a target and if they are unable to respond or instead respond with an error, their response is recorded as incorrect (Nickels 2002, Greenwood et al. 2010). If a patient fails to

name the target correctly within the given time limit, a phonemic cue is provided by the clinician to prime it, which is usually the initial phoneme of the target (Hickin et al. 2001). Processing of the phonemic cue is said to activate its own underlying phonological representation as well as the phonological representations of all words beginning with it (which includes the target word, of course). Following the cue, the subject is then given another time limit within which they must try the target once more. The phonological representation of the target word is said to remain partially active at this stage, allowing phonological encoding to be accomplished more easily (Schriefers et al. 1990, Meyer & Schriefers 1991). Responses are recorded as incorrect or correct in the same manner as before and the percentage of correctly named items post-treatment is compared with the percentage of words correctly named pre-treatment to measure the treatment's efficacy.

In light of the existence of cheremes, I am keen to suggest that the phonemic approach to intervention for spoken word retrieval could easily be adapted for use with signed language users provided that cheremes can prime signs in the way that phonemes (their spoken language equivalent) can prime words. If this is so, the same methodology could be followed, but rather than clinicians providing patients with a phonemic cue when they have been unable to produce the target, they could offer a cheremic cue instead. The cheremic cue could then prime the target sign by activating its underlying phonological representation, making it more accessible and therefore easier to retrieve from memory.

7. Conclusion

To summarise, aphasia is an impairment of language that follows damage to the language area of the left cerebral hemisphere. Extensive research into aphasia in hearing individuals has identified robust patterns of breakdown across a number of spoken languages. We know from this research that the two most common forms of aphasia include Broca's aphasia (a non-fluent expressive aphasia) and Wernicke's aphasia (a fluent receptive aphasia) but that the symptom most commonly associated with all of the aphasias is anomia, an impairment of language which causes word retrieval difficulties. People with anomia may be unable to access words at all, or produce frequent word errors like semantic paraphasias such as calling a chair a *table* (Howard & Orchard-Lisle 1984, Kay & Ellis 1987, Hillis et al. 1990) or phonological paraphasias such as calling a cat a *tat* (Atkinson et al. 2005).

By contrast, there has been scant research into the impact of aphasia on sign language, but from the few studies that have been conducted we know that many of the patterns described above also occur in signed languages; the non-fluent/fluent dichotomy has been described by Poizner et al. (1987) for example and there have even been individuals reported to have sign finding problems as a symptom of their aphasia (Chiarello et al. 1982, Marshall et al. 2004). There are more specific parallels in the nature of aphasic errors, however. For example, in sign aphasias we observe semantic errors that are caused by semantic anomia where signers sign a semantically related word instead of the target word and phonological errors that are due to phonological anomia which entail signers making use of the wrong phonological parameter(s) or chereme(s) in a sign.

Up until recently, the emphasis of previous studies has been placed on the treatment of semantic anomia in spoken languages (Hickin et al. 2002). Treatment of phonological anomia in spoken languages has been studied too, but to a lesser extent, and treatment of phonological anomia in signed languages has been largely overlooked. Phonological anomia in spoken languages is often treated with phonemic cueing tasks in which phonemic cues are provided to patients to phonologically prime a target word and facilitate its retrieval, since phonological priming has been shown for some time now to improve lexical access (Marslen-Wilson & Welsh 1978, Goldinger et al. 1992, Coltheart et al. 2001, Carreiras et al. 2005).

In previous psycholinguistic studies of sign processing, however, the effects of priming on lexical access have been decidedly mixed. I argued that this may well be because in such studies, distinctions between signs are based on parameter values, and signs are selected as primes by virtue of them sharing the same handshape, movement or location as the target word. This means that variation within each of the parameters is neglected, and it would be naïve to believe that sign processing is insensitive to such variation. In psycholinguistic studies of word processing, however, prime stimuli are selected due to them sharing their word-initial phoneme with the target. Since Stokoe (1960) has argued for the existence of cheremes (the signed equivalent of phonemes), I suggested that if we better control for variation within parameters by selecting signs that share a chereme rather than a phonological parameter value with the target as prime stimuli, results would be more reliable and may well evince the positive priming effects that we see in spoken language studies. That said, I went on to suggest that if this was the case we could very easily adapt the current phonemic cueing tasks used to treat spoken language phonological anomia to make them suitable for signed language users. Rather than providing phonemic cues to improve lexical access, we could provide cheremic cues instead. It would be interesting to see whether this could help signers with phonological anomia with their word retrieval, though additional research to test the success of cheremic priming in signed languages would of course have to be conducted before we could attempt to use it clinically. It is about time deaf people had access to phonological therapy in the way that hearing people do, and given the rise of signed language phonology, there really is no excuse for them not to.

References

Albert, M., Goodglass, H., Helm, N., Rubens, A. & Alexander, M. (1981). *Clinical aspects of dysphagia.* New York: Springer-Verlag.

Allport, D. (1985). Distributed memory, modular subsystems and dysphasia. In Newman, S. & Epstein, R. (eds.) *Current perspectives in dysphasia*, 32-60. Edinburgh: Churchill Livingstone.

Atkinson, J., Marshall, J., Woll, B. & Thacker, A. (2005). Testing comprehension abilities in users of British Sign Language following CVA. *Brain and Language* 94, 233-248.

Bavelier, D., Corina, D., Jezzard, P., Clark, V., Karni, A., Lalwani, A., Rauschecker, J., Braun, A., Turner, R. & Neville, H. (1998). Hemispheric specialization for English and ASL: Left invariance – right variability. *Neuroreport* 9, 1537-1542.

Benson, D. (1979). *Aphasia, alexia and agraphia.* New York: Churchill Livingstone.

Blumstein, S. (1973). *A phonological investigation of aphasic speech.* The Hague: Mouton.

Bock, K. & Levelt, W. (1994). Language production: grammatical encoding. In Gernsbacher, M. (ed.) *Handbook of psycholinguistics*, 945-984. San Diego, CA: Academic Press.

Boyle, M. & Coelho, C. (1995). Application of semantic feature analysis as a treatment for aphasic dysnomia. *American Journal of Speech Language Pathology* 4, 94-98.

Braun, A., Guillemin, A., Hosey, L. & Varga, M. (2001). The neural organization of discourse: An H215O-PET study of narrative production in English and American Sign Language. *Brain* 124, 2028-2044.

Brentari, D. (1998). *A prosodic model of sign language phonology*. Cambridge, MA: Massachusetts Institute of Technology Press.

Brookshire, R. (1997). *An introduction to neurogenic communication disorders.* St. Louis, MO: Mosby.

Budd, M., Kortte, K., Cloutman, L., Newhart, M., Gottesman, R., Davis, C., Heidler-Gary, J., Seay, M. & Hillis, A. (2010). The nature of naming errors in primary progressive aphasia versus acute post-stroke aphasia. *Neuropsychology* 24, 581-589.

Burke, D., MacKay, D., Worthley, J. & Wade, E. (1991). On the tip of the tongue: What causes word finding problems in young and older adults? *Journal of Memory and Language* 30, 542-579.

Butterworth, B. (1989). Disorders of phonological encoding. *Cognition* 42, 261-286.

Byng, S. (1995). What is aphasia therapy? In Code, C. & Müller, D. (eds.) *The treatment of aphasia from theory to practice*, 3-17. London: Whurr.

Capek, C., Waters, D., Woll, B., MacSweeney, M., Brammer, M., McGuire, P., David, A. & Campbell, R. (2008). Hand and mouth: Cortical correlates of lexical processing in British Sign Language and speechreading English. *Journal of Cognitive Neuroscience* 20, 1220-1234.

Caramazza, A. (1997). How many levels of processing are there in lexical access? *Cognitive Neuropsychology* 14, 177-208.

Carlson, N. (2012). *Physiology of behavior.* 11th ed. Essex: Pearson Education Limited.

Caron, J. (1989). *Las regulaciones del discurso: Psicolingüística y pragmática del lenguaje* [The rules of discourse: Psycholinguistics and pragmatics in language]. Madrid: Gredos.

Carreiras, M., Dunabeitia, J. & Perea, M. (2007). Reading words, numbers and symbols. *Trends in Cognitive Science* 11, 454-455.

Carreiras, M., Ferrand, L., Grainger, J. & Perea, M. (2005). Sequential effects of phonological priming in visual word recognition. *Psychological Science* 16, 585-589.

Carroll, D. (2008). *Psychology of language.* 5th ed. Ontario: Thomson Wadsworth.

Catani, M. & Thiebaut de Schotten, M. (2008). A diffusion tensor imaging tractography atlas for virtual in vivo dissections. *Cortex* 44, 1105-1132.

Chiarello, C., Knight, R. & Mandel, M. (1982). Aphasia in a prelingually deaf woman. *Brain* 105, 29-51.

Chomsky, N. & Halle, M. (1968). *The sound pattern of English*. New York: Harper and Row.

Code, C. (1991). *The characteristics of aphasia.* Hove: Psychology Press.

Collins, A. & Loftus, E. (1975). A spreading activation theory of semantic processing. *Psychological Review* 82, 407-428.

Collins, A. & Quillian, M. (1969). Retrieval time from semantic memory. *Journal of Verbal Learning and Verbal Behavior* 8, 240-247.

Coltheart, M., Rastle, K., Perry, C., Langdon, R. & Ziegler, J. (2001). DRC: A dual route cascaded model of visual word recognition and reading aloud. *Psychological Review* 108, 204-256.

Corina, D. (1998). The processing of sign language: Evidence from aphasia. In Stemmer, B. & Whitaker, H. (eds.) *Handbook of neurolinguistics*, 313-329. New York: Academic Press.

Corina, D. (2000). Some observations regarding paraphasia in American Sign Language. In Emmorey, K. & Lane, H. (eds.) *The signs of language revisited: An anthology to honor Ursula Bellugi and Edward Klima*, 493-507. Mahwa, NJ: Lawrence Erlbaum Associates.

Corina, D. & Emmorey, K. (1993). Lexical priming in American Sign Language. Paper presented at the 34th annual meeting of the Psychonomics Society. Washington.

Corina, D. & Hildebrandt, U. (2002). Psycholinguistic investigations of phonological structure in American Sign Language. In Meier, R., Cormier, K. & Quinto-Pozos, D. (eds.) *Modality and structure in signed and spoken languages*, 88-111. Cambridge: Cambridge University Press.

Corina, D., McBurney, S., Dodrill, C., Hinshaw, K., Brinkley, J. & Ojemann, G. (1999). Functional roles of Broca's area and SMG: Evidence from cortical stimulation mapping in a deaf signer. *NeuroImage* 10, 570-581.

Corina, D., Poizner, H., Bellugi, U., Feinberg, T., Dowd, D. & O'Grady Batch, L. (1992). Dissociation between linguistic and nonlinguistic gestural systems: A case of compositionality. *Brain and Language* 43, 414-447.

Corina, D., San Hose-Robertson, L., Guillemin, A., High, J. & Braun, A. (2003). Language lateralization in a bimanual language. *Journal of Cognitive Neuroscience* 15, 718-730.

Corina, D., Chiu, Y., Knapp, H., Greenwald, R., San Jose-Robertson, L. & Braun, A. (2007). Neural correlates of human action observation in hearing and deaf subjects. *Brain Research* 1152, 111-129.

Crofts, B., Nickels, L., Makin, K., Taylor, C. & Moses, M. (2004). A comparison of the effectiveness of 'semantic' and 'phonological' tasks in the facilitation of word production in aphasia. In Murdock, B., Goozee, J., Whelan, B. & Docking, K. (eds.) *Proceedings of the 26th world congress of the international association of logopedics and phoniatrics*, 1-5. Brisbane, Australia.

Cutting, J. & Ferreira, V. (1999). Semantic and phonological information flow in the production lexicon. *Journal of Experimental Psychology: Learning, Memory and Cognition* 25, 318-344.

Damasio, A. (1992). Aphasia. *Medical Progress* 326, 531-539.

Damasio, H. (1989). Neuroimaging contributions to the understanding of aphasia. In Boller, F. & Grafman, J. (eds) *Handbook of neuropsychology*, 3-46. Amsterdam: Elsevier.

Davis, A. & Pring, T. (1991). Therapy for word finding deficits: More on the effects of semantic and phonological approaches to treatment with dysphasic patients. *Neuropsychological Rehabilitation* 1, 135-145.

Dell, G. (1986). A spreading-activation theory of retrieval in sentence production. *Psychological Review* 93, 283-321.

Dell, G., Schwartz, M., Martin, N., Saffran, E. & Gagnon, D. (2007). Lexical access in aphasic and nonaphasic speakers. *Psychological Review* 104, 801-838.

Devinsky, O. & D'Esposito, M. (2004). *Neurology of cognitive and behavioral disorders.* Oxford: Oxford University Press.

Douglas, E. & Richardson, J. (1959). Aphasia in a congenital deaf-mute. *Brain* 83, 68-80.

Drew, R. & Thompson, C. (1999). Model-based semantic treatment for naming deficits in aphasia. *Journal of Speech, Language and Hearing Research* 42, 972-989.

Dufour, S. & Peereman, R. (2003). Lexical competition in phonological priming: Assessing the role of phonological match and mismatch lengths between primes and targets. *Memory and Cognition* 31, 1271-1283.

Dye, M. & Shih, S. (2006). Phonological priming in British Sign Language. In Goldstein, L., Whalen, D. & Best, C. (eds.) *Papers in laboratory phonology*, 241-264. Berlin: Mouton de Gruyter.

Ellis, A., Kay, J. & Franklin, S. (1992). Anomia: Differentiating between semantic and phonological deficits. In Margolin, D. (ed.) *Cognitive neuropsychology in clinical practice*, 207-228. New York: Oxford University Press.

Emmorey, K., McCullough, S. & Brentari, D. (2003). Categorical perception in American Sign Language. *Language and Cognitive Processes* 18, 21-45.

Emmorey, K., Mehta, S. & Grabowski, T. (2007). The neural correlates of sign versus word production. *NeuroImage* 36, 202-208.

Fay, D. & Cutler, A. (1977). Malapropisms and the structure of the mental lexicon. *Linguistic Inquiry* 8, 505-520.

Fick, D., Agostini, J. & Inouye, S. (2002). Delirium superimposed on dementia: A systematic review. *Journal of the American Geriatrics Society* 50, 1723-1732.

Fitzgerald, M. (1996). *Neuroanatomy: Basic and clinical.* Philadelphia: Saunders.

Fridriksson, J., Baker, J. & Moser, D. (2009). Cortical mapping of naming errors in aphasia. *Human Brain Mapping* 30, 2487-2498.

Fromkin, V. (1971). The non-anomalous of anomalous utterances. *Language* 47, 27-52.

Garrett, M. (1975). The analysis of sentence production. In Bower, G. (ed.) *The psychology of learning and motivation: Advances in research and theory*, 133-177. New York: Academic Press.

Glaser, W. & Glaser, M. (1989). Context effects on Stroop-like word and picture processing. *Journal of Experimental Psychology: General* 118, 13-42.

Goldinger, S. (1999). Only the shadower knows: Comment on Hamburger & Slowiaczek (1996). *Psychonomic Bulletin and Review* 6, 347-351.

Goldinger, S., Luce, P. & Pisoni, D. (1989). Priming lexical neighbors of spoken words: Effects of competition and inhibition. *Journal of Memory and Language* 28, 501-518.

Goldinger, S., Luce, P., Pisoni, D. & Marcario, J. (1992). Form-based priming in spoken word recognition: The roles of competition and bias. *Journal of Experimental Psychology: Learning, Memory and Cognition* 18, 1210-1238.

Goodglass, H. & Wingfield, A. (1997). *Anomia: Neuroanatomical and cognitive correlates.* Burlington: Elsevier Science.

Greenwood, A., Grassly, J., Hickin, J. & Best, W. (2010). Phonological and orthographic cueing therapy: A case of generalised improvement. *Aphasiology* 24, 991-1016.

Hamburger, M. & Slowiaczek, L. (1996). Phonological priming reflects lexical competition. *Psychonomic Bulletin & Review* 3, 520-525.

Hamburger, M. & Slowiaczek, L. (1999). On the role of bias in dissociated phonological priming effects: A reply to Goldinger (1999). *Psychonomic Bulletin & Review* 6, 352-355.

Hamilton, R., Sanders, L., Benson, J., Faseyitan, O., Norise, C., Naeser, M., Martin, P. & Coslett, H. (2010). Stimulating conversation: Enhancement of elicited propositional speech in a patient with chronic nonfluent aphasia following transcranial magnetic stimulation. *Brain and Language* 113, 45-50.

Helm-Estabrooks, N., Albert, M. & Nicholas, M. (2014). *Manual of aphasia and aphasia therapy.* 3rd ed. Austin, TX: PRO-ED Inc.

Hickin, J., Best, W., Herbert, R., Howard, D. & Osborne, F. (2001). Treatment of word retrieval in aphasia: Generalisation to conversational speech. *International Journal of Language and Communication Disorders* 36, 13-18.

Hickin, J., Best, W., Herbert, R., Howard, D. & Osborne, F. (2002). Phonological therapy for word-finding difficulties: A re-evaluation. *Aphasiology* 16, 981-999.

Hickok, G., Bellugi, U. & Klima, E. (1998). The neural organization of language: Evidence from sign language aphasia. *Trends in Cognitive Sciences* 2, 129-136.

Hickok, G., Houde, J. & Rong, F. (2011). Sesorimotor intergration in speech processing: computational basis and neural organization. *Neuron* 69, 407-422.

Hickok, G., Love-Geffen, T. & Klima, E. (2002). Role of the left hemisphere in sign language comprehension. *Brain and Language* 82, 167-178.

Hillis, A. & Caramazza, A. (1994). Theories of lexical processing and theories of rehabilitation. In Humphreys, G., & Riddoch, M. (eds.) *Cognitive neuropsychology and cognitive rehabilitation* 449-482. Hillsdale, NJ: Lawrence Erlbaum Associates.

Hillis, A., Rapp, B., Romani, C., & Caramazza, A. (1990). Selective impairment of semantics in lexical processing. *Cognitive Neuropsychology* 7, 191-244.

Howard, D. (2000). Cognitive neuropsychology and aphasia therapy: The case of word retrieval. In Papathanasiou, I. (ed.) *Acquired Neurogenic Communication Disorders: A Clinical Perspective*, 76-99. London: Whurr.

Howard, D. & Orchard-Lisle, V. (1984). On the origin of semantic errors in naming: Evidence from the case of a global aphasic. *Cognitive Neuropsychology* 1, 163-190.

Howard, D., Patterson, K., Franklin, S., Orchard-Lisle, V. & Morton, J. (1985). Treatment of word retrieval deficits in aphasia: A comparison of two therapy methods. *Brain* 108, 817-829.

James, L. & Burke, D. (2000). Phonological priming effects on word retrieval and tip-of-the-tongue experiences in young and older adults. *Journal of Experimental Psychology: Learning, Memory and Cognition* 26, 1378-1391.

Kay, J. & Ellis, A. (1987). A cognitive neuropsychological case study of anomia. Implications for psychological models of word retrieval. *Brain* 110, 613-629.

Kay, J., Lesser, R. & Coltheart, M. (1992). *PALPA: Psycholinguistic assessments of language processing in aphasia.* Hove: Lawrence Erlbaum.

Kearns, K. (2005). Broca's aphasia. In LaPointe, L. (ed.) *Aphasia and related neurogenic language disorders,* 117-141. New York: Thieme.

Keller, S., Crow, T., Foundas, A., Amunts, K. & Roberts, N. (2009). Broca's area: Nomenclature, anatomy, typology and asymmetry. *Brain and Language* 109, 29-48.

Kertesz, A. (1982). *Western aphasia battery.* San Antonio: The Psychological Corporation.

Kertesz, A., Harlock, W. & Coates, R. (1979). Computer tomographic localization, lesion size, and prognosis in aphasia and nonverbal impairment. *Brain and Language* 8, 34-50.

Kempen, G. & Huijbers, P. (1983). The lexicalization process in sentence production and naming: Indirect election of words. *Cognition* 14, 185-209.

Kiran, S. & Thompson, C. (2003). The role of semantic complexity in treatment of naming deficits: Training semantic categories in fluent aphasia by controlling exemplar typicality. *Journal of Speech, Language and Hearing Research* 46, 608-622.

Klima, E. & Bellugi, U. (1979). *The signs of language.* Cambridge: Harvard University Press.

Kolb, B. & Whishaw, I. (2003). *Fundamentals of human neuropsychology.* New York: Worth Publishers.

La Heij, W., Dirkx, J. & Kramer, P. (1990). Categorical interference and associative priming in picture naming. *British Journal of Psychology* 81, 511-525.

Lecours, A. & Lhermitte, F. (1969). Phonemic paraphasias: Linguistic structures and tentative hypotheses. *Cortex* 5, 193-228.

Leischner, A. (1943). Die 'aphasie' der taubstummen [Aphasia of the deaf]. *Archiv fur Psychiatrie und Nervenkrankheiten* 115, 469-548.

Levelt, W. (1989). *Speaking: From intention to articulation.* Cambridge, MA: Massachusetts Institute of Technology Press.

Levelt, W., Schriefers, H., Vorberg, D., Meyer, A., Pechmann, T. & Havinga, J. (1991). The time course of lexical access in speech production: A study of picture naming. *Psychological Review* 98, 122-142.

Mackay, D. & Burke, D. (1990). Cognition and aging: A theory of new learning and the use of old connections. In Hess, T. (ed.) *Aging and cognition: Knowledge organization and utilization,* 213-263. Amsterdam: North Holland.

MacSweeney, M., Capek, C., Campbell, R. & Woll, B. (2008). The signing brain: The neurobiology of sign language. *Trends in Cognitive Sciences* 12, 432-440.

MacSweeney, M., Woll, B., Campbell, R., McGuire, P., David, A., Williams, S., Suckling, J., Calvert, A. & Brammer, M. (2002). Neural systems underlying British Sign Language and audio-visual English processing in native users. *Brain* 125, 1583-1593.

Marcel, A. (1983). Conscious and unconscious perception: Experiments on visual masking and word recognition. *Cognitive Psychology* 15, 197-237.

Marshall, J., Atkinson, J., Smulovitch, E., Thacker, A. & Woll, B. (2004). Aphasia in a user of British Sign Language: Dissociation between sign and gesture. *Cognitive Neuropsychology* 21, 537-554.

Marshall, J., Pound, C., White-Thomson, M. & Pring, T. (1990). The use of picture/word matching tasks to assist word-retrieval in aphasic patients. *Aphasiology* 4, 167-184.

Marshall, R., Lazar, R. & Mohr, J. (1998). Aphasia. *Medical Update for Psychiatrists* 3, 132-138.

Marslen-Wilson, W. & Welsh, A. (1978). Processing interactions and lexical access during word recognition in continuous speech. *Cognitive Psychology* 10, 29-63.

Marslen-Wilson, W. & Zwitserlood, P. (1989). Accessing spoken words: The importance of word onsets. *Journal of Experimental Psychology: Human Perception and Performance* 15, 576-585.

Manasco, H. (2014). *Introduction to neurogenic communication disorders.* Burlington, MA: Jones & Bartlett Learning.

McClelland, J. & Rumelhart, D. (1981). An interactive activation model of context effects in letter perception: Part I. An account of basic findings. *Psychological Review* 88, 375-407.

McDonough, K. & Trofimovich, P. (2008). *Using priming methods in second language research.* New York: Routledge.

Meyer, A. & Damien, M. (2007). Activation of distractor names in the picture-picture interference paradigm. *Memory and Cognition* 35, 494-503.

Meyer, A. & Schriefers, H. (1991). Phonological experimentation in picture-word interference experiments: Effects of stimulus onset asynchrony and types of interfering. *Journal of Experimental Psychology: Learning, Memory and Cognition* 17, 1146-1160.

Miceli, G., Amitrano, A., Capasso, R. & Caramazza, A. (1996). The treatment of anomia resulting from output lexical damage: Analysis of two cases. *Brain and Language* 52, 150-174.

Naeser, M. & Hayward, R. (1978). Lesion localization in aphasia with cranial computed tomography and the Boston Diagnostic Aphasia Exam. *Neurology* 28, 545-551.

Neely, J. (1977). Semantic priming and retrieval from lexical memory: Role of inhibitionless spreading activation and limited-capacity attention. *Journal of Experimental Psychology: General* 106, 226-254.

Neely, J. (1991). Semantic priming effects in visual word recognition: A selective review of current findings and theories. In Besner, D. & Humphreys, G. (eds.) *Basic processes in reading: Visual word recognition*, 264-336. Hillsdale, NJ: Erlbaum.

Nettleton, J. & Lesser, R. (1991). Therapy for naming difficulties in aphasia: Application of a cognitive neuropsychological model. *Journal of Neurolinguistics* 6, 139-157.

Neville, H., Bavelier, D., Corina, D., Rauschecker, J., Karni, A., Lalwani, A., Braun, A., Clark, V., Jezzard, P. & Turner, P. (1998). Cerebral organization for language in deaf and hearing subjects: Biological constraints and effects of experience. *Proceedings of the National Academy of Sciences* 95, 922-929.

Newman, A., Bavelier, D., Corina, D., Jezzard, P. & Neville, H. (2002). A critical period for right hemisphere recruitment in American Sign Language processing. *Nature Neuroscience* 5, 76-80.

Nickels, L. (2002). Improving word finding: Practice makes (closer to) perfect? *Aphasiology* 16, 1047-1060.

Nickels, L. & Best, W. (1996). Therapy for naming deficits (part I): Principles, puzzles and progress. *Aphasiology* 10, 21-47.

Ottenheimer, H. (2012). *The anthropology of language: An introduction to linguistic anthropology.* Belmont, CA: Cengage Learning.

Peterson, R. & Savoy, P. (1998). Lexical selection and phonological encoding during language production: Evidence for cascaded processing. *Journal of Experimental Psychology: Learning, Memory and Cognition* 24, 539-557.

Pitt, M. & Shoaf, L. (2002). Revisiting bias effects in word-initial phonological priming. *Journal of Experimental Psychology: Human Perception and Performance* 28, 1120-1130.

Poizner, H., Klima, E. & Bellugi, U. (1987). *What the hands reveal about the brain.* Cambridge, MA: Massachusetts Institute of Technology Press.

Posner, M. & Snyder, C. (1975). Facilitation and inhibition in the processing of signals. In Rabbitt, P. & Dornic, S. (eds.) *Attention and Performance V*, 669-682. New York: Academic Press.

Potter, M. & Faulconer, B. (1975). Time to understand pictures and words. *Nature* 253, 437-438.

Pring, T., White-Thomson, M., Pound, C., Marshall, J. & Davis, A. (1990). Picture/word matching tasks and word retrieval: Some follow up data and second thoughts. *Aphasiology* 4, 479-483.

Raymer, A. & Ellsworth, T. (2002). Response to contrasting verb retrieval treatments: A case study. *Aphasiology* 16, 1031-1045.

Raymer, S., Thompson, C., Jacobs, B. & Le Grand, H. (1993). Phonological treatment of naming deficits in aphasia: Model-based generalization analysis. *Aphasiology* 7, 27-54.

Roelofs, A. (1992). A spreading-activation theory of lemma retrieval in speaking. *Cognition* 42, 107-142.

Roelofs, A. (1997). The WEAVER model of word-form encoding in speech production. *Cognition* 64, 249-284.

Sandler, W. (1989). *Phonological representation of the sign: Linearity and non-linearity in American Sign Language*. Dordrecht: Foris.

Sandler, W. & Lillo-Martin, D. (2006). *Sign language and linguistic universals.* Cambridge: Cambridge University Press.

Schriefers, H., Meyer, A. & Levelt, W. (1990). Exploring the time course of lexical access in language production: Picture-word interference studies. *Journal of Memory and Language* 29, 86-102.

Senghas, A., Kita, S. & Özyürek, A. (2004). Children creating core properties of language: Evidence from an emerging sign language in Nicaragua. *Science* 305, 1779-1782.

Slowiaczek, L. & Hamburger, M. (1992). Prelexical facilitation and lexical interference in auditory word recognition. *Journal of Experimental Psychology: Learning, Memory and Cognition* 18, 1239-1250.

Slowiaczek, L., Nusbaum, H. & Pisoni, D. (1987). Phonological priming in auditory word recognition. *Journal of Experimental Psychology: Learning, Memory and Cognition* 17, 459-477.

Smith, E., Shoben, E. & Rips, L. (1974). Structure and process in semantic memory: A featural model for semantic decisions. *Psychological Review* 81, 214-241.

Snodgrass, J. & Vanderwart, M. (1980). A standardized set of 260 pictures: Norms for name agreement, image agreement, familiarity and visual complexity. *Journal of Experimental Psychology: Human Learning and Memory* 6, 174-215.

Spinelli, E., Segui, J. & Radeau, M. (2001). Phonological priming in spoken word recognition with bisyllabic targets. *Language and Cognitive Processes* 16, 367-392.

Stokoe, W. (1960). *Sign language structure: An outline of the visual communication systems of the American deaf*. Revised 1978. Silver Spring, MD: Linstok Press.

Stokoe, W., Casterline, D. & Croneberg, C. (1965). *A dictionary of American sign languages on linguistic principles.* Washington, D.C.: Gallaudet College Press.

Tureen, L., Smolik, E. & Tritt, J. (1951). Aphasia in a deaf mute. *Neurology* 1, 237-289.

Van der Hulst, H. (1993). Units in the analysis of signs. *Phonology* 10, 209-241.

Van der Kooij, E. (2002). *Phonological categories in sign language of the Netherlands: The role of phonetic implementation and iconicity.* Utrecht: Landelijke Onderzoekschool Taalwetenschap.

Woll, B., Sutton-Spence, R. & Elton, F. (2001). Multilingualism, the global approach to sign languages. In Ceil, L. (ed.) *The sociolinguistics of sign language*, 8-32. Cambridge: Cambridge University Press.

Woollams, A., Cooper-Pye, E., Hodges, J. & Patterson, K. (2008). Anomia: A doubly typical signature of semantic dementia. *Neuropsychologica* 46, 2503-2514.

Vanessa McIntosh
School of English Literature, Language and Linguistics
Percy Building
Newcastle University
Newcastle upon Tyne
NE1 7RU
United Kingdom

v.m.mcintosh@newcastle.ac.uk

QAF IN MOSUL ARABIC: LEVELLING OR RESISTANCE?

ABDULKAREEM YASEEN
(Newcastle University)

Abstract

The literature on phonological variation and change has reported that a process of dialect levelling amongst Arabic varieties spoken in urban settings has been in operation. Iraqi Arabic dialects and those spoken in neighbouring areas that are akin to them are believed to be undergoing such a process as a result of different factors. This article assesses whether this process is occurring in the variety of Arabic spoken in Mosul, Iraq.

Notwithstanding various comments referring to the supplanting of Mosul Arabic (henceforth MA), this paper presents evidence that the variety appears to be resisting the levelling that is currently occurring in other parts of Iraq. The results show that [q], one of the traditional features of MA, is still robustly maintained by speakers of MA stratified by gender, age and social class. The article suggests that the social meaning attached to this sound and the type of relationships maintained within the community of MA are the best possible factors that can be adduced to explain its preservation since this particular sound defines, at least linguistically, the community of MA.

1. Introduction

Recent decades have seen considerable theoretical as well as methodological strides in the variationist research, notably in the realm of phonological variation and change. Labov's Martha's Vineyard study (1963) has sparked an upsurge of research interest in elucidating the multi-dimensional trajectories of phonological change through exploring inter- and intra-speaker variation (Foulkes & Docherty 1999). A key more recent offshoot of this strand of research is the theme of levelling. Levelling, or supralocalisation, is a phenomenon whereby natives of certain traditional dialects gradually abandon their traditional linguistic features, whether by choice or under duress, in favour of competing equivalents of a broader regional usage. Thus, features which make different dialects distinctive smooth out or completely vanish (Trudgill 1986: 98, Williams & Kerswill 1999: 149). Hinskens (1993: 11) defines levelling as a process in which the number of structures distinguishing a dialect from other varieties decreases, including those features socially perceived as prestigious and standard.

As a contact-induced process, levelling is seen as a process facilitated by *inter alia* increased mobility and immigration in a particular area. Dialect levelling (or more appropriate in this study, accent levelling) has become a focus of research and discussion in particular with reference to British English. Scholars such as Trudgill (1986), Kerswill (1996a) and Britain (1997) have extensively examined levelling as they tracked changes triggered by dialect contact, which has in turn been brought about by the social upheavals that have been taking place in Britain since the twentieth century.

Dialect levelling has also been a recurrent theme in the literature, the applicability of which is extending well beyond the Anglocentric sphere. A number of studies followed the British pattern to include many parts of Europe such as Norway (e.g. Kerswill 1996b, Haug Hilton 2010, Røyneland 2009) and France (e.g. Armstrong 2001, 2002, Temple 2001, Esch 2002, Hornsby 2002, Pooley 2002, Lodge 2004, Boughton 2005) as well as areas beyond Europe such as the Arab world (e.g. Gibson 2002, Manfredi 2012) to name but a few.

2. Iraqi Arabic dialects

Of the various varieties spoken in Iraq, Iraqis are known for speaking a nation-wide variety called Iraqi, also known as Mesopotamian, Arabic. Iraqi Arabic (henceforth IA) is a continuum of Arabic varieties spoken in the Mesopotamian basin of Iraq, which extends beyond Iraq's territory to include adjacent parts of neighbouring Syria, Turkey and Iran. IA is predominantly centralised in Iraq and is the mother language to over 80% of the Iraqi people (Peoples & Bailey 2011: 298).

Blanc (1964) established a bipartite linguistic classification of IA that has remained in fashion throughout much of the work on the spectrum of Arabic dialects spoken in the Mesopotamian area. He charted the division of the dialects of Mesopotamia into two main dialect groups, *qeltu* and *gilit*[8] dialects, coined from the realisation of the 1st person singular past tense of the verb 'say'. Jastrow (1978) classified the *qeltu* dialects into three further groups: the Tigris, Euphrates, and Anatolian group. The *qeltu* group includes Maslawi Arabic (henceforth MA), as well as ethnically-based dialects such as Christian Baghdad (CBA) and Jewish Baghdad Arabic (JBA) (See Diagram 1 below).

**Diagram 1. Iraqi Arabic dialect classification based on Blanc (1964) and Jastrow (2006)**
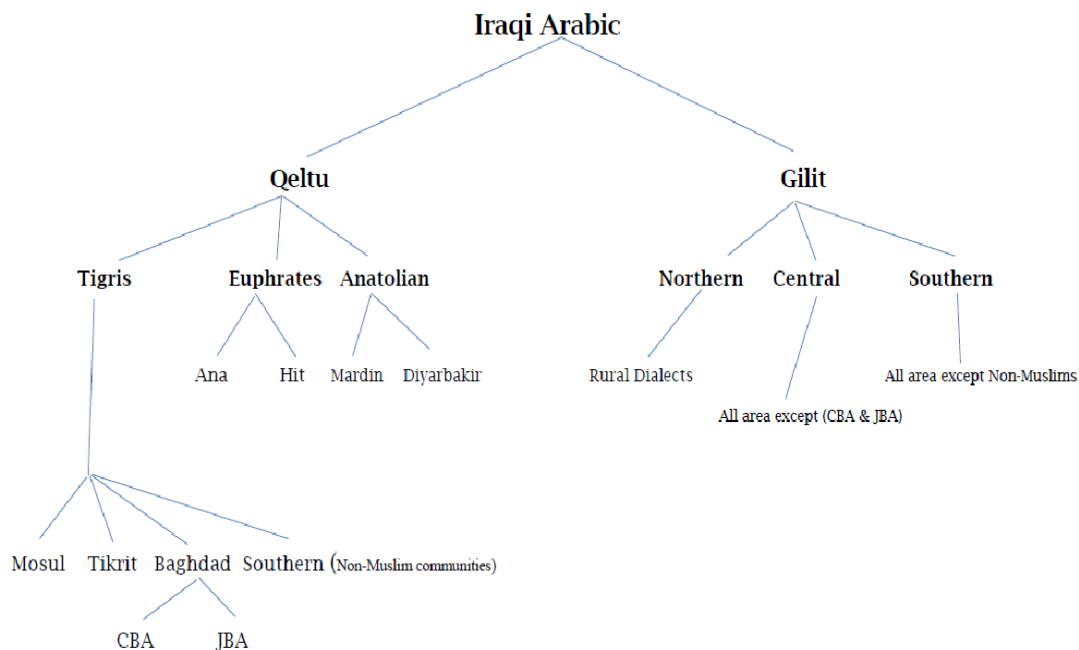
There is, in fact, no definite geographical feature dividing the two dialect areas and there are areas of variation where both may be heard. However, Holes (2007: 124) notes that an isogloss can be drawn between these two main Iraqi dialect groups (as illustrated in Figure 1); it runs approximately between the town of Falluja on the Euphrates River and the town of Samarra on the Tigris.

The *qeltu* dialects are spoken by Muslims north of Baghdad (e.g. Mosul and Tikrit) and by Christians and Jews in the whole dialect area which extends from Baghdad up to northern

---

[8] This is the nearest transliterated form of this term to how it is pronounced in Iraqi Arabic although other forms exist (e.g. *gelet* as in Abu Haidar 2007).

Iraq, northeastern Syria and southern Turkey. The majority *gilit* dialects are spoken in the rest of Iraq, and parts of Iran and are akin to Najdi Arabic, a variety spoken in northeastern Arabia (Biadsy et al. 2009: 55).

Figure 1. The distribution of *qeltu* and *gelet* dialects in Iraq and beyond (based on Collin 2009: 250)
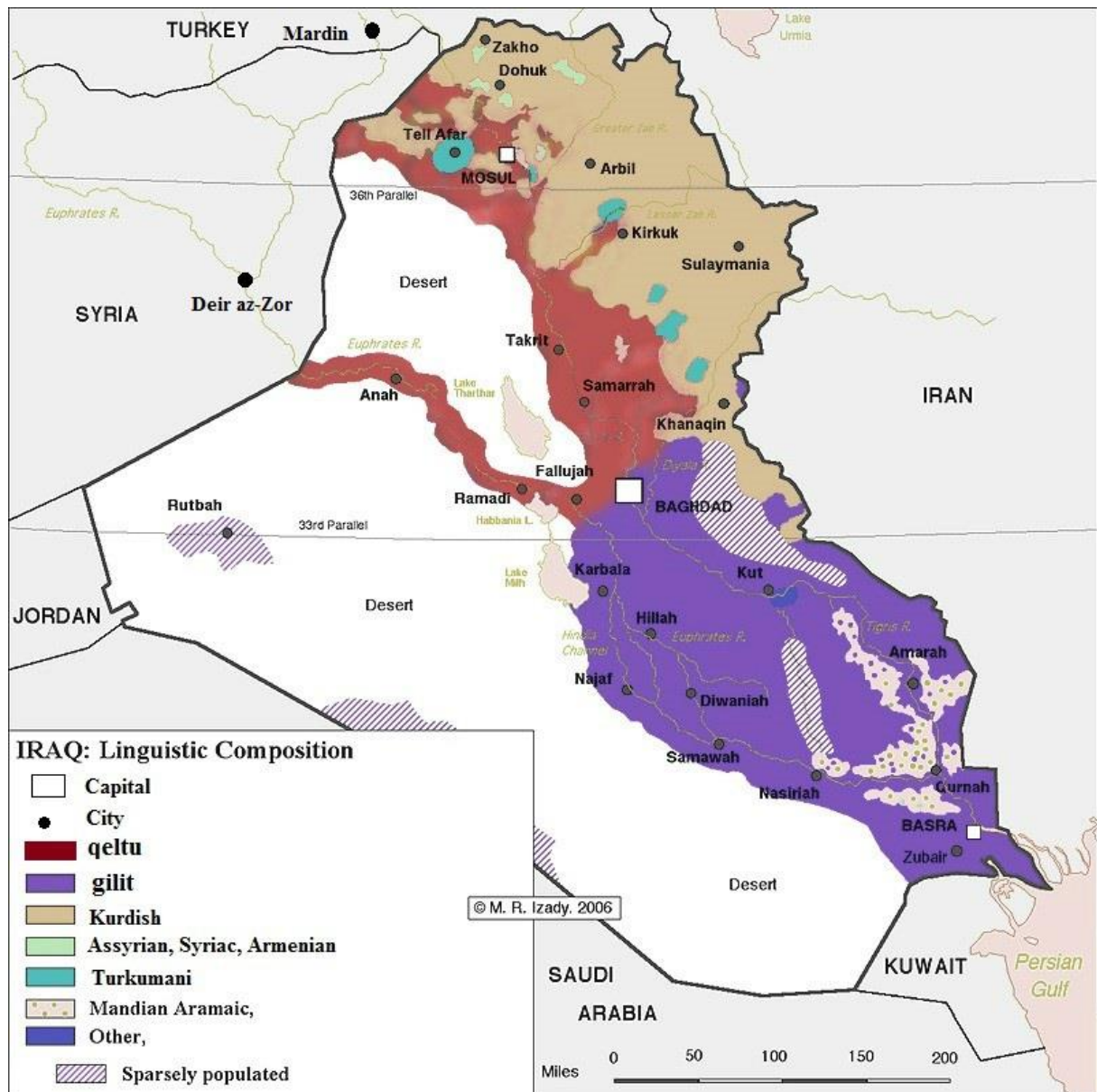
3. Mosul

The capital of Nineveh province, Mosul is a city in northern Iraq. It is Iraq's third largest city and the major metropolitan centre in the northern region of Iraq (Figure 2). Situated some 363 km northwest of Baghdad, Mosul is sandwiched in a hilly area between the mountains in the north and east and the Al-jazeera desert in the south and the west. The Tigris River divides the city into two parts. These two parts are locally known as *Sahel Al-ayman* 'Right bank' and *Sahel Al-ayasar* 'Left bank'. Both banks contain a sprawling web of historic as well as newly-built neighbourhoods.

Mosul's society is often depicted as a kaleidoscope of different ethnicities and religious groups that include Arabs, Christians, Kurds, Yazidis, Turkmen, Kaka'is and Shabaks. Corresponding to this variety of inhabitants is a number of languages ranging from Arabic, Neo-Aramaic, and Kurdish to Shabaki and Turkmen. Arabs constitute the majority of the inhabitants in Mosul and their areas are found on both banks of the city. The areas of the other ethnic groups can be found largely on the right bank. Although these ethnic groups have their own languages, they still speak Arabic, which acts as the lingua franca that binds the different population groups of Mosul into a whole.

Figure 2. Map showing the geographical location of Mosul in Iraq

3.1. Maslawi Arabic (MA)

MA is the Arabic variety spoken by the people of the city of Mosul in northern Iraq, 352 km north of Baghdad. Thanoon (2010) claims that the origin of the Mosul dialect can be traced back to the Arab tribe, *Bani Tamim*, who settled in Iraq in pre-Islamic times. MA is influenced by the languages of the ethnic groups of Mosul such as the Kurds, Turkmen, and Christians (Ibid.). Irrespective of their ethnic or religious backgrounds, all Maslawis speak MA with some differences. Blanc (1964: 164) notes that MA is akin to CBA in that both share noticeable similarities. Still, each variety has some exclusive linguistic features that distinguish one from the other.

4. Levelling in MA

Levelling is believed to be occurring in *qeltu* varieties, not least in MA, in that some of the defining linguistic features of the *qeltu* group are losing ground to supralocal ones (e.g. *gilit*). Although no previous surveys of dialect levelling have been carried out on MA, anecdotal evidence, however, can serve as an indicator of the occurrence of this process. Early remarks germane to the issue of levelling, albeit impressionistic in nature, can be found in Palva (1983: 101) and Abu-Haidar (2007) who claim that levelling is underway in the *qeltu* dialect group. The quotes below are separated by more than two decades, yet they both agree on the occurrence of levelling in *qeltu* speaking areas, including Mosul. Palva (1983: 101) states:

> Due to the radical change in the modern society, local dialects are today exposed to substantial linguistic interference. On the other hand, qeltu dialects represent almost everywhere a geographically recessive type. In Iraq and in the Syrian town Der iz-Zor they are losing ground to the dominant gilit dialects, and in Anatolia to Turkish and Kurdish. In many places in Anatolia the socio-religious minorities have already become extinct or have left the area, and many dialects will become extinct during the next generation.

A similar observation is also made by Collin (2009) whose personal communication with scholars like Farida Abu-Haidar and Clive Holes has led him to conclude that levelling tendencies are afoot in Mosul as well as other cities in northern Iraq and that the boundary between the dialects of Mosul and Baghdad is drifting further up to the north:

> Tikrit was once a qeltu-speaking town. Nowadays, the majority speak gelet. Mosul is going that way too. Kirkuk pre-1979 was predominantly Turkoman-speaking. By the time of the invasion in 2003, the Turkomans had become a minority and you could find all kinds of southern and central gelet varieties throughout Kirkuk. (Abu Haidar, p.c. cited in Collin (2009))

Collin goes further to refer to Abu Haidar's claim that the influence of Baghdad plays a major role in the current change in those areas given the status of MBA as a prestigious norm for the young to imitate. Capitals are typically depicted as modern, cultural centres whose dialects are nationally perceived as the prestigious varieties. Baghdad is no exception. Its dominant *gilit* variety is viewed as the nation-wide variety spoken by and intelligible to a large part of the Iraqi people. That is why this variety is diffusing to major population centres in the northwest of Iraq like Mosul (Ibid.). Al-damluji (2014) also claims that more "foreign" variants are gaining currency in the dialect of Mosul, which seems to be "fighting for survival" amid the influence of the Bedouin-type dialects of Iraq.

The alleged recessiveness of MA is also drawing public interest, receiving an airing in the public press, as it frequently appears in the local as well as national papers. For instance, Elyas (2013) argues that Maslawis are "fearing the extinction of their dialect" owing to reasons similar to those sketched in the next section. Still on the local level, Maslawis in general note that their accent is not heard very much nowadays, attributing this to the influence of non-MA speaking folks who have flooded the city in recent years for different reasons such as urbanisation, and sectarian tensions (more details in Section 5 below). This resulted in a demographic change in Mosul in which non-Maslawis have become the majority, according to Maslawis themselves.

5. Factors behind dialect levelling in Mosul

The precipitating factors often cited in research and commentaries on the occurrence of levelling in MA are the unprecedented events (sketched below) that took place in Iraq in general and more particularly in Mosul during the past few decades. The social changes brought about by such events have collectively led to the disintegration of the traditional community of Mosul and consequently led to the retreat in the use of MA. These factors can be described as follows.

5.1. Socio-psychological reasons
5.1.1 Mobility, contact and accommodation

Mosul has attracted immigrants from far and wide during the last three decades. The influx of newcomers (e.g. workers, farmers, and others) to Mosul from other parts of Iraq has at times been noticeable, locals claim, given the recognition of Mosul as a major economic and cultural centre of the northern region of Iraq. The newcomers who flooded Mosul have had a fragmenting effect on the social and demographic structure of Mosul. Such an effect has manifested as disintegration of closely-knit social networks and burgeoning of weak ties, which, according to Milroy & Milroy (1985), constitute an easy channel through which linguistic changes could make inroads into a community. Another linguistic corollary of the social network theory of Milroy & Milroy (Ibid.) is that tightly-knit networks, on the other hand, tend to preserve localised linguistic forms.

   In this light, the change in the social structure of Mosul brought Maslawis into contact with people speaking different varieties. As a result, linguistic accommodation, visibly on the part of Maslawis towards the newcomers' speaking habits has surfaced. Locals claim the variety is now more tolerant of non-MA linguistic elements, which are being absorbed by Maslawi youth. Mosul, then, appears to be an example of a community in which dialect contact is the norm.

5.1.2. Baghdadisation

Baghdadisation, as Duri (1978) put it, is 'the attachment to everything that is Baghdadi'. On the linguistic level, this means the convergence of speakers of other Iraqi dialects on Baghdadi Arabic (BA) – usually reflecting the wish of young people to sound like Baghdadis. Maslawis themselves comment that Mosul's young people nowadays tend to Baghdadise (or, as Maslawis say, *yitbaghdad*) in their speech.

5.2. Socio-political factors
5.2.1. Urbanisation

Urbanisation has been one of the main social changes that developed in the Arab world during the past half century (Miller 2004: 177). The effects of these demographic changes were felt more readily in the oil-rich countries whose economic boom led to the urbanisation of large areas and consequently promoted linguistic variation and change (Bassiouney 2009: 114). This is true of many Arab countries including Iraq.

5.2.2 Bedouinisation

A major consequence of the urbanisation movement is the so-called Bedouinisation. Already in operation in Baghdad (Abu Haidar 2006a), Mosul has been the recipient of people of Bedouin background over the past decades. This movement ensued for several reasons such as:

a)   Rural to city movement.
b)   Sectarian conflicts that brought people of different parts of Iraq to Mosul.
c)   Arabisation policy: A government-run policy that involved the displacement of non-Arabic speaking people of Mosul and other areas during the past decades and repopulating their areas with Arabic speaking people of rural and Bedouin backgrounds, which further increased the number of Bedouinised groups in Mosul.

5.2.3. Media

The putative role of the media is also reported in enhancing the process of dialect levelling (Gibson 2002: 25). Since the third gulf war, there has been an upsurge in the emergence of television channels directed at the youth market in which non-standard registers and accents (e.g. southern *gilit* varieties) have, to the detriment of other varieties, become *de rigueur*. In addition to this, there has been greater exposure to Modern standard Arabic (henceforth, MSA) through its widespread use in the media.

6. This study

This paper reports preliminary findings of a larger study on phonological variation and change in Mosul. The study is concerned with the phonological patterns and the trajectory of change in MA. This paper looks at the variable (q), as one of the main differentiating variables between *qeltu* and *gilit*, to assess whether the process of levelling is occurring in this sound and explain why it is (not) happening.

6.1. The variable

Abstractly corresponding to the 21st letter of the Arabic alphabet (ق) (qaf), the sound (q) and its dialectal manifestations have received a good amount of attention in Arabic phonology (Al-Ani 1976: 48). It is one of the features upon which Blanc's (1964) *qeltu-gilit* classification is based. Like all *qeltu* dialects, MA is characterised by the retention of the MSA voiceless, unaspirated stop [q] in all environments, even in those where it is usually [g] in *gilit*, as in *sooq* 'market' [sɔ:q] v. [su:g]. However, *gilit* has both variants [g] and [q] in certain distributions (for instance, *haq* 'right' [ħaq]).

(qaf) is also one of the interesting variables in Arabic given its geographical, social, and stylistic dimensions in the Arabic speaking communities. Hachimi (2005: 124) notes that dialectologists have employed the different realisational variants of (q) to delineate isoglosses to mark dialect boundaries of Arabic. This distinction has been in existence since as early as the pre-Islamic era (Ibid.). Blanc (1964: 29) notes that Arab sources appear to agree as to the existence of, for instance, a sedentary vs. nomad dichotomy based on the different realisations of (q). The [q] vs. [g] distinction is still an important marker in establishing the Bedouin and sedentary dialectal dichotomy in the Arabic-speaking world (Cadora 1992). This is typified by the use of the voiced velar stop [g] as being characteristic of Bedouin varieties while the voiceless uvular stop [q] is typical of its sedentary (urban) counterparts. A good example is presented by Blanc (1964) whose classification of the group dialects of IA was based upon [q] and [g] where the former is characteristic of the sedentary *qeltu* group of dialects and the latter is a feature of the *gilit* dialects.

6.1.1. Previous accounts of (qaf)

(q) is one of the recurrently visited variables in the literature on Arabic speaking communities (e.g. Blanc 1964, Cadora 1970, AI-Ani 1976, Abdel Jawad 1981, Holes 1987, Haeri 1991, 1997, Al-Wer 2007). These studies have shown that not only does this variable dichotomise communities into urban and rural but also it correlates, albeit differently from one dialect to another, with a set of social variables such as gender, social class, education, urban/rural backgrounds and stylistic levels (Hachimi 2005: 126). For instance, in Amman, Jordan, the increasing use of the standard variant [q] was found to be associated with formality of style and level of education, and was also found to be favoured by men and *fellahin* 'rurals' (Ibid.). This same variable is also now supplanting the traditional variants, [g], [ʔ] and [k] in the speech of the Jordanian people in Irbid (Al-Khatib 1988: 87).

6.1.2. Variants of (q)

Having developed differently in the Arabic-speaking world, (q) showcases several variants ranging from unvoiced, voiced, palatalised, or even [k], with each variant of these dating well back in history (Edzard 2006: 2). A main variant of (q) is the voiceless uvular stop [q], which is one of the consonants of the phonemic inventory of MSA (Ryding 2005: 15). This variant is still maintained in a number of urban dialects in the Arab world. It is well reported in different varieties in Syria, Tunisia, Morocco and Iraq. The voiceless uvular stop also exists in relatively few other places in Arabia – particularly in parts of Oman and Yemen (Edzard 2006).

Another common variant of (q) is the voiced velar stop [g]. This variant can be heard in the *gilit* varieties of IA, the Arabian Gulf countries, and many other Bedouin varieties in the Arab world like Upper Egypt *Said Misr* (Hachimi 2005).(q) is also debuccalised to a glottal stop [ʔ]. This variant can be found in Cairene Arabic and several other urban varieties spoken in the Levant countries such as Syria and Lebanon (Ibid.).

6.1.3. Phonological change involving /q/

Researchers attendant on the phonological variation and change of this variable from a historical perspective (e.g. Cantineau 1937, Garbell 1958, Johnstone 1963, Blanc 1964, Ferguson 1996, Edzard 2006) agree that [q] has witnessed several changes that date back over centuries. These changes have been reported in the different varieties of Arabic. Sibawayh, in his book *al-Kitab*, which is considered the first comprehensive and systematic grammar of the Arabic language, classified the consonants of the sound system of classical Arabic into two main groups: *majhurah* 'voiced' and *mahmusah* 'voiceless'. At first, [q] was classified as a voiced sound *majhour* (i.e. [g]). Later, probably in the eighth or ninth century, [g] appears to have been devoiced to become (q) (Blanc 1964: 29). Blanc (1969) also argues that this sound has gone through a linguistic push chain mechanism in which the Semitic /g/ as part of a triangle of velar phonemes (g, k, q), was fronted to [gj], [j] and largely developed in Bedouin settings into [g] (Edzard 2006).

This sound has been reported to be undergoing different phases of change in different parts of the Arab world surfacing with various manifestations. To illustrate this change in the behaviour of (q) variants, let us have a look at the findings reported for this variable in the Arab world. On the one hand, Ferguson (1996: 195) notes that [q] is receding in the face of standard prestigious norms, although with different interpretations of the notion of standard and prestige depending on the community studied.

In his study on the descriptive grammar of the Palmyra (also known as Tadmur) dialect, Cantineau (1934) found that Palmyra's localised form [q] was being supplanted by the regional

standard prestige variant [ʔ]. In Bahrain, speakers of rural Shiites background are shifting from their traditional [q] to [g], a characteristic of the urban educated Sunnis (Holes 1987).

In Iraq, [q] is perceived as an urban variant and is almost a typical realisation in *qeltu* varieties CBA and JBA, although these two communities have dwindled in number and the remaining members of these communities have already started adopting the *gilit* [g] (Abu-Haidar 1991). Al-Ani (1976) notes that the preservation of [q] is a firmly established feature in both Mosul and Tikrit. He reasons that the people of these chief *qeltu* areas are conservative in nature and their lives have seen little change. He also notes that these areas have been exporting rather than importing immigrants to other areas, especially to the capital city, Baghdad.

That said, this is not the whole picture of qeltu in terms of the (q) sound, however. The realisational distribution of [q] in other *qeltu* type varieties such as those spoken in Hit and Ana is rather different from Mosul and Tikrit in that it is in an intermediate stage, Al-Ani (1976) claims. A great deal of lexical items with (q), especially in Hit, are observed as [g] rather than [q]. Furthermore, recent local commentaries (e.g. Abu Haidar 2007) and personal observations of the speech of Tikritis, for instance, suggest that [g] in Tikrit is gaining more ground at the expense of the traditional [q].

7. Methodology

The study was carried out from a variationist perspective. This tradition grows out of questions to explore the possible contribution of a constellation of linguistic and extra-linguistic factors in shaping the phonological variability and change of language varieties. Variationist fieldwork approaches have their roots in other areas of language study, including historical and comparative linguistics, and traditional dialectology (Chambers & Trudgill 1998). However, it was the ground-breaking work of Weinreich, Labov and Herzog (1968) that established the modern variationist enterprise (Schilling 2013: 4). The variationist tradition is borne on the premise that dialectal variation is not random. Rather, it is governed by what Labov and his colleagues called "orderly heterogeneity" whereby regular patterns can be discerned by correlating social structure with the linguistic structure. This involves studying populations stratified by social parameters such as age, ethnicity, gender and socioeconomic status. These parameters are subject to different distributions across communities. Moreover, different patterns have emerged from each study conducted since these parameters, although often showing common linguistic patterns, are interpreted and thus approached differently by researchers (Tagliamonte 2011: 6).

7.1. Data

The dataset upon which the remarks of this paper are based was collected from 16 speakers of MA. A total of 539 tokens of (qaf) were coded. A ceiling of one single iteration of the same lexical item that contains this sound was adopted to ensure a variety of lexical items for each speaker rather than a sample dominated by repetitive words. Established elicitation techniques (e.g. informal sociolinguistic interviews, a map task, and picture-naming) were used to collect the data. The sociolinguistic interview is one of the most common methods of data collection used by variationists (Milroy and Gordon 2008). A typical sociolinguistic interview usually comprises questions relating to topics such as demography, community, neighbourhood, etc. and is gradually steered towards other preferable issues (Tagliamonte 2006: 39). It is not unproblematic to collect enough tokens of certain linguistic features in sociolinguistic interviews, especially if one aims to keep the effects of the fieldworker to a minimum (Hilton 2010: 131). Therefore, different types of elicitation methods were used to obtain enough tokens. The first, and main, part of each recording session consisted of informal conversation between

the informant and the author as a part of the sociolinguistic interview sketched above. This part often started with a conversational module in which a number of exploratory questions centred on topics of general interest such as demography, community, neighbourhood, etc. This technique serves as an assessment of whether the interviewee shows willingness to talk about any particular topic and then the interview progresses into more modules allowing informants to extend on any topic that particularly interests them, and to tell stories or narratives (Meyerhoff, Adachi, Nanbakhsh, & Strycharz, 2012: 130). These modules can be re-adjusted and combined by the interviewer into a larger conversational network in which the course of natural conversation is ensured so as to elicit more extended amounts of impromptu, carefree speech (Schilling 2013: 93).

The second part was composed of talk generated by two other techniques of collecting data used in the variationist tradition – the picture description task and map task. In the former, target tokens were elicited by presenting informants with a set of pictures depicting an event or a process, which the informants were asked to describe or discuss each picture. Regarding the latter, it involved two participants (i.e. the informant and the author) holding a map with the first participant's map showing landmarks and routes while some of these were missing on the second participant's map and vice versa. The first participant drew the route on his map as described by the second participant. The details in the maps were presented in a way to stimulate corrections and some discussion until the target was reached.

7.2. Speakers

The 16 speakers recruited for this study were born and raised in Mosul. All of them had recently arrived in the UK, with their families, for education and business purposes. MA speakers are locally pigeonholed as *qiqu* – a shibboleth used to describe the prototypical MA speakers derived from the use of voiceless uvular stop [q], a well-known feature of *qeltu*. They are also called *qhah* [qħaħ]: an MA-specific word meaning 'natives' to refer to deeply-rooted natives of Mosul. The speakers were divided by gender (10 males and 6 females) and grouped into three age brackets representing three main broadly-defined life stages: youth (18-30), middle age (30-45), and old (50+). Six young speakers were recruited while five speakers were recorded for each of the middle and old age cohorts. These three age groups span the three generations of MA speakers. In the generational differences that may or may not surface across these three generations, in particular between young and old groups, one can discern whether or not a change is occurring. The speakers were from two social classes prevailing in the MA-speaking community: lower middle class and middle class. Speaker selection was facilitated by judgment sample, using the widespread "friend of a friend" technique (Milroy 1987: 66). This technique involved approaching a local community member to whom the nature of the research is disclosed. This person was asked to suggest potentially suitable informants who can fulfil the criteria.

All the informants hail from MA speaking areas, which are largely centred in the old alleys in central Mosul on both sides of the city (see Figure 3). These areas are surrounded by the seven old gates of Mosul, all within the right bank area (*Sahel Al-ayman*), and whose catchment area includes neighbourhoods and alleys such as *Midan, Dawwasa, Sarjkhana, Nabi Sheet, Alnajjar, Bab altoob, Jamsheed, Farooq, Alrifa'i, Khatoonia, and Ahmadia*. Certain MA-speaking neighbourhoods located in the left bank also include *Alfaisaliya, Kafa'at, Alhadbaa', Sumer, Andalus* and *Aldhubbat*.

Figure 3. A map showing the MA speaking area of Mosul (adapted from Mappery 2015)

7.3. Procedures

Recording sessions were conducted for the informants in Newcastle, Manchester and Leeds. None of the speakers were given any information about the purpose of the recordings until after all the sessions had been completed. The recordings were carried out in quiet conditions in the subjects' homes. Obtruding echoes and outside noises were minimised in the recording environment, by using a small room. This was provided with soft surfaces such as upholstery and folded curtains that can absorb sound waves, thereby reducing echo. All the sessions were recorded at a sampling rate of 96.0 kHz with a 24-bit resolution and were then transferred onto a computer disk and saved as .wav files.

The recorder used for conducting these sessions was an Edirol R-09HR High-Resolution recorder. Prior to performing the recordings, a test recording was conducted for each speaker for a few minutes to ensure better functionality for the whole process before starting the main session.

The tokens extracted from the pool of data were auditorily and acoustically analysed for this study using the latest version of PRAAT open-source freeware phonetic analysis software (Boersma and Weenink 2014). Foulkes and Docherty (2006) note that integrating instrumental methods with auditory analysis can yield some new details unobservable through impressionistic measurement and thus can help the researcher to determine the variant produced.

8. Results and discussion

All the 539 tokens of (q) were realised with [q] by all speakers of all the social categories (i.e. gender, age, and social class) incorporated in the sample. All the extracted tokens contain at least one instance of (q) that is realised as a variant other than [q] in other dialects in Iraq to gauge whether there is any sign of change/drift towards those varieties with respect to this

particular variable. Table 1 shows a sample of the words produced by the speakers and how they were realised by them as against how they are realised in *gilit*.[9]

Table 1. Table showing a sample of the words extracted for analysis

Word	*qeltu*	*gilit*	Gloss
quul	[quːl]	[guːl]	say
qasseer	[qasˤɪːɣ]	[gisˤɪːr]	short
qaed	[qɛːʕid]	[gaːʕid]	sitting
lahaqtu	[laħħaqtu]	[laħħagit]	I followed
urooq	[ʔɪʕɣɔːq]	[ʔɪʕruːg]	a type of bread
qroon	[qruːn]	[gruːn]	horns
maslooq	[maslɔːq]	[masluːg]	boiled
fooq	[fɔːq]	[fɔːg]	up
sooq	[sɔːq]	[suːg]	market
qasaseeb	[qasˤɛːsˤiːb]	[gisˤaːsˤiːb]	butchers
sandooq	[sˤanduːq]	[sˤanduːg]	box
yethooq	[ʔɪðuːq]	[ʔɪðuːg]	tasting
aqraba	[ʕaqqabi]	[ʕagruba]	scorpion
yeaalqoon	[yiʕalquːn]	[yiʕalguːn]	hanging
tqoom	[tqɔːm]	[tguːm]	standing up
yetbaq	[yitˤbuq]	[yitˤubg]	standing by
qumtu	[qumtu]	[gumit]	I did
teaa'liq	[tiʕliq]	[tiʕlig]	jumble
qussa	[qusˤa]	[gusˤa]	forehead
qloob	[qluːb]	[gluːb]	hearts

One single exception was the word *garayib* 'relatives', which was realised with [g] by all the speakers indicating that this word is adopted that way since no alternative words appear to be available in the lexicon of MA, it is suggested. The preservation of [q] shows us that MA appears to be resisting the levelling of this variant, unlike the case reported in other *qeltu* varieties such as Hit, Tikrit, and Ana as well as other *qeltu* varieties spoken outside the political boundaries of Iraq. This result also corroborates Al-Ani's (1976) statement that Mosul is considered a stronghold for [q]. In addition, the retention of this defining feature provides evidence that MA is not destined to be dominated by *gilit* as was predicted by Abu Haidar (2007) and Palva (1983), at least as far as (q) is concerned. This means that, contrary to its sister CBA (Abu-Haidar 1991), it could survive as an established variety, at least for generations to come, unless conditions (e.g. disappearance of its speakers) similar to those of both CBA and JBA occur.

The reason for this preservation, it is suggested, is that this variant has its own social connotations that are strong enough to keep it intact among competitive alternatives. It acts as a symbol of Maslawi identity, which the MA-speaking community attempts to assert in view of the diluted environment of present-day Mosul. Participants have commented that their dialect remains part and parcel of their identity, and thus gives them pride to maintain it and to be recognised with it despite the fact they are being dominated by non-Malsawi people. It must also be remembered that this variant in particular is so special for Maslawis that even the

---

[9] The pronunciations for gilit are not part of the data collected for this study. Rather, they were provided by the author, as a native speaker of gilit, to give the reader a flavour of the differences between the two dialects concerning the sound /q/.

shibboleths *qiqu* and *qhah* by which they are known are based on it. Therefore it survives being supplanted even when the community is experiencing external pressures and concomitant linguistic change is expected in the variety.

Another possible reason is the nature of the relationships that exist between the members of the MA community. While the newcomers from several parts of Iraq have infiltrated into Mosul, it is reported that the MA-speaking community is still clustered around certain areas in central Mosul. Moreover, the affinity relationships in the community have not been disturbed. Some of my participants mentioned that all *qiqu* families know each other and still maintain relationships over their life span. Therefore, these factors have collectively maintained, to some extent, closely-knit ties that connect the community. Under such conditions, Milroy & Milroy (1985) note, linguistic traditions most likely tend to remain in use.

9. Conclusion

This paper was an investigation of the process of dialect levelling in the Arabic variety spoken in Mosul, Iraq. It was a contribution from a larger project investigating phonological variation and change in the same variety. Earlier in this paper, it was shown that several *qeltu*-type varieties are currently undergoing change in this particular sound due to numerous external factors. Mosul, which has witnessed a number of social upheavals, is thought to be on the march, and thus inevitably stands as a prime site for levelling. The study attempted to assess the levelling of [q], a process that has been reported in some *qeltu*-varieties in Iraq and beyond.

Given the almost categorical realisation of (q) as [q] by all speakers of all socially-stratified groups incorporated in the study, there is enough evidence of phonological resistance rather than levelling in MA in which [q] is still robust. It is thus not losing its marked features like its sister *qeltu* dialects in Hit and Tikrit as reported in the literature despite the social shakeup that resulted from the events mentioned earlier. While the fragmentation of the hitherto cohesive demographic fabric of Mosul has visibly intensified as a result of those events, the linguistic behaviour of MA speakers, unlike that of the natives of Hit and Tikrit, appears to be resisting such a pressure by maintaining close relationships and this is manifested in the preservation of identity-defining dialectal features (e.g. [q]).

References

Abdel-Jawad, H. R. E. (1981). *Lexical and phonological variation in spoken Arabic in Amman.* Ph.D dissertation, University of Pennsylvania.

Abu-Haidar, F. (1991). *Christian Arabic of Baghdad*. Wiesbaden: Harrassowitz.

Abu Haidar, F. (2006). Bedouinization. In Versteegh, K. (ed.), *Encyclopedia of Arabic language and linguistics,* 1, 269-274. Boston: Brill.

Al-Ani, S. (1976). The development and distribution of the Arabic sound 'qaf' in Iraq. In Berkes, N. & Donald, P. L. (eds.), *Essays on Islamic civilization,* 48-56. Leiden: Evert Jan Brill.

Al-damluji, S. (2014). *Mosul in the early twentieth century*. Amman, Jordan: Altaliaa Scientific House.

Al-Khatib, M. A. A. (1988). *Sociolinguistic change in an expanding urban context: a case study of Irbid city, Jordan.* Ph.D dissertation, University of Durham.

Al-Wer, E. (2007). The Formation of the dialect of Amman. In Miller, C. (ed.), *Arabic in the city: issues in dialect contact and language variation,* 5, 55-76. London: Routledge.

Armstrong, N. (2001). *Social and stylistic variation in spoken French: a comparative approach.* Amsterdam: John Benjamins Publishing.

Armstrong, N. (2002). Nivellement et standardisation en anglais et en français [Grading and standardization in English and French]. *Langage et société* 102(4), 5-32.

Bassiouney, R. (2009). *Arabic sociolinguistics*. Edinburgh: Edinburgh University Press.

Biadsy, F., Hirschberg, J., & Habash, N. (2009). *Spoken Arabic dialect identification using phonotactic modeling.* Paper presented at the EACL 2009 Workshop on Computational Approaches to Semitic Languages.

Blanc, H. (1964). *Communal dialects in Baghdad* (Vol. 10). Cambridge: Harvard University Press.

Blanc, H. (1969). *The fronting of Semitic g and the qāl-gāl dialect split in Arabic.* Paper presented at the International Conference on Semitic Studies held in Jerusalem.

Boughton, Z. (2005). Accent levelling and accent localisation in northern French: comparing Nancy and Rennes. *Journal of French Language Studies* 15(03), 235-256.

Britain, D. (1997). Dialect contact, focusing and phonological rule complexity: the koineisation of Fenland English. *University of Pennsylvania Working Papers in Linguistics* 4(1), 141–70.

Cadora, F. J. (1970). Some linguistic concomitants of contactual factors of urbanization. *Anthropological Linguistics* 12, 10-19.

Cadora, F. J. (1992). *Bedouin, village and urban Arabic: an ecolinguistic study* (Vol. 18). Leiden: Brill.

Cantineau, J. (1934). *Le dialecte arabe de Palmyre: grammaire* [The Arabic dialect of Palmyra: grammar] (Vol. 1). Beyrouth: Institut français de Damas.

Cantineau, J. (1937). *Etudes sur quelques parlers de nomades arabes d'Orient. (deuxième article)* [Studies on some dialects of Arab nomads of the East (second article)]. Paris: Librairie Larose.

Chambers, J. K., & Trudgill, P. (1998). *Dialectology*. Cambridge: Cambridge University Press.

Collin, R. O. (2009). Words of war: the Iraqi tower of Babel. *International Studies Perspectives* 10(3), 245-264.

Duri, A. (1978). *Mawrid*. State Cultural Affairs House, Ministry of Culture and Arts, Iraq: Aljahidh Press.

Edzard, L. (2006). qaf. In Versteegh, K. (ed.). *Encyclopedia of Arabic language and linguistics,* 4, 1-4. Boston: Brill.

Elyas, S. (2013). Mosul locals fear loss of their local language, http://www.niqash.org/en/articles/society/3248/mosul-locals-fear-loss-of-their-local-language.htm [Retrieved on 4th September 2015].

Esch, E. (2002). My dad's auxiliaries. In Jones, M. C. & Esch, E. (eds.), *Language change: the interplay of internal, external, and extra-linguistic factors,* 81-110. Berlin: Mouton de Gruyter.

Ferguson, C. A. (1996). *Sociolinguistic perspectives: papers on language in society, 1959-1994*. Oxford: Oxford University Press.

Foulkes, P., & Docherty, G. (1999). *Urban voices: accent studies in the British Isles*. London: Arnold.

Foulkes, P., & Docherty, G. (2006). The social life of phonetics and phonology. *Journal of Phonetics* 34(4), 409-438.

Garbell, I. (1958). Remarks on the historical phonology of an East Mediterranean Arabic dialect. *Word: journal of the International Linguistic Association* 14(2-3), 303-337.

Gibson, M. (2002). Dialect levelling in Tunisian Arabic: towards a new spoken standard. In Rouchdy, A. (ed.), *Language contact and language conflict in Arabic: variations on a sociolinguistic theme,* 24-40. London: Routledge Curzon.

Hachimi, A. (2005). *Dialect leveling, maintenance and urban identity in Morocco Fessi immigrants in Casablanca.* Ph.D dissertation, University of Hawaii at Manoa.

Haeri, N. (1991). Sociolinguistic variation in Cairene Arabic: palatalization and the /qaf/ in the speech of men and women. Ph.D dissertation, University of Philadelphia.

Haeri, N. (1997). The reproduction of symbolic capital: language, state, and class in Egypt. *Current anthropology* 38(5), 795-816.

Hilton, N. H. (2010). Regional dialect levelling and language standards: changes in the Hønefoss dialect. Ph.D dissertation, University of York, UK.

Hinskens, F. L. M. P. (1993). *Dialect levelling in Limburg: structural and sociolinguistic aspects*. Tübingen: Niemeyer (Linguistische Arbeiten).

Holes, C. (1987). *Language variation and change in a modernising Arab state: the case of Bahrain* (Vol. 7). Taylor & Francis.

Holes, C. (2007). Colloquial Iraqi Arabic. In Postgate, J. (ed.), *Languages of Iraq, ancient and modern,* 123-134. London: British School of Archaeology in Iraq.

Hornsby, D. (2002). Dialect contact and koinéization: the case of northern France. *Contributions to the Sociology of Language [CSL]* 86, 19-28.

Jastrow, O. (1978). Die mesopotamisch-arabischen qaltu-Dialekte. Band I: Phonologie und Morphologie [The Mesopotamian-Arabian qaltu dialects. Volume I: phonology and morphology]. *Abhandlungen für die Kunde des Morgenlandes Wiesbaden* 43(4), 1-330.

Jastrow, O. (2006). Iraq. In Versteegh, K., Elgibali, A., Woidich. M. & Zaborski, A. (eds.), *Encyclopedia of Arabic language and linguistics*, 414-427. Leiden/Boston: Brill.

Johnstone, T. M. (1963). The affrication of "kaf" and "gaf" in the Arabic dialects of the Arabian Peninsula. *Journal of Semitic Studies* 8(2), 210-226.

Kerswill, P. (1996a). Dialect levelling, koinéisation and the speech of the adult migrant. *University of Pennsylvania Working Papers in Linguistics* 3(1), 237-246.

Kerswill, P. (1996b). Milton Keynes and dialect levelling in south-eastern British English. *English: History, Diversity and Change* 1, 292-300.

Labov, W. (1963). The social motivation of a sound change. *Word: Journal of the International Linguistic Association* 19(3), 273-309.

Lodge, R. A. (2004). *A sociolinguistic history of Parisian French*. Cambridge: Cambridge University Press.

Manfredi, S. (2012). Dialect mixing and dialect levelling in Kordofanian Baggara Arabic (Western Sudan). In Barontini, A., Vicente, A., Zimari, K. & Caubet, D. (eds.), *Dynamiques langagières en arabophonies: variations, contacts, migrations et créations artistiques. Hommage offert à Dominique Caubet par ses élèves et collègues*, 141-162. Zaragoza: Universidad de Zaragoza, Área de Estudios Árabes e Islámicos.

Mappery.com (2015). Mosul Area Map. http://www.mappery.com/map-of/Mosul-Area-Map [Retrieved on 30 September, 2014]

Meyerhoff, M., Adachi, C., Nanbakhsh, G., & Strycharz, A. (2012). Sociolinguistic fieldwork. In Thieberger, N. (ed.), *The Oxford handbook of linguistic fieldwork*, 121-146. Oxford: Oxford University Press.

Miller, C. (2004). Variation and changes in Arabic urban vernaculars. In Woidich, M., Haak, M., de Jong, R. D. & Versteegh, C. H. M. (eds), *Approaches to Arabic dialects: collection of articles presented to Manfred Woidich on the occasion of his sixtieth birthday*, 177-206. Leiden: Brill.

Milroy, J., & Milroy, L. (1985). Linguistic change, social network and speaker innovation. *Journal of Linguistics* 21(02), 339-384.

Milroy, L. (1987). *Observing and analysing natural language: a critical account of sociolinguistic method*. Oxford: Blackwell.

Milroy, L., & Gordon, M. (2008). *Sociolinguistics: method and interpretation*. Oxford: Wiley.

Palva, H. (1983). Review of Otto Jastrow, Die mesopotamisch-arabischen qeltu-Dialekte. Band I: Phonologie und Morphologie [The Mesopotamian-Arabian qaltu dialects. Volume I: phonology and morphology]. Wiesbaden 1978. *Acta Orientalia* 43, 100-104.

Peoples, J., & Bailey, G. (2011). *Humanity: an introduction to cultural anthropology*. Belmont, CA: Cengage Learning.

Pooley, T. (2002). The depicardization of the vernaculars of the Lille conurbation. *Contributions to the Sociology of Language [CSL]* 86, 29-62.

Røyneland, U. (2009). Dialects in Norway: catching up with the rest of Europe? *International Journal of the Sociology of Language* 2009, 7-30.

Ryding, K. C. (2005). *A reference grammar of modern standard Arabic*. Cambridge: Cambridge University Press.

Schilling, N. (2013). *Sociolinguistic fieldwork*. Cambridge: Cambridge University Press.

Tagliamonte, S. A. (2006). *Analysing Sociolinguistic Variation*. Cambridge: Cambridge University Press.

Temple, R. (2001). The interaction between speaker sex and changes in regional patterns of voicing in metropolitan French: a case of over-levelling. In Hintze, M. A., Pooley, T. & Judge, A. (eds.), *French accents: phonological and sociolinguistic perspectives* 148-164. London: Centre for Information on Language Teaching and Research.

Tagliamonte, S. (2011). *Variationist Sociolinguistics: Change, Observation, Interpretation*. Malden, MA: Wiley.

Thanoun, T. (2010). *Abhath fi tarikh Almosul almuasser* [Research on the contemporary history of Mosul]. Mosul: Ibn Al-Atheer House.

Trudgill, P. (1986). *Dialects in contact*. Oxford: Blackwell.

Williams, A., & Kerswill, P. (1999). Dialect levelling: change and continuity in Milton Keynes, Reading and Hull. In Foulkes, P. & Docherty, G. *Urban voices: accent studies in the British Isles*, 141-162. London: Arnold.

Weinreich, U., Labov, W., & Herzog, M. I. (1968). *Empirical foundations for a theory of language change*. Austin: University of Texas Press.

Abdulkareem Yaseen
School of Education, Communication & Language Sciences
Newcastle University King George VI Building
Newcastle Upon Tyne
NE1 7RU
United Kingdom

a.y.a.ahmed@newcastle.ac.uk

THE IMPACT OF AUDIO-VISUAL MEDIA ON ENGLISH PHONOLOGICAL ACQUISITION AND DEFOSSILIZATION BY LEARNERS IN CHINA

CHEN ZHANG
(University of Sheffield)

Abstract

This paper investigates the effectiveness of 'audio-visual media' as language input for English phonological acquisition and defossilization among Chinese-speaking learners of English in the Chinese context. Supported by the exemplar theory (Johnson 1996, Pierrehumbert 2001), the necessity of native-like phonological input is greatly emphasized. However, in the current Chinese context, the phonological input is generally non native-like. Also, given the fact that adult learners in China have a fossilized English phonology, it is important to find an alternative form of input. Audio-visual media have the potential to become the alternative, on account of being a provider of native-like English phonological characteristics and a provider of visual cues. The impact of audio-visual media was tested in a natural self-learning setting by implementing a pre-test and a post-test on vowel quality. Eighteen Chinese English learners participated. Over a half-month period, nine subjects received input from an English TV series, while the other nine did not receive any input. The results suggest that audio-visual media can have a positive impact on L2 (i.e. second language) phonological acquisition. Audio-visual media can also help Chinese L2 English learners break the deadlock in fossilized L2 phonology, and thus push forward the development of their interlanguage phonology.

1. Introduction

The present preliminary study aims to test the impact of audio-visual media as language input for English phonological defossilization by L2 English learners in China. This project is motivated by the need for native-like phonological input in second language research and the current situation of English phonological input in China. Native-like input is a necessary component in L2 phonological acquisition (see Section 2.1). However, considering the nature of L2 English education in China, the English input that Chinese-speaking learners of English receive is non native-like. The L2 learner's interlanguage has obvious characteristics from Chinese transfer and has a tendency to be fossilized (Xiang 2009, Chen 2011). The non native-like phonological input can aggravate learners' phonological fossilization and impede the process of defossilization. Considering this problem, it is essential that we find alternative L2 English input to better facilitate phonological acquisition and defossilization processes among Chinese-speaking learners of English (Zhang Y 2013). Due to the fact that audio-visual media simultaneously provide both unmodified target language input in the sound tracks and moving visual images, it would appear to be the best alternative form of input to traditional classroom teaching in the Chinese context, and this is tested in this study.

　　The impact of audio-visual media has gained increasing attention in the past few years. However, it is still a new research topic which needs to be comprehensively investigated. Few studies focus on testing the impact of audio-visual media on L2 phonological acquisition. Therefore, to fill this research gap, an intervention experiment, which consists of a pre-test and an immediate post-test, was conducted to test the effectiveness of audio-visual media. The main research question addressed was 'Can audio-visual media have a positive impact on L2 acquisition of phonology?'

This paper has six sections. In Section 2, the research background, including the necessity of native-like language input in L2 phonological acquisition, is presented. The problem of English input in China and phonological fossilization is also considered alongside the alternative L2 input of audio-visual media. Section 3 introduces the research methodology of the experiment. Next, in Section 4, the data analyses and results are outlined. The discussion is presented in Section 5 and this paper is concluded in Section 6.

2. Research background
2.1. Native-like input in L2 phonological acquisition

The past few decades have witnessed a great development of language input research in the field of second language acquisition (Gass 1997, Ellis, R 1985, Ellis, R 1994, Krashen 1982, Krashen 1985, VanPatten 2004, Carroll 2007, Schwartz 1993, Rast 2008, Ellis N 2009, Piske & Young-Scholten 2009). The necessity of language input has been commonly demonstrated and acknowledged (Gass 1997). Learners cannot acquire the target language in a vacuum. Specifically, phonological input, as an integral part of language input, is undoubtedly essential in phonological acquisition. Flege (2009: 175) defines phonological input as 'all L2 vocal utterances the learner has heard and comprehended'. Both native-like and non native-like L2 utterances are included in this definition. Comparing these two kinds of utterances, Jenkins (2000: 207) claimed that native-like utterances are unnecessary, with regard to intelligibility, when using English to communicate internationally. Thus the demand for native-like input in L2 phonological acquisition is reduced. However, as far as I am concerned, native-like input is far more important than non native-like input in L2 phonological acquisition. This viewpoint is supported by the exemplar theory (Johnson 1996, Pierrehumbert 2001).

The notion of exemplar was first introduced from psychological models for the study of speech sounds by Johnson (1996). This exemplar model works for both phonological perception and production. In this model, every phonological category is represented by 'a large cloud of remembered tokens of that category' in the mind (Pierrehumbert 2001: 140). The categories are systematically organized. The sounds which are perceived as similar phonological units are stored closely, while the dissimilar units are stored far away from each other. If new phonological tokens are perceived, they will be categorized into the 'clouds' according to their similarity to the existing tokens. If the 'cloud' has more numerous and activated exemplars, it will have advantages over the competition (Pierrehumbert 2001: 142). Different 'clouds' of sound units which are perceived from the linguistic input establish the recognition of the phonology of the target language. Phonological input is stored as exemplars and production originates from them. In exemplar theory the authenticity of phonological input is of great significance. Only if this is native-like can L2 learners locate the 'clouds' in the right place and thus facilitate their acquisition process in both perception and production.

From the above explanation, it is obvious that native-like input is of great necessity in L2 phonological acquisition. Therefore, phonological input should involve as much native-like input as possible.

2.2. The main problem with L2 English input in China and phonological fossilization

There is seemingly a strong desire among the Chinese public to acquire a foreign language proficiently for the purpose of worldwide political, economic and cultural communication. Governmental policies also encourage the public to learn a foreign language (The Ministry of Education of P. R. China 2011). In China, English has an unshakable status, and is considered as a very important worldwide communication tool. There has been an upsurge of enthusiasm towards English language learning. Though it appears that the whole of Chinese society realizes

the importance of English education and seems to be active in popularizing English among the public, there are admittedly still many problems in current English language teaching and learning. Among the problems, the lack of native-like phonological input is the most prominent one.

In the Chinese context, the phonological input generally comes from 'teacher-talk' in classrooms. It is the main source of English input that L1 (i.e. first language) Chinese learners can receive. According to Li & Wang (2006), more than eighty per cent of English input that the learners receive comes from the English language classroom. Liu (2010) obtained a similar result on the percentage of classroom input contribution (seventy-eight per cent) in the overall L2 English input in China. 'Teacher-talk' plays an important role in L2 target language learning. However, as is mentioned in Young-Scholten's work (1995: 111), 'teacher-talk' is 'typically L1 accented' in foreign language classrooms. A similar problem is also found in Chinese L2 English classrooms, and it appears to be the main notable hindrance in phonological acquisition. The overwhelming majority of the language teachers in China are native speakers of Chinese but not native speakers of English. The speech that they deliver in classroom teaching is not native-like English, especially in terms of phonology (Jia 2011, Li R 2009). The English teachers in China are also English learners themselves. Most of them have not had the opportunity to go to English speaking countries (Li X 2007, Deng 2007, Xu & Guo 2009, Zhang J 2009, Chai & Lu 2009, Tu 2012) and they rarely interact with native speakers of English. Therefore, their English phonology is far from native-like, and their speech still carries very obvious interlanguage characteristics which come from Chinese phonology and individual dialectal accents (Yang & Ren 2009). Another problem which exists in 'teacher-talk' in the Chinese classroom is that the speech that teachers deliver is exaggeratedly modified. Adjustments frequently occur in pronunciation in language classrooms (Henzl 1979). In the Chinese context, teachers usually slow down the speech speed, use long pauses when waiting for a reaction from students, and accentuate certain words in the speech for emphasis (Li H 2011). These phonological modifications can potentially make the input much more comprehensible for L2 English learners (Loschky 1994, Krashen 2003) and might be beneficial for learners when understanding the teacher's talk. However, they do not benefit the learners in phonological acquisition. In a word, 'teacher-talk' in China fails to present native-like English phonological characteristics. Thus, it is inefficient in facilitating English phonological acquisition. Given the fact that non-native 'teacher-talk' is the main source of phonological input in China, the lack of native-like phonological input is widespread.

In China, many English learners have a fossilized English phonology due to the lack of native-like phonological input. This phenomenon can be explained from the two measures of L2 phonological input proposed by Moyer (2009), that is the 'age of onset' and the 'length of residence' respectively. With respect to 'age of onset', it refers to the age when a L2 learner starts learning the target language. In Moyer's claim (2009), early 'age of onset' is beneficial for L2 phonological acquisition. This viewpoint is supported by the 'critical period' hypothesis in L2 phonological acquisition (Scovel 1998, 2000), whereby puberty is the dividing point for speech learning ability. Before puberty, language learners have strong speech learning abilities. In contrast, after puberty, the speech learning ability declines due to neural maturation. If learners receive exposure to native-like L2 phonological input before puberty, they can possibly acquire the L2 phonology successfully. In China, English learners start learning English in the classroom from primary school at the age of six (The Ministry of Education of P. R. China 2011). In other words, they do have an early 'age of onset'. However, this advantageous effect is not manifested in their English phonological acquisition, because the phonological input they receive before puberty is generally non native-like. Their knowledge of English phonology obtained from 'teacher-talk' has traits of Chinese phonology and dialectal accent, and the learner's interlanguage is obviously influenced by L1 transfer. It is argued that L1 transfer is a

prominent affecting factor in L2 phonological acquisition and production (Hansen 2006, cf. Altenberg & Vago 1987, Benson 1988). Thus, the L1 Chinese learner's English phonology does not develop much before puberty. Considering the fact that exposure to non-native input is continuously given after the critical period, when their speech learning ability has declined, the L2 learner's English phonology is highly likely to be fossilized. In terms of 'length of residence', most L2 English learners in China do not have the experience of living in English speaking countries to receive native-like phonological input. This is restricted for various social and economic reasons. By 2010 in China, there were 400 million L2 English learners. It is expected that there will be a billion English learners by 2020 (Cao 2012). Given the huge number of Chinese speaking learners of English, it is impossible that many will have the opportunity to go abroad and become immersed in English speaking contexts. Therefore, generally speaking, L2 English learners in China have no 'length of residence'. Without 'length of residence' in English speaking countries, learners cannot be exposed to native-like English phonology, and thus cannot benefit from it to reorganize the exemplar 'clouds' and defossilize their English phonology.

Considering the lack of native-like phonological input and learner fossilization in L2 English phonology, it is desirable to find an alternative form of input which contains unmodified native-like English to facilitate L2 acquisition and defossilization. With the development of the internet and the opening-up of the media policy in China, audio-visual media in English are now widely accessible to English learners in China. Audio-visual media contain a large amount of unmodified native-like phonological input, and so it is an ideal alternative to 'teacher-talk'.

2.3. Audio-visual media as alternative phonological input

As language input, audio-visual media has several characteristics. Firstly, it mostly contains rich phonological characteristics spoken by native speakers of English. The quantity of audio-visual media is abundant and is available to the majority of English learners in China. The native-like input from audio-visual media has the potential to defossilize the L2 learner's English phonology. Secondly, audio-visual media simultaneously combines moving visual images with a sound-track. This combination makes the language input much easier for the learners to comprehend and process, because non-linguistic visual cues make the input more noticeable (Plass & Jones 2005: 472). Visual cues from audio-visual media can also help L2 learners recognise the exemplars and process the speech provided in the input. The learners can receive the input, including, for example, information about the lip contour, mouth location, mouth scale and orientations from the actor and the actress. Fanelli et al. (2009) affirm the role of audio-visual fusion in speech processing. They argue that video information provided together with the audio can significantly promote the accuracy of recognition. This viewpoint has also been approved by Macaluso et al. (2004) and Niyogi et al. (1999). Based on these characteristics, it is promising that audio-visual media can positively influence L2 English phonological acquisition by L1 adult Chinese learners. Besides these characteristics, audio-visual media can also convey cultural information, which is interesting to explore. However, since it is outside the research scope of the current study, cultural information is not discussed in this paper.

Past studies on the impact of audio-visual media mainly focus on other aspects of L2 acquisition, such as incidental vocabulary learning (Neuman & Koskinen 1992, Yuksel & Tanriverdi 2009, Karakas & Saricoban 2012). Just a few studies have been carried out on L2 phonological acquisition. From the current literature, Davis and Kim (2001) examine the impact of visual speech on the acquisition of L2 Korean by ten native-speakers of English. They compared the effectiveness of audio with visual speech (i.e. video) information (e.g. mouth

movements, lip contours) and the effectiveness of audio without visual speech information. The results revealed that the fusion of audio-visual cues is beneficial when English speakers learn Korean. Similar results were found in Ortega-Llebaria et al. (2001). In their study, thirty-six Spanish learners of English significantly improved in English phonology with the help of audio-visual (or video) stimuli. However, evidence demonstrating the impact of audio-visual media in the Chinese context is rare. Therefore, the current study contributes to the literature in this field and provides pedagogical implications.

Based on the characteristics and implications of previous studies, it is hypothesized that audio-visual media can have a positive impact on L2 English phonological acquisition by L1 Chinese learners.

3. Research methodology
3.1. Experimental focus and phonological language background

In the experiment, the focus was on aspects of vowel quality including monophthongs and diphthongs. The corresponding phonological language background information is presented as follows.

There are five vowels in Chinese, namely /i/, /y/, /u/, /ə/, /a/. When high vowels occur before another vowel, they behave as glides /j, ɥ, w/. /i/ and /u/ can also follow a non-high vowel to form a diphthong. For instance, /ai/ and /ou/ are diphthongs which are commonly found. However, compared to English, the Chinese diphthongs are less diphthong-like, based on the degree of Formant 1 (F1) and Formant 2 (F2) changes (Zheng & Zheng 2011).

It is also worth noting that there are language-specific phonological constraints in different languages (Klima & Bellugi 1979). There are several combinations of consonants and vowels which exist in English. However, these combinations are impossible in Chinese. These English-specific phonological constraints may be difficult for L1 Chinese-speaking learners. For instance, /hi:/ is a commonly seen consonant-vowel combination in English. However, it is impossible in Chinese. Chinese speaking learners might have difficulties in pronouncing these combinations without exposure to native-like English input. Thus, they might possibly refer to their L1 knowledge of phonology and choose the closest pronunciation (e.g. pronounce /hi:/ as /hei/). Because of this, there might be a chance that the degree of F1 and F2 changes of these vowels in the learner's interlanguage is higher than it should be in English.

3.2. Research hypotheses and predictions

Given the above discussion of audio-visual media as L2 input, the experimental focus and the nature of learners' L1 phonology, the research hypotheses of the current study are as follows:

1)   Audio-visual media can have a positive impact on English phonological acquisition and defossilization of monophthongs.
2)   Audio-visual media can have a positive impact on English phonological acquisition and defossilization of diphthongs.

Given the nature of audio-visual media, it is predicted that it may positively influence the L1 Chinese learners' pronunciation and defossilization of monophthongs and diphthongs.

3.3. Research subjects

Twenty-one Chinese L2 English learners were selected to participate in the experiment. After the pre-test, three participants were screened out, on the grounds that they watched films and

TV series for relatively greater amounts of time than the other subjects. There remained eighteen participants who had similar linguistic and educational backgrounds. At the time of testing, the average age of the participants was twenty-two. They had their English education in an EFL context from the age of six and they had been exposed to 'teacher-talk' in L2 English classrooms for about 16 years. All the participants were in China at the time of testing, none of the subjects had been to an English-speaking country and their L2 English input was mainly from classroom instruction.

3.4. Research materials

In this study, Received Pronunciation (RP) was used as the criteria of measurement. By definition, RP refers to the standard accent spoken in the United Kingdom (Wells 1982). When choosing the input stimuli and the baseline material, the main concern was whether it contained RP. In terms of the input stimuli, a comic TV series made by BBC Two in 2009 called *Miranda* was chosen. The main actress, whose name is Miranda Hart, speaks RP (Another Tongue 2015). The first two seasons were adopted in the experiment. Each season included six episodes and each episode lasts for about thirty minutes. The total input time of treatment was six hours. The chosen TV series provided sufficient consistent phonological input in stimuli during the given period and kept the participants interested in the experiment. Also, the TV series provided contained visual components, such as lip contours, facial expressions and body gestures from the actor and the actress. Considering the negative influence that orthographic input might have in L2 phonological acquisition (Young-Scholten 1995, Bassetti 2009), English subtitles were not included in the given TV series. With regard to the baseline material, since the Queen's English shows strong characteristics of RP (Cruttenden 2014: 79), the *Queen's address to Parliament during the Diamond Jubilee* was chosen.

3.5. Research procedures

This study was carried out in a natural self-learning setting. This intervention study consisted of a pre-test and an immediate post-test. A delayed post-test was considered when designing the experiment. However due to the limitations of the research time scale, it was not conducted. In the pre-test stage, all eighteen participants were required to record their reading of the Queen's address. After the recordings were collected, the participants were randomly allocated to either the experimental group (N=9) or the control group (N=9). During the input period of two weeks, the experimental group was asked to watch the given TV series, whereas the control group was not. The experimental group watched the TV series (two seasons) just once via the internet on home computers individually. They were told to watch the given TV series at their own pace and report to the researcher. The control group had no exposure to English input during the input period. Meanwhile, none of the participants in either the experimental group or the control group received any classroom instruction in English or other types of audio-visual media input. Immediately after the input period, all the participants were required to record their reading of the Queen's address again. They were not given time for pronunciation rehearsal. The recordings collected in the pre-test and the immediate post-test are the main data from the experiment. Table 1 below presents the general research procedures.

Table 1. Research procedure

|	Pre-test	Stimuli input period (two weeks)	Immediate post-test
Control group (N=9)	√	×	√
Experimental group (N=9)	√	√	√

4. Data analyses and results

All the data were analyzed in the phonological analysis software 'Praat'. Statistical analyses were conducted using SPSS. The next two subsections present the data analyses and results from the perspective of vowel quality, including monophthongs and diphthongs.

4.1. Vowel quality - monophthongs

In the examination of monophthongs, it was found that /ɔ:/ had a tendency to be more open in participants' interlanguage in the pre-test. This was observed based on the F1 values from the word 'taught' in the Queen's address. Generally speaking, the F1 value represents the vowel openness. Compared to close vowels, open vowels have higher F1 values. From Figure 1 shown below, it is obvious that the F1 values of both the experimental group (Category 2 in yellow) and the control group (Category 1 in green) are much higher than the baseline (Category 0 in blue). A statistical comparison made in the F1 values revealed that there was no significant between-group difference in the pre-test (p=0.957 > 0.05). Therefore, the performances of the two groups are valid for comparison from the starting point. After the input period, the experimental group's mean F1 value was much lower than the control group's F1 value and much closer to the baseline in the immediate post-test. Though the changes that the two groups made from pre-test to immediate post-test were not statistically different (p=0.117 > 0.05), I argue that audio-visual media had a slight positive impact according to the level of mean F1 values.

Figure 1. Between-group comparison: mean F1 values of /ɔ:/
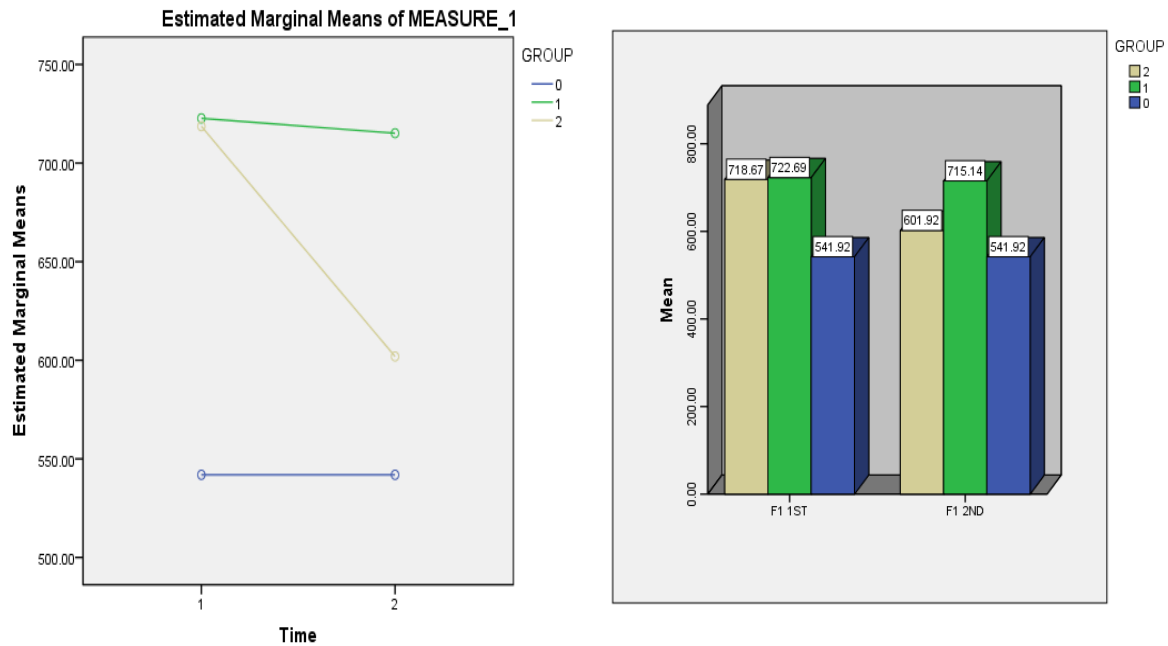
Apart from the finding of /ɔ:/, it was found that the monophthong /i/ in the word 'office' in the Queen's address had a tendency to be diphthongal in the pre-test. This finding was based on the great rate of change in F1 (representing the vowel openness) and F2 (representing the vowel frontness) values in the pre-test. In the analysis, the diphthongal extent was measured by using the rate of change (ROC), calculated using the equation '(End F value- Start F value)/Duration' (Deterding 2000). Big absolute ROC values indicate great changes. Figure 2 and Figure 3 below present the between-group comparisons of mean F1 ROC values and mean F2 ROC values of /i/ respectively.

Figure 2. Between-group comparison: mean F1 ROC values of /i/
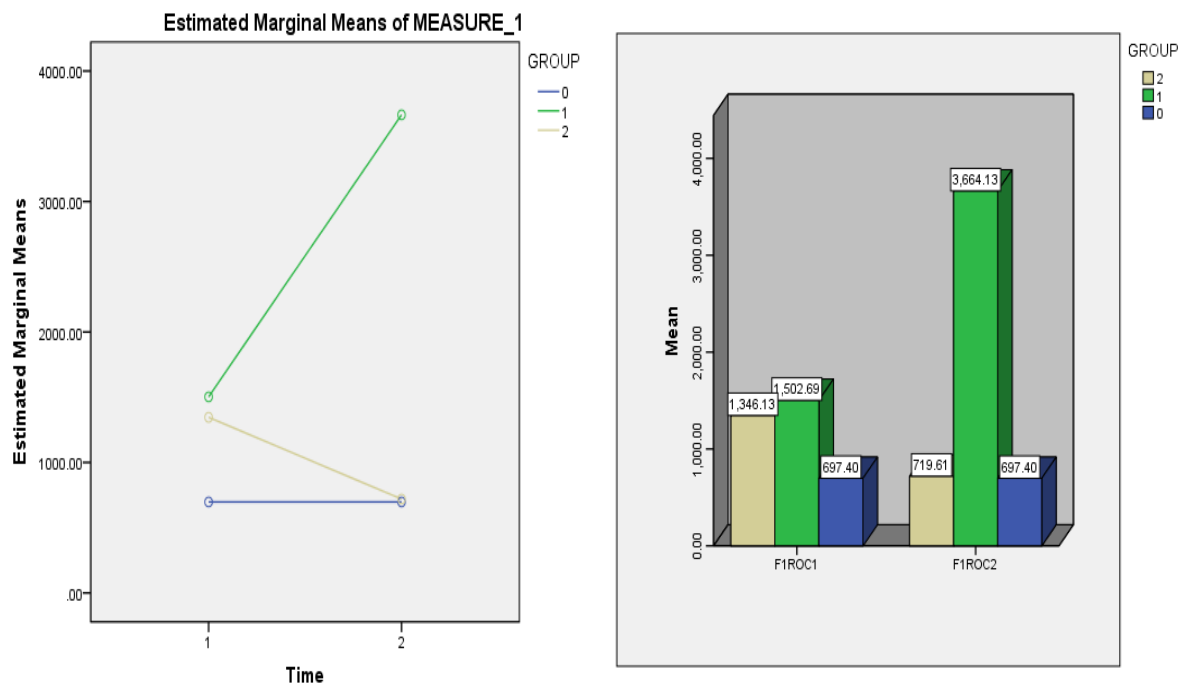
From Figure 2, it is obvious that in the pre-test, both the experimental group (Category 2 in yellow) and the control group (Category 1 in green) had a larger ROC value in F1, compared to the baseline (Category 0 in blue). Statistical analysis demonstrated the group homogeneity in the pre-test (p=1.000 > 0.05). After the audio-visual input period, the experimental group's F1 ROC value dropped and was very much closer to the baseline. However, the control group's F1 ROC value continued to be larger in the immediate post-test. Though the statistical results did not show a significant difference between the experimental group and the control group in the immediate post-test (p=0.318 > 0.05), the F1 rate of change of the experimental group did become smaller.

Figure 3. Between-group comparison: mean F2 ROC values of /i/
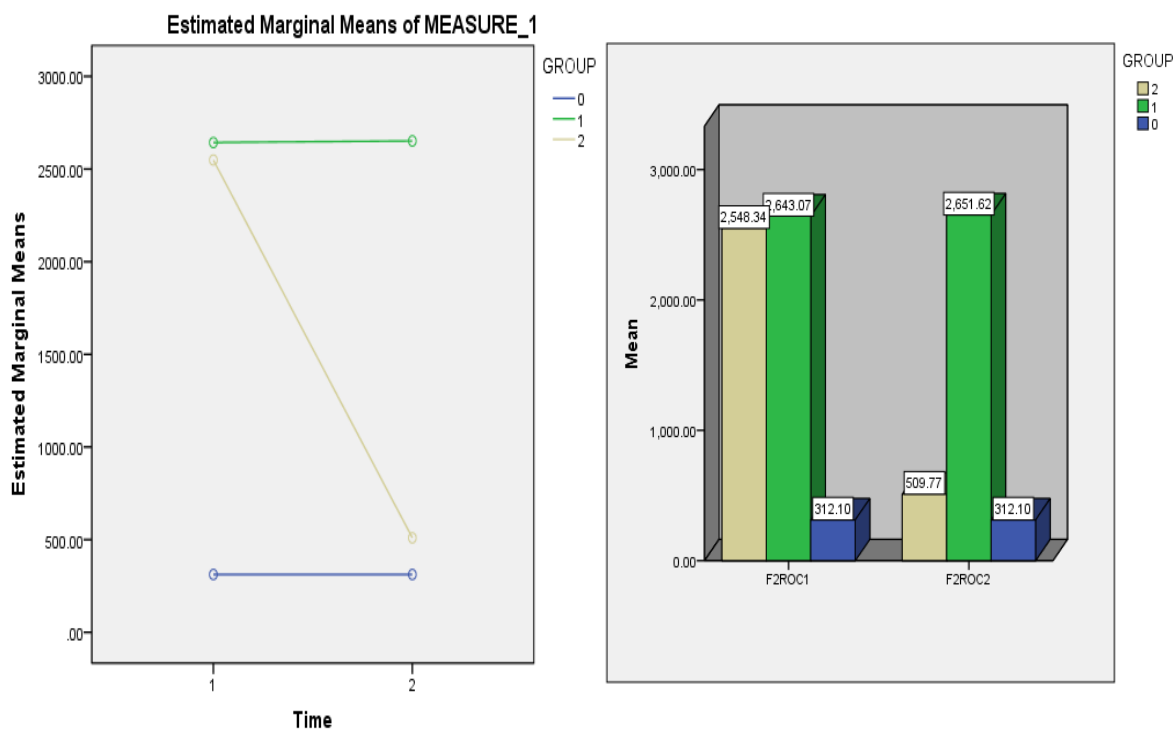
As is shown in Figure 3, in the pre-test stage the F2 ROC values of the experimental group and the control group were very similar. Statistical results show that the two groups were homogeneous in the pre-test (p=1.000 >0.05). After the input period, in the immediate post-test, the experimental group's F2 ROC value significantly dropped down to the baseline level (p=0.018 < 0.05). Also, there was a significant difference between the experimental group and the control group in the immediate post-test (p=0.009 < 0.05).

When summarizing the results of F1 ROC values and F2 ROC values, it is obvious from Figure 2 and Figure 3 that /i/ pronounced by the experimental group is less diphthongal in the immediate post-test, compared to the control group. Thus, the audio-visual media input positively influenced the pronunciation of /i/ in the word 'office'.

4.2. Vowel quality - diphthongs

When examining the diphthongs, it was found that /əu/ was less diphthongal in the participants' interlanguage in the pre-test. In English, there is a clear glide from /ə/ to /u/ when pronouncing the diphthong /əu/. The F1 value, which represents the vowel openness, and the F2 value, which represents the vowel frontness, are both expected to change greatly. However, the interlanguage characteristics in the pre-test showed that the participants generally glided from /ə/ to /ɔ/, and then ceased to glide to /u/. Their F2 values changed as is the way in English, but their F1 values did not change much. Figure 4 below presents the between-group comparison of F1 ROC values for the diphthong /əu/.

Figure 4. Between-group comparison: mean F1 ROC values of /əu/
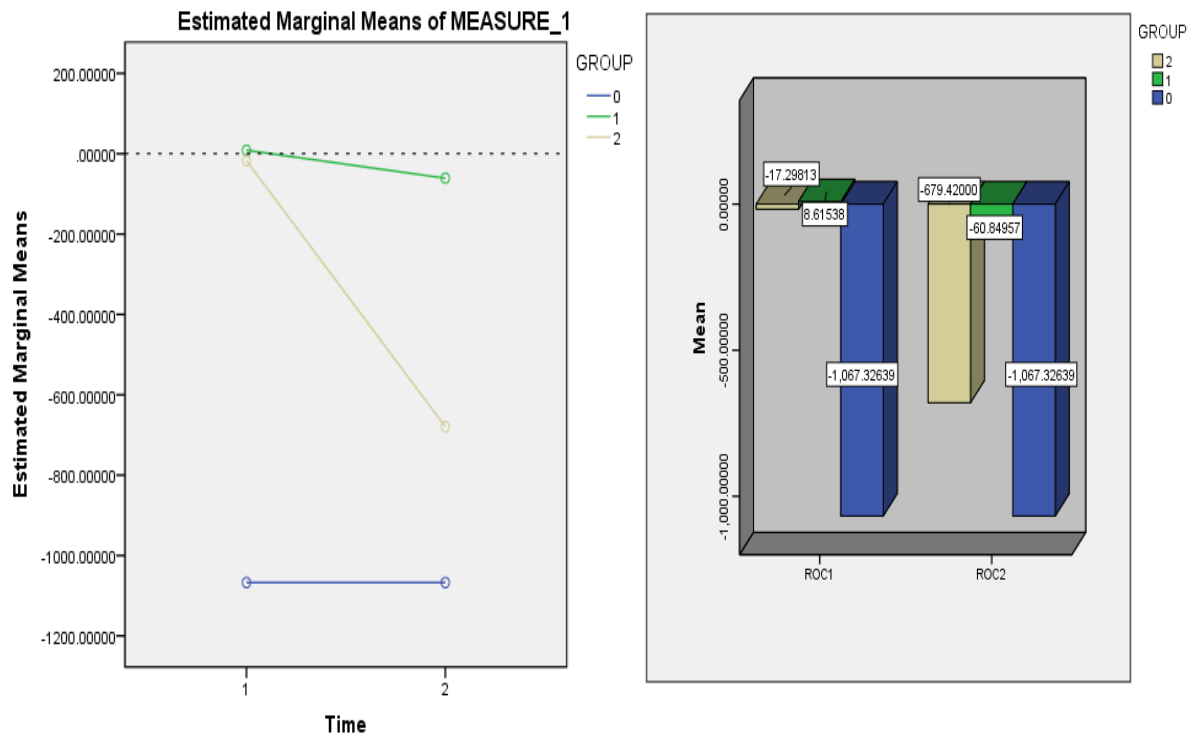
From Figure 4 we can see that the between-group difference in the pre-test is very subtle. This is also manifested in the statistics, where no significant difference is found between the experimental group (Category 2 in yellow) and the control group (Category 1 in green) at the starting point (p=0.929 > 0.05). Figure 4 also shows that the mean absolute F1 ROC value of the control group does not vary much from the pre-test to the immediate post-test. In comparison, the mean absolute F1 ROC value of the experimental group increases drastically from the pre-test to the immediate post-test and reaches towards the baseline value (Category 0 in blue). Statistical analyses reveal that the experimental group significantly improved in pronouncing /əu/ (p=0.002 < 0.05) from the pre-test to the immediate post-test. Apart from this, another significant difference was found in the between-group comparison in the immediate post-test (p=0.015 < 0.05). It is apparent that the audio-visual media appear to have had a positive effect on the English pronunciation of diphthong /əu/.

5. Discussion

This study investigated the effectiveness of audio-visual media on English phonological defossilization by Chinese L2 adult learners. The results shown in the previous section provided evidence for the hypothesis that audio-visual media can influence L2 English phonological acquisition positively, with regard to vowel quality (monophthongs and diphthongs). The effectiveness can be accounted for by the reduction of negative first language transfer, the visual speech aid and the input of native-like phonological exemplars.

As can be seen in the investigation of diphthongs, obvious transfer from L1 Chinese was detected in the participants' interlanguage. Their pronunciation of diphthongs was less diphthongal according to the F1 ROC values in the pre-test. As explained in Section 3.1, Chinese diphthongs are less diphthongal than English ones, and the characteristics of Chinese diphthongs are retained when participants speak English. Another interlanguage characteristic

from Chinese transfer was manifested in the diphthongal /i/. As stated in Section 3.1, there are several combinations of vowels and consonants which exist in English that cannot be realized in Chinese. Among them, /fi/ is an impossible combination in Chinese. Chinese learners would look for a similar combination in their knowledge of Chinese phonology and thus pronounce /fi/ as /fei/. Audio-visual media, which provide high quality native-like English input in sufficient amounts, weaken the negative transfer from the first language. As is shown in the results, /əu/ became more diphthongal and /i/ became less diphthongal in the experimental group's performance in the immediate post-test, and their pronunciations became much closer to the standard baseline. Audio-visual media can potentially break the fossilized deadlock in the interlanguage caused by negative first language transfer, and thus push forward the continuum of phonological acquisition.

In the data analysis of the diphthongs, it was found that in the pre-test, the participants glided from /ə/ to /ɔ/ and then ceased to glide to /u/ at the end. In other words, the participants moved the vowel from the front to the back correctly. However, they failed to control the vowel openness, as is the way in English. After the input stimuli of audio-visual media was given, the experimental group significantly improved in the control of vowel openness. I argue that this might be attributed to the visual speech aid provided in the audio-visual media. With the provided stimuli, it is hypothesized that participants in the experimental group could see the mouth movement and contours from the actors and the actress, which offered cues to vowel openness. By referring to the lip contours, the participants who received the stimuli glided from the front position to the back and at the same time closed the mouth in order to pronounce /əu/ correctly. This further provides demonstrative evidence for Davis and Kim's study (2001).

The results also show that /ɔ:/ has a tendency to be more open in participants' interlanguage in the pre-test. After the input period, the experimental group performed better than the control group, according to the mean F1 values, though not significantly so. This indicates that the experimental group might have conquered the problem of vowel openness with the help of audio-visual media. In my opinion, this might be due to the fact that audio-visual media can provide learners with effective and native-like phonological exemplars. The phonological inventory in the interlanguage might be positively influenced, and so be more likely to be reorganized, thus promoting the process of L2 English phonological acquisition.

This study also offers practical implications. To compensate for the lack of native-like phonological input in China, audio-visual media can be widely applied in classroom teaching, as well as individual learning. Chinese learners of L2 English can benefit from the language input in audio-visual media without being immersed in English-speaking countries.

Due to the scale of the project, this study has some limitations. Admittedly, the input period (2 weeks) is considered to be a little short. If the period could be extended, the effects of the input from films and TV series would be much more salient and more data about, for example, consonants or perhaps sentence intonation, could be elicited. Also, if time allowed, delayed post-tests could also be conducted, in order to see whether the audio-visual media can have a long-term effect.

6. Conclusion

The present study investigated the effectiveness of audio-visual media as language input for L2 English phonological defossilization by adult learners in China. An intervention study which consisted of a pre-test and an immediate post-test was conducted. The results revealed that audio-visual media might potentially have a positive influence on the acquisition and defossilization of vowel quality, with regard to monophthongs and diphthongs. The positive impact was attributed to the nature of audio-visual media input, which can weaken negative first language transfer, provide a visual speech aid and offer native-like phonological

exemplars. This suggests that audio-visual media can tackle the current problem of L2 input in China and can be widely applied in classroom teaching and individual learning.

References

Altenberg, E. P. & Vago, R. M. (1987). Theoretical implications of an error analysis of second language phonology production. In Ioup, G. & Weinberger, S. H. (eds.), *Interlanguage phonology: the acquisition of a second language sound system*, 148-164. Cambridge, MA: Newbury House Publishers.

Another Tongue. (2015). *Miranda Hart*. http://www.anothertongue.com/artist/miranda-hart [Retrieved on 15 April, 2015].

Bassetti, B. (2009). Orthographic input and second language phonology. In Piske, T. & Young-Scholten, M. (eds.), *Input matters in SLA*, 191-206. Clevedon: Multilingual Matters.

Benson, B. (1988). Universal preference for the open syllables as an independent process in interlanguage phonology. *Language Learning* 38, 221-235.

Cao, D. 曹迪 (2012). 语言教育政策与文化软实力 [Language education policy and cultural soft power]. *语文学刊/ Yuyan xuekan* 4, 7-8.

Carroll, S. (2007). Autonomous induction theory. In VanPatten, B. & Williams, J. (eds.), *Theories in second language acquisition: an introduction*. Mahwah, NJ: Lawrence Erlbaum Associates.

Chai, R. 柴瑞琴 & Lu, Y. 卢玉芳 (2009). 二语习得课堂中的教师话语分析 [Analysis of teacher talk in the L2 classroom]. *长春理工大学学报(社会科学版) Changchun ligong daxue xuebao (shehui kexue ban)* 22(4), 648-649.

Chen, J. 陈君君 (2011). 二语学习者英语语音的石化现象研究 [Research on the English phonological fossilization of L2 learners]. *黑龙江教育学院学报 Heilongjiang jiaoyu xueyuan xuebao* 30(7), 147-148.

Cruttenden, A. (2014). *Gimson's pronunciation of English.* Oxon: Routledge.

Davis, C. & Kim, J. (2001). Repeating and remembering foreign language words: implications for language teaching systems. *Artificial Intelligence Review* 16, 37-47.

Deng, L. 邓利蓉 (2007). 教师话语在二语习得课堂上的作用 [The use of teacher talk in the L2 classroom]. *西昌学院学报(社会科学版) Xichang xueyuan xuebao (shehui kexue ban)* 19(2), 142-145.

Deterding, D. (2000). Measurements of the /eɪ/ and /əʊ/ vowels of young English speakers in Singapore. In Brown, A, Deterding, D. & Low, E. L. (eds.), *The English language in Singapore: research on pronunciation,* 93-99. Singapore: Singapore Association for Applied Linguistics.

Ellis, N. (2009). Input and second language acquisition: the roles of frequency, form, and function introduction to the special issue. *The Modern Language Journal* 93(3), 329-335.

Ellis, R. (1985). *Understanding second language acquisition*. Oxford: Oxford University Press.

Ellis, R. (1994). *The study of second language acquisition.* Oxford: Oxford University Press.

Fanelli, G., Gall, J. & Gool, L. J. V. (2009). Hough transform-based mouth localization for audio-visual speech recognition, in '*British Machine Vision Conference*'. British Machine Vision Association.

Flege, J. (2009). Give input a chance! In Piske, T. & Young-Scholten, M. (eds.), *Input matters in SLA*, 175-190. Clevedon: Multilingual Matters.

Gass, S. (1997). *Input, interaction and the second language learner.* Mahwah, NJ: Lawrence Erlbaum Associates.

Hansen, J. G. (2006). *Acquiring a non-native phonology: linguistic constraints and social barriers.* London: Continuum.

Henzl, V. (1979). Foreigner talk in the classroom. *International Review of Applied Linguistics* 17, 159-165.

Jenkins, J. (2000). *The phonology of English as an international language*. Oxford: Oxford University Press.

Jia, H. 贾红安 (2011). 浅谈当代大学英语教师的综合素质 [Brief discussion of essential qualities of contemporary college English teachers]. *价值工程 Jiazhi gongcheng* 28, 162.

Johnson, K. (1997). Speech perception without speaker normalization. In Johnson, K. & Mullennix, J. (eds.), *Talker variability in speech processing,* 145-165. San Diego: Academic Press.

Karakas, A. & Saricoban, A. (2012). The impact of watching subtitled animated cartoons on incidental vocabulary learning of ELT students. *Teaching English with Technology* 4, 3-15.

Klima, E. & Bellugi, U. (1979). *The signs of language*. Cambridge, MA: Harvard University Press.

Krashen, S. (1982). *Principles and practice in second language acquisition.* Oxford: Pergamon Press.

Krashen, S. (1985). *The input hypothesis: issues and implications.* New York: Longman.

Krashen, S. (2003). *Explorations in language acquisition and use: the Taipei lectures*. Portsmouth: Heinemann.

Li, H. 李卉 (2011). 高校英语课堂教师话语: 问题与对策 [Teacher talk in higher education English classes: problems and solutions]. *内蒙古农业大学学报(社会科学版) Neimenggu nongye daxue xuebao (shehui kexue ban)* 13(1), 144-145.

Li, L. 李力 & Wang, H. 王海波 (2006). 大学英语口语教学问题的探讨与对策 [A discussion of oral English teaching problems and counter-measures in colleges]. *浙江万里学院学报 Zhejiang Wanli xueyuan xuebao* 19(4), 149-151.

Li, R. 李瑞 (2009). 浅析英语课堂上教师的语言输入 [Brief discussion on teacher's language input in English classrooms]. *内江科技 Neijiang keji* 4, 60-61.

Li, X. 李霞 (2007). 教师话语及其对英语课堂教学的影响 [The influence of teacher talk in the EFL classroom]. *沈阳农业大学学报(社会科学版) Shenyang nongye daxue xuebao (shehui kexue ban)* 9(4), 580-583.

Liu, X. 刘晓明 (2010). 非英语专业大学生语言输入现状调查与研究 [An investigation of the current situation concerning language input among non-English major students in Chinese universities]. *科技信息 Keji xinxi* 29, 240-241.

Loschky, L. (1994). Comprehensible input and second language acquisition. *Studies in second language acquisition* 16, 303-323.

Macaluso, E., George, N., Dolan, R., Spence, C. & Driver, J. (2004). Spatial and temporal factors during processing of audiovisual speech: a PET study. *Neuroimage* 21(2), 725-732.

Moyer, A. (2009). Input as a critical means to an end: quantity and quality of experience in L2 phonological attainment. In Piske, T. & Young-Scholten, M. (eds.), *Input matters in SLA*, 159-174. Clevedon: Multilingual Matters.

Neuman, S. & Koskinen, P. (1992). Captioned television as comprehensible input: effects of incidental word learning from context for language minority students. *Reading Research Quarterly* 27(1), 94-106.

Niyogi, P., Petajan, E. & Zhong, J. (1999). Feature based representation for audio-visual speech recognition. In *proceeding of audio-visual speech processing*, Santa Cruz, USA.http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.399.3465&rep=rep1&type=pdf [Retrieved on 15 April, 2015].

Ortega-Llebaria, M., Faulkner, A. & Hazan, V. (2001). Auditory–visual L2 speech perception: effects of visual cues and acoustic–phonetic context for Spanish learners of English. *Speech, Hearing and Language: work in progress* 13, 39–51.

Pierrehumbert, J. B. (2001). Exemplar dynamics: word frequency, lenition and contrast. In Bybee, J. & Hopper, P. (eds.), *Frequency and the emergence of linguistic structure,* 137-158. Amsterdam: John Benjamins.

Piske, T. & Young-Scholten, M. (2009). *Input matters in SLA*. Clevedon: Multilingual Matters.

Plass, J. & Jones, L. (2005). Multimedia learning in second language acquisition. In Mayer, R. (ed.), *The Cambridge Handbook of Multimedia learning,* 467-488. Cambridge: Cambridge University Press.

Rast, R. (2008). *Foreign language input: initial processing.* Clevedon: Multilingual Matters.

Schwartz, B. (1993). On explicit and negative data effecting and affecting competence and linguistic behaviour. *Studies in Second Language Acquisition* 15, 147-163.

Scovel, T. (1998). *A time to speak: a psycholinguistic inquiry into the critical period for human speech.* Cambridge, M.A.: Newbury House.

Scovel, T. (2000). A critical review of critical period research. *Annual Review of Applied Linguistics* 20, 213-223.

The Ministry of Education, P. R. China. (2011). *English course standard of compulsory education.* Beijing: Beijing Normal University Press.

Tu, J. 屠静 (2012). 教师话语对中学二语习得课堂教学的影响 [The influence of teacher talk on second language teaching in middle schools]. *中国科教创新导刊 Zhongguo kejiao chuangxin dao kan 8, 154.*

VanPatten, B. (2004). Input processing in second language acquisition. In VanPatten, B. (ed.), *Processing instruction: theory, research and commentary.* Mahwah, NJ: Lawrence Erlbaum Associates.

Wells, J. C. (1982). *Accents of English*. Cambridge: Cambridge University Press.

Xiang, T. 项湍舟 (2009). 英语语音学习过程中的"石化现象 [Fossilization in English pronunciation learning]. *宜春学院学报 Yichun xueyuan xuebao* 31(1), 149-151.

Xu, L. 徐丽仙 & Guo, H. 郭弘 (2009). 教师话语研究及其对英语课堂教学的启示 [Research on teacher talk and its implications for English classroom teaching]. *内蒙古师范大学学报(教育科学版) Neimenggu shifan daxue xuebao (jiaoyu kexue ban)* 22(3), 112-115.

Yang, Y. 杨艳慧 & Ren, T. 任婷 (2009). 大学英语精读课教师话语研究 [Research on teacher talk in college English intensive reading]. *吉林工程技术师范学院学报 Jilin gongcheng jishu shifan xueyuan xuebao* 25(2), 6-8.

Young-Scholten, M. (1995). The negative effects of 'positive' evidence on L2 phonology. In Eubank, L., Selinker, L. & Sharwood Smith, M. (eds.), *The current state of interlanguage*, 107-122. Amsterdam: Benjamins.

Yuksel, D. & Tanriverdi, B. (2009). Effects of watching captioned movie clip on vocabulary development of EFL learners. *The Turkish Online Journal of Educational Technology* 8, 48-54.

Zhang, J. 张洁君 (2009). 大学英语课堂教师话语存在的问题及解决策略 [The problems of teacher talk in college English classroom teaching and the solutions]. *考试周刊 Kaoshi zhoukan* 3(3), 126-127.

Zhang, Y. 张尧 (2013). 英语电影对大学生学习英语的重要性 [The importance of English movies to English learning among college students]. *创新教育 Chuangxin jiaoyu* 6, 60.

Zheng, X. R. 郑鲜日 & Zheng, X, J. 郑晓杰 (2011). RP 英语和 PTH 汉语双元音声学特征对比研究 [Research on the comparison of the acoustic characteristics of diphthongs between RP English and PTH Chinese]. *东疆学刊 Dongjiang Journal* 28(3), 99-105.

Chen ZHANG
School of English
University of Sheffield
4.80b Jessop West
1 Upper Hanover Street
SHEFFIELD
S3 7RA
UK

czhang17@sheffield.ac.uk