



The
University
Of
Sheffield.

Shef*field*
Ling*uistics*
P*ost-*
G*raduate*
C*onference*

track changes

The Postgraduate Journal of the
Faculty of Arts & Humanities

Special Issue: Inaugural ShefLing PGC Proceedings

4-5 February 2014



The
University
Of
Sheffield.

The
Alumni
Foundation.

Track Changes Special Issue: ShefLing PGC Proceedings

About ShefLing

ShefLingPGC is organised by postgraduate students from across all departments of language and linguistics. This conference is designed to give linguistics postgraduates from all research areas an opportunity to present and discuss their research.

Student Organising Team

Sarah Gelcich (School of English & School of Medicine)

Helen Geyer (Cognitive Studies)

Hannah Leach (School of English)

Dario Lecic (Russian and Slavonic Studies, School of Languages and Cultures)

Suk May Low (School of English)

Ayumi Nonomiya (School of English)

Chen Zhang (School of English)

Editorial Board

Duncan B Burnes (School of English)

Richard Finn (School of English)

Sean McCorry (School of English)

Anthony C Prince (School of English)

Daniel I Southward (School of English)

Christine Wallis (School of English)

Track Changes Special Issue: ShefLing PGC Proceedings

Copyright Information

All rights reserved. Apart from fair use for the purpose of research, private study, criticism, or review, no part of this publication may be reproduced, translated, stored, or transmitted in any form. It is the Author's responsibility to secure the necessary copyright permissions for any images/maps/diagrams, etc. that have been reproduced within their work. Copyright for individual articles is held by the author.

University of Sheffield, 2014

Track Changes Special Issue:

ShefLing PGC Proceedings

Contents

Global Languages in Local Spaces: Rethinking	1-13
Multilingualism in the Linguistic Landscape of Toulouse	
Will Amos	
Movement and Direction in the Chinese V-V Construction	14-28
Zhishuang Chen	
Multilingual Sheffield? A Micro-Study of the Linguistic Landscape of London Road, Sheffield	29-41
Kirsten Ehinger	
Tracing the Emergence of Sociolinguistic Variation in L2 Speech: Polish Migrants in Edinburgh	42-54
Daniel Lawrence	

Bringing Theoretical L2 Acquisition Research Findings 55-65

to the Language Classroom: A Materials Development

Challenge

Elaine Lopez

A Contrastive Rhetorical Study of Critical Essays in 66-77

English by Tunisian EFL Students and American

Students

Enneifer Salwa

Dubbing Can Be an Agent in Pragmatic Language 78-89

Change: Evidence from the Use of *esatto* in Italian as

an Agreement Marker

Lorella Viola

@Sum1, #Everyone: A Corpus-Based Analysis of 90-101

Language Performances and Attitudes on Twitter

Matthew Voice

Global Languages in Local Spaces: Rethinking multilingualism in the linguistic landscape of Toulouse

Will Amos

University of Liverpool

H.W.G.Amos@Liverpool.ac.uk

1. Introduction

This paper explores data collected in 2013 for a MA on the Linguistic Landscape (hereafter LL) of Toulouse. Over the past 15 years, the Linguistic Landscape has emerged as an innovative way to evaluate languages in contact in public spaces. First used to investigate the reactions of Canadian school pupils to signs written in French and English (Landry & Bourhis, 1997), the term has since been incorporated into larger-scale studies of multilingualism on signs (advertisements, billboards, posters, shop names, etc.) in towns and cities throughout the world.¹ In much of this research, English has recurrently appeared as one of the most common languages visible in the LL.

This has led many researchers to dub English an ‘international language’ (Cenoz & Gorter, 2009; Huebner, 2006; Piller, 2001, 2003). This description is significant in the ‘world languages’ discussion, in which Quirk (1985, 1990) and Crystal (1997) argue that standard English is spreading throughout the world, transcending linguistic, cultural, and national borders. In a reactive argument emerging in tandem, Kachru (1985, 1986, 1992) persistently makes the case for what he describes as many diverse ‘World Englishes’, by which he claims that the spread of English is subject to local variation. This research adopts the position that the former depiction of a single, global language only applies in a non-descriptive sense. In other words, whilst in abstract terms the idea of English (for this we can argue the sociocultural value that is

¹ See edited volumes by Shohamy & Gorter (2009); Shohamy, Ben-Rafael, & Barni (2010); Gorter, Marten, & Van Mensel (2012); and Hélot, Barni, Janssens, & Bagna (2012).

often attached to it) may be ‘spreading’, its modes of expression clearly take on many structural guises, and are subject to distinct variation.

Whichever position the researcher takes, it is undoubtedly a concern that many LL studies do not qualify English beyond the description of a ‘global language’. English is recurrently found to be the second or third most common language in any given LL, where its visibility has been variously explained in terms of prestige and style (Ross, 2008); modernity (Kasanga, 2010); creativity and humour (Mettewie, Lamarre, & Van Mensel, 2012); success and sophistication (Piller, 2001, 2003); and wealth (Dimova, 2007). Elsewhere, Kelly-Holmes (2000), Huebner (2006), Edelman (2009), and others speak of the ‘fetishisation’ of English, in which it is used to promote ideas, ideals, lifestyles, and products by appealing to potential consumers through its connotation with Anglo-American culture.

2. Reconceptualising multilingualism

As the LL continues to expand beyond the traditional boundaries of language to incorporate the study of diverse semiotic resources, the issue of connotations within languages has begun to spark interest. In wider sociolinguistics, the introduction of terms such as ‘metrolingualism’ (Otsuji & Pennycook, 2010) and ‘polylanguaging’ (Jørgensen, Karrebæk, Madsen, & Møller, 2011) has challenged the traditional boundaries between what we consider to be languages. Within the LL, the issue of language coding likewise remains underdeveloped: although the research methodology for the MA project classified signs according to nine types — business names, business signs, graffiti, information, instructions, labels, legends or slogans, street signs, and trademarks — these were contextual distinctions, and made no provision for analysing signs featuring elements of multiple languages simultaneously. Whilst the methodology measured multilingualism according to the degree of translation between the languages, the dataset included a significant number of signs whose languages were difficult to interpret. One of the aims of this paper, therefore, is to exemplify some of the shortfalls of many existing methodological models, and posit some possible directions for future research into language coding.

3. The data

The data collected for this project found English to be the second most visible language in the Toulouse LL. Out of a corpus of 770 signs, English was visible on 16% of them, and it featured on almost three quarters of multilingual signs. English and French was the most common combination of languages on multilingual signs (see fig. 1). In terms of sign types, the English-French combination was seen predominantly on information signs (28%), slogans (10%), and business names (8.5%), in light of which we can deduce that English is used chiefly outside the boundaries of informational and instructional contexts.

FIGURE 1: PRESENCE OF LANGUAGES ON MONO- AND MULTILINGUAL SIGNS (%)

	Monolingual	Multilingual	Total
French	87.9	91.5	87.7
English	9.8	73.2	15.7
Occitan	0.1	16.9	1.7
Japanese	0.6	8.5	1.3
Italian	0.4	4.2	0.8
Arabic	0.3	2.8	0.5
German	0	4.2	0.4
Latin	0.4	0	0.4
Portuguese	0	2.8	0.3
Irish	0.1	1.4	0.3
Catalan	0.1	0	0.1
Spanish	0.1	0	0.1
n	694	71	770 ²

These data suggest that the English in the Toulouse LL is not aimed at internationals or foreigners (Huebner, 2006; Smalley, 1994); but rather that its tokenistic cultural connotations (Bhatia, 1992; Schlick, 2002) mark it out as written for consumption by the predominantly French-speaking monolingual public.

² Five signs in the dataset featured illegible graffiti, and it was unclear which or how many languages they contained.

FIGURE 2: IDÉAL MODELS

Fig. 2 illustrates how sign writers use code mixing to incorporate elements of English into their writing. The acute accent in ‘Idéal’ identifies the word as French, yet ‘models’ is missing both a grave accent and a second ‘e’, and could therefore be classified as English. The word order (adjective preceding noun) is also markedly un-French, which further intensifies the English flavour to the sign. Whilst these factors may lead us to treat this sign as English, or as moderately bilingual, however, other signs featuring the shop name call this classification into doubt:

FIGURE 3: DOOR STICKER

FIGURE 4: OPENING TIMES

The apostrophe in fig. 3 suggests a possible desire for an Anglicised business name, ‘apostrophe-s’ being a typically English grammatical marker. However, the absence of an apostrophe in the same business name in fig. 4 complicates any assumptions we can make about a uniform strategy for the shop’s name. Furthermore, the treatment of French on the bottom line of fig. 4, where the days of the week are atypically capitalised, is ambiguous, which leads us even to question the French classification of the sign. Despite the suggestion that this shop and its signs contain English elements, it is unclear whether this capitalisation indicates a deliberate decision to project English stylistic features onto French, or a deviation from the norms of standard French (intended or accidental).

The traditional notion of multilingualism, where multiple languages appear alongside one another, is further challenged in this next example:

FIGURE 5: MY JEAN IS MY BOYFRIEND

Fig. 5 shows part of an advertisement visible in the window of *Jennyfer*, a mid-priced clothing chain found in many major towns and cities in France. Following the research methodology this item was recorded as multilingual, though its treatment of language does not evince a clear distance between both codes. An initial glance suggests that this is an example of a common practice in French advertising, where a sign is written in English then translated into French via an asterisk, in order to conform to official legislation which prescribes the use of French on all public inscriptions. This would imply that we categorise the sign as multilingual. Closer analysis, however, reveals that the English assumes a French identity, superseding its role as a marker of Anglophone cultures (Bhatia, 1992; Schlick, 2002). Indeed, this sign exhibits a play-on-words that is delivered in English, but which celebrates a linguistic function that is undoubtedly French. The French word ‘jean’ is a homonym which refers both to the trouser and the common male forename *Jean*. This sign was first interpreted as an example of what Piller (2003: 173) terms ‘ludicrous’ English: given that the trouser is normally pluralised in standard English (‘jeans’), I considered this an inaccurate pun that any competent English speaker would deem invalid. From the perspective of a French-speaker, however, ‘jean’ resonates in French, and its etymology (and modification to the singular) is inconsequential. Thus, whilst the phrase ‘my jean is my boyfriend’ may appear English throughout, the specificities of the pun are markedly French. On the one hand we can describe this as bilingual creativity; on the other as the reimagining of an English term in a French construction, a process Kachru (1992: 59) refers to as ‘acculturation’. This covert superiority of one code over another challenges current research, which often goes no further than classifying such items as ‘multilingual’.

FIGURE 6: SO COLISSIMO

Similarly, fig. 6 features the trademark ‘Colissimo’, which indicates the most rapid parcel delivery service offered by *La Poste*, the national postal service. The brand name is a play on words where the French ‘colis’ refers to a parcel and the Italian suffix ‘-issimo’ implies a superlative. The intention is to suggest that this is the greatest, best, and fastest parcel service, further reinforced by the English adverb ‘so’. It is impossible to speculate how far readers might recognise the Italian and English features on this sign, but it does indicate the choice of *La Poste* to include multilingual elements in its advertising strategies.

In terms of methodology, it is unclear how this item should be coded. The most straightforward solution might be to count it as French: it is in the window of a French shop, displaying the well-known emblem of a French company, in France. However, given its Italian and English features, ‘so Colissimo’ is arguably only one third French. We must consider, therefore, how far the Italian and English features remain tied to their languages of origin, or nativised into the host language. In other words, we must

ask whether this sign evinces any ‘collocation’ (Barthes, 1977: 33) with these two linguistic cultures, or whether this link is an intangible ‘myth’, as Berger (1972: 140) has argued.

It is my contention that this sign does not exemplify English and Italian as native-speakers might understand them, but rather a localised — or French — ‘foreignness’ that belongs within the expanding French linguistic paradigm. ‘So Colissimo’ contains elements of foreign languages chosen to be aimed specifically at French readers. These may be termed clichéd semiotic artefacts. This can be particularly attributed to the word ‘so’, which also appears on other items in the dataset:

FIGURE 7: SO EXTREME



FIGURE 8: SO CUTE



The use of ‘so’ does not necessarily index English-speaking cultures — as to what nations, peoples, or social environments this could refer is itself difficult to determine — but rather a sense of cosmopolitanism, both for the products and the intended

consumer. The signs in figs. 6, 7, and 8 are not English or Italian; rather they use commercial-linguistic stereotypes that conjure ideas, concepts, and designs of transnational contexts. The argument that these linguistic elements occur within the boundaries of French chimes with Hill's (1995) description of 'mock language': the intensive use of the adjective 'so' may be designed to appeal to the multiculturalism of those who can read and understand it; yet it may also be intended to identify English as an undesirable Other, by subversively mocking it for its simplicity or triviality. That the English, particularly when printed in italics (fig. 7), reads almost as sarcastic to a native-speaker suggests some distance between the standard language and this particular brand of nativised English.

From the perspective of the sign author, it is possible to argue that as long as the provenance of nativised elements remains identifiable, 'incorrect' language does not detract significantly from the semiotic impact of a sign, where its connotation is more important than its meaning. Moreover, nativising foreign languages allows for a more creative range of linguistic inputs than those available within the structured boundaries of standard French. This implies that languages can combine in complex ways beyond the criteria of methodologies based on a 'language A OR language B' coding system.

4. Occitan

Nativisation not only concerns languages traditionally considered the vehicles of globalisation, but also the regional language Occitan, which is denied exclusivity in the LL by its dependence on French. This subordination is spatial, where it is consistently written below the official language; but also lexical, where the distance between the languages is never substantial:

FIGURE 9: RUE LAFAYETTE STREET SIGNS**FIGURE 10: RUE GENTY-MAGRE STREET SIGNS**

Any argument that these two languages are competing or mutually exclusive is unconvincing, since it is reasonable to assume that the many French monolinguals in contemporary Toulouse understand most or all of the information on these Occitan signs. The French-speaker is aware that this is not his language, and that the sign symbolises a different linguistic culture; yet at the same time it is able to speak to him, and thus the linguistic gap between French and Occitan is bridged. Moreover, we can reason that Occitan — whose presence is legally permitted only when it appears alongside the official language, French — is deprived of any individuality, giving the impression that it exists as a variable of and within the French landscape. Thus, while we have seen that so-called ‘foreign’ languages may be localised into French, in the same way we can say that Occitan becomes nationalised.

5. Conclusion

These data illustrate that ‘multilingualism’ does not always signify a collection of independent codes (written side by side or separately) since languages in the LL combine and envelop one another. I propose, therefore, that ‘multilingualism’ is not defined exclusively in terms of vocabulary or lexicon. Borrowings, code-switching, and plays on words are all functions through which a supposedly ‘English’, ‘Italian’, or ‘Occitan’ sign may reveal its inherent ‘Frenchness’. We can therefore define a new brand of multilingualism, in which a range of semiotic elements combine to signify a blend of languages and linguistic cultures.

In terms of future directions, measuring the degree of ‘foreignness’ on individual signs offers interesting potential for further investigation as we seek to understand the ways multiple codes interact, and how they combine to form varying manifestations of what we call languages in our public spaces. Although the hegemony of French is largely uncontested in the Toulouse LL, the way multilingualism is treated and transformed constructs peoples’ perceptions of languages, their roles in various contexts, how the state manages them, and ultimately what we consider their value to be.

References

- Barthes, R. (1977). Rhetoric of the image. In S. Heath (Ed. & Trans.), *Image, music, text* (pp. 32–51). New York: Hill and Wang.
- Berger, J. (1972). *Ways of seeing*. London: Penguin Classics.
- Bhatia, T. K. (1992). Discourse functions and pragmatics of mixing: Advertising across cultures. *World Englishes*, 11(2-3), 195–215. doi:10.1111/j.1467-971X.1992.tb00064.x
- Cenoz, J., & Gorter, D. (2009). Language economy and linguistic landscape. In E. Shohamy & D. Gorter (Eds.), *Linguistic landscape: Expanding the scenery* (pp. 55–69). London: Routledge.
- Crystal, D. (1997). *English as a global language*. Cambridge: Cambridge University Press.
- Dimova, S. (2007). English shop signs in Macedonia. *English Today*, 23(3-4), 18–24.

- Edelman, L. J. (2009). What's in a name? Classifications of proper names by language. In E. Shohamy & D. Gorter (Eds.), *Linguistic landscape: Expanding the scenery* (pp. 141–154). London: Routledge.
- Gorter, D., Marten, H. F., & Van Mensel, L. (Eds.). (2012). *Minority languages in the linguistic landscape*. Basingstoke: Palgrave Macmillan.
- Hélot, C., Barni, M., Janssens, R., & Bagna, C. (Eds.). (2012). *Linguistic landscapes, multilingualism and social change*. Frankfurt am Main: Peter Lang.
- Hill, J. H. (1995). Mock Spanish: A site for the indexical reproduction of racism in American English. In D. J. Glick (Ed.), *Language and culture, symposium 2*. Retrieved from <http://language-culture.binghamton.edu/symposia/2/part1/index.html>
- Huebner, T. (2006). Bangkok's linguistic landscapes: Environmental print, codemixing and language change. *International Journal of Multilingualism*, 3(1), 31-51.
- Jørgensen, J. N., Karrebæk, M. S., Madsen, L. M., & Møller, J. S. (2011). Polylanguaging in Superdiversity. *Diversities*, 13(2), 23–37.
- Kachru, B. B. (1985). Standards codification and sociolinguistic realism: The English language in the outer circle. In R. Quirk & H. G. Widdowson (Eds.), *English in the World: Teaching and learning the language and literatures* (pp. 11–30). Cambridge: Cambridge University Press.
- Kachru, B. B. (1986). *The alchemy of English: The spread, functions and models of non-native Englishes*. Oxford: Pergamon Press.
- Kachru, B. B. (1992). *The other tongue: English across cultures*. Chicago: University of Illinois Press.
- Kasanga, L. A. (2010). Streetwise English and French advertising in multilingual DR Congo: Symbolism, modernity, and cosmopolitan Identity. *International Journal of the Sociology of Language*, 206, 181–205.
- Kelly-Holmes, H. (2000). Bier, parfum, kaas: Language fetish in European advertising. *European Journal of Cultural Studies*, 3(1), 67–82.
- Landry, R., & Bourhis, R. Y. (1997). Linguistic landscape and ethnolinguistic vitality: An empirical study. *Journal of Language and Social Psychology*, 16(1), 23–49.

- Mettewie, L., Lamarre, P., & Van Mensel, L. (2012). Clins d'oeil bilingues dans le paysage linguistique de Montréal et Bruxelles: Analyse et illustration de mécanismes parallèles. In C. Hélot, M. Barni, R. Janssens, & C. Bagna (Eds.), *Linguistic landscapes, multilingualism and social change* (pp. 201–216). Frankfurt am Main: Peter Lang.
- Otsuji, E., & Pennycook, A. (2010). Metrolingualism: fixity, fluidity and language in flux. *International Journal of Multilingualism*, 7(3), 240–254.
- Piller, I. (2001). Identity constructions in multilingual advertising. *Language in Society*, 30(02), 153–186.
- Piller, I. (2003). Advertising as a site of language contact. *Annual Review of Applied Linguistics*, 23, 170–183.
- Quirk, R. (1985). The English language in a global context. In R. Quirk & H. G. Widdowson (Eds.), *English in the World: Teaching and learning the language and literatures* (pp. 1–6). Cambridge: Cambridge University Press.
- Quirk, R. (1990). Language varieties and standard language. *English Today*, 6 (1), 3–10.
- Ross, N. J. (2008). Signs of International English. *English Today*, 13(02), 29–33. Retrieved from http://journals.cambridge.org/abstract_S0266078400009597
- Schlick, M. (2002). The English of shop signs in Europe. *English Today*, 18(02), 3–7. doi:10.1017/S0266078402002018
- Shohamy, E., Ben-Rafael, E., & Barni, M. (Eds.). (2010). *Linguistic landscape in the city*. Bristol: Multilingual Matters.
- Shohamy, E., & Gorter, D. (Eds.). (2009). *Linguistic landscape: Expanding the scenery*. New York: Routledge.
- Smalley, W. A. (1994). *Linguistic diversity and national unity: Language ecology in Thailand*. Chicago: University of Chicago Press.

Movement and Direction in the Chinese V-V Construction

Zhishuang Chen
University of York
 zc643@york.ac.uk

1. Introduction

In Chinese, a verb sequence V_1-V_2 in which V_2 signals the result of V_1 is traditionally called a ‘resultative verb compound’ (RVC) (Li & Thompson 1981). For example, *da-si* ‘hit-die’.

(1) Ta **da-si** le lang.

he hit-die PERF wolf

‘He hit the wolf and the wolf died.’

The focus of this paper is a special kind of resultative verb compound, called ‘directional resultative verb compound’. In a directional RVC, V_1 denotes movement / displacement, while V_2 denotes a direction. For example, *song-lai* ‘send-come’.

(2) Ta **song-lai** le shu.

he send-come PERF book

‘He sent the book (towards the speaker).’

Note that the second verb *lai* means ‘come’ when it is used alone. However, when it is used in this verb sequence, it only denotes a direction which is ‘towards the speaker’.

More examples of directional RVCs are shown below.

(3) Ta **song-qu** le shu.

he send-go PERF book

‘He sent the book (away from the speaker).’

(4) Ta **zou-shang** le louti.

he walk-ascend PERF stairs

‘He walked up the stairs.’

- (5) Ta **zou-jin** le fangjian.
 he walk-enter PERF room
 ‘He walked into the room.’

Directional RVCs display many interesting characteristics which have not received significant attention in the contemporary literature. In Section two it will be shown that some of the directional RVCs can have two word orders when taking an object and that these word orders show different syntactic and semantic characteristics. In Section three, I will propose that the two word orders reflect different merge structures. With this approach the syntactic and semantic variation can be explained neatly.

2. The Puzzle

2.1 The Word Order Problem

Descriptive work by Li and Thompson (1981) has already stated that when non-directional RVCs take an object they only allow the object to follow the whole compound, as is shown below:

- (8) da-si lang
 hit-die wolf
 ‘hit the wolf and the wolf died’

This holds for some directional RVCs as well.

- (9) zou-shang louti
 walk-ascend stairs
 ‘walk up stairs’

However, when V₂ is *lai* ‘come’ or *qu* ‘go’, directional RVCs also allow the object to intervene between the two verbal elements, as is shown in (10).

- (10) a. song-lai tang
 send-come soup
 b. song tang lai

send soup come
 ‘bring the soup’

Thus we have two word orders here: $V_m V_d O$ and $V_m O V_d$.³

The descriptive work only uses one translation for both orders; it seems that the two orders have exactly the same meaning. Yet, by observing the data more carefully, I will demonstrate that the two orders are not equivalent in syntax or semantics.

2.2 Syntactic and Semantic Variation of the Two Word Orders

The syntactic and semantic variation of the two word orders can be illustrated from three perspectives: (1) compatibility with a location object; (2) position of perfective marker *le*; (3) telicity.

2.2.1 Compatibility with a location object

Interestingly, when the object denotes a location, only the $V_m O V_d$ order is grammatical, while the $V_m V_d O$ order becomes ungrammatical, as is shown in (11) and (12).

- (11) shang shan lai
 ascend mountain come
 ‘come up the mountain’
- (12) *shang lai shan
 ascend come mountain

Other location objects such as *xuexiao* ‘school’, *he* ‘river’, etc. also disallow the $V_m V_d O$ order. The right order should be: *chu xuexiao qu* ‘exit school go’, *guo he lai* ‘cross river come’.

³ V_m = movement verb; V_d = directional verb

2.2.2 Position of perfective marker *le*

le is a perfective aspect marker which occurs directly after verbs or predicative adjectives to indicate that the action has been realized.⁴ This *le* appears in different positions in the two word orders.

In the case of the V_m V_d O order, *le* should follow the whole compound. It cannot intervene between V_m and V_d.

- (13) Ta song-(**le*)-lai le tang
 he send-(**PERF*)-come PERF soup
 ‘He sent the soup here.’

In the case of the V_m O V_d order, *le* can occur immediately after the V_m, after the V_d or after both. (14) shows the case with *le* appearing after both V_m and V_d.

- (14) Ta song le tang lai le.
 he send PERF soup come PERF
 ‘He sent the soup here.’

2.2.3. Telicity

Telicity is a semantic property of a verb or verb phrase and it is one of the properties used to distinguish Vendler’s (1967) four situation types. A telic event has a natural finishing point, like ‘eat an apple’, ‘run a mile’. Any event which does not have a natural finishing point is atelic, like ‘walk’, ‘push a cart’.

The two orders seem to show the difference in telicity reading as well. See below:

- (15) Ta song-lai tang, #keshi hai mei dao.

⁴ In the literature there are two kinds of *le*. One which appears immediately after verbs is widely accepted as perfective marker, the other, which occurs at the end of sentences, is viewed as a sentence final particle. When the sentence finishes with an intransitive verb followed with a *le*, this *le* is viewed as ambiguous between aspect marker and sentence final particle. The *le* spoken about here is the aspect marker *le* (including the ambiguous one), meaning the *les* directly after verbs.

- he send-come soup, but still not arrive
 ‘He brought the soup, #but it has not arrived.’
- (16) Ta song tang lai, keshi hai mei dao.
 he send soup come, but still not arrive
 ‘He has come with the soup, but it has not arrived.’

I add the same assertion ‘but it has not arrived’, which provides a finishing point, at the end of (15) ‘send-come soup’ and (16) ‘send soup come’. (15) becomes infelicitous while (16) does not. As can be seen from the translation, ‘send come soup’ indicates that the soup has arrived while ‘send soup come’ only indicates the subject has set off to send the soup; he may have arrived or be on the way. This is the first time in this paper that I show the translations of the two orders are not exactly the same, differing from the previous descriptive works. I further suggest that the clash with the assertion containing a finishing point is due to the fact that ‘send-come soup’ already entails a finishing point while ‘send soup come’ does not. In other words, the $V_m V_d O$ order is telic whilst the $V_m O V_d$ order is atelic. To prove this, we need some tests.

Dowty (1979) proposed a few tests for telicity and I have adapted some of them here. The first is the in-PP test. Basically, in-PP is grammatical with telic events but ungrammatical with atelic events. For example:

- (17) John painted a picture in an hour.

- (18) *John walked in an hour.

Applying it to our data, namely *song-lai tang* and *song tang lai*, we get:

- (19) Ta yi-xiaoshi-nei song lai tang.
 he one-hour-in send come soup
 ‘He brought the soup in one hour.’
- (20) *Ta yi-xiaoshi-nei song tang lai.
 he one-hour-in send soup come

‘He came with the soup in one hour.’⁵

Thus, the in-PP test tells us that the $V_m V_d O$ order has a telic reading whilst the $V_m O V_d$ order has an atelic reading.

The second method of testing is the entailment test. According to Dowty, if ϕ is a telic verb/verb phrase, then $x \phi ed$ for y time does not entail that $x \phi ed$ was true during any time within y .

(21) John painted a picture for an hour.

‘John painted a picture’ is not necessarily true at any time during that hour.

However, if ϕ is an atelic verb/verb phrase, then $x \phi ed$ for y time entails that any time during y , $x \phi ed$ was true.

(22) John walked for an hour.

‘John walked’ is true at any time during that hour.

Applying this to our data, we have:

(23) Ta song-lai tang song le yi-xiaoshi.

he send-come soup send PERF one-hour

‘He spent an hour bringing the soup.’⁶

‘He send-come soup’ was not necessarily true at any time during that one hour.

(24) Ta song tang lai song le yi-xiaoshi.

he send soup come send PERF one-hour

‘He has been bringing the soup for an hour.’

‘He send soup come’ was true at any time during that one hour.

⁵ (20) is grammatical only when it means ‘He will come with the soup in one hour.’, i.e. when ‘in one hour’ means ‘one hour from now’. But the in-PP test cares about the past reading of the in-PP. In this sense (20) is ungrammatical. Another way to make (20) grammatical is by inserting a perfective *le* after *sang*. This is probably because *le* transfers the sentence from atelic to telic.

⁶ Chinese for-PP structure is constructed by copying the first verb and then adding a temporal NP.

The entailment test gives the same result as the in-PP test.

The last test I use is the ‘almost’ test. The adverb ‘almost’ has different effect on telic/atelic events.

- (25) John almost painted a picture.

A telic event like (25) gives two readings with ‘almost’: Firstly, that John had the intention of painting a picture, but then changed his mind and did nothing at all. Or, Secondly, John did begin working on the painting and he almost but not quite finished it.

- (26) John almost walked.

An atelic event with ‘almost’ like (26) only has one reading: John did not walk at all.

We insert the Chinese ‘almost’ *jihu* into the two orders:

- (27) Ta jihu song-lai tang le.
 he almost send-come soup SFP⁷
 ‘He almost brought the soup.’

Reading: either He did not do the ‘soup-sending’ at all or He did start out sending the soup, but it did not arrive for some reason.

- (28) Ta jihu song tang lai le.
 he almost send soup come SFP
 ‘He almost came to send the soup.’

Reading: He did not do the ‘soup-sending’ at all.

The ‘almost’ test shows the same result as the other two tests.

⁷ SFP refers to sentence final particle.

So far all the three telicity tests confirm my point that the $V_m V_d O$ order is telic whilst the $V_m O V_d$ order is atelic

2.3 Interim Summary

I summarize the findings obtained for the syn-sem variation of the $V_m V_d O$ order and the $V_m O V_d$ order below:

(29)

	Compatibility with location object	Position of <i>le</i>	Telicity reading
$V_m V_d O$	NO	$V (le^*)V le O$	telic
$V_m O V_d$	YES	$V (le) O V (le)$	atelic

Based on this table, two questions arise: How are the two word orders derived? How can we account for the syntactic and semantic differences observed?

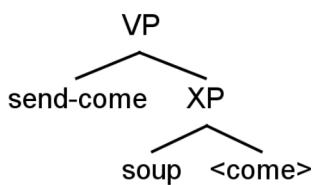
3. Analysis

3.1 Two Merge Structures

Having seen the characteristics the $V_m V_d O$ order and the $V_m O V_d$ order present, I hypothesize that the two orders reflect different original merge structures.

Specifically, when presenting the $V_m V_d O$ order, V_d merges with O first and then raises to V_m . Therefore for a phrase like *song lai tang* ‘send come soup’, minimally there should be a merge structure like this:

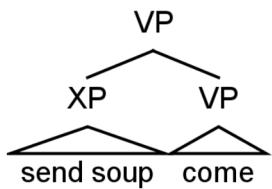
(30)



‘Soup’ and ‘come’ merge first, projecting an XP which can be interpreted as a result phrase. In other words the structure is interpreted as ‘someone sends the soup and the soup comes’.

When presenting the $V_m O V_d$ order, V_d is the main predicate while the constituent consisting of $[V_m O]$ is an adjunct adjoined to V_d . So a phrase like *sing tang lai* ‘send soup come’ would have a merge structure like this:

(31)



The raise of ‘come’ in (30) could be attributed to the same reason that resultative predicates raise, suggested by Sybesma (1999).

I will not go into these details in this paper because this minimal structure is enough to account for the syntactic and semantic variation.

3.2 Account for the Syn-sem Variation

3.2.1 Compatibility with Location Object

We have already seen in (11) and (12) that when the object is a location, only the $V_m O V_d$ order is available. Why is that?

Based on my hypothesis, the ungrammatical order ‘ascend come mountain’ should have this merge structure:

(32) [VP ascend [XP mountain come]]

There is a problem in this derivation. ‘Mountain’ cannot be an argument of ‘come’ because in that case, (32) would mean something like ‘someone ascends the mountain and the mountain comes’, which does not make sense. This differs from the non-

location object case like ‘send come soup’ which can be interpreted as ‘...send a soup and the soup comes’. The only possible argument for ‘come’ in (32) should be the subject of the sentence. Therefore, ‘mountain’ and ‘come’ cannot merge in the first place, thus the derivation crashes, which results in the ungrammaticality of the $V_m V_d O$ order.

In the case of the grammatical order ‘ascend mountain come’, however, the merge structure is like below:

- (33) [VP [XP ascend mountain] [VP come]]

(33) is actually interpreted as ‘someone comes ascending the mountain’. ‘Ascend’ and ‘mountain’ form an adjunct in which ‘mountain’ is selected by ‘ascend’ as its internal argument.⁸ The derivation can converge successfully. Therefore, the $V_m O V_d$ order is grammatical.]

3.2.2 Position of *le*

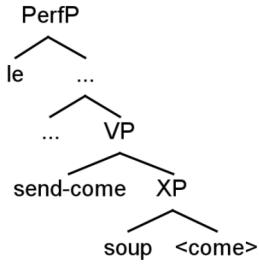
I will now move on to explain why *le* cannot intervene between V_m and V_d in the case of the $V_m V_d O$ order, as (13) shows, which I repeat below:

- (34) Ta song-(**le*)-lai le tang.
 he send-(**PERF*)-come PERF soup
 ‘He sent the soup here.’

The tree diagram for (34) is as follows:

⁸ The Chinese *shang* ‘ascend’ is a transitive verb

(35)



According to my proposal, we first construct VP and raise ‘come’ to adjoin to ‘send’. Then we continue to merge *le* projecting PerfP. At this point, we do not have any chance to insert *le* in between ‘send’ and ‘come’. The only possible recourse is to raise the whole complex verb, thus demonstrating the order of ‘send come *le*’.

However, we also observe that in the case of the $V_m O V_d$ order, *le* can occur immediately after the V_m , after the V_d or after both. I use the example (14) with two *les* (repeated below) to illustrate this point.

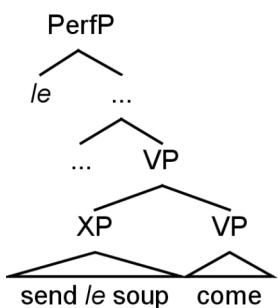
(36) Ta song le tang lai le.

he send PERF soup come PERF

‘He sent the soup here.’

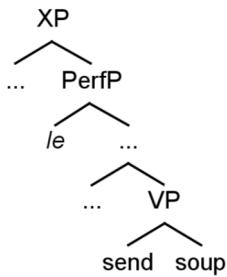
The tree diagram (36) is as below:

(37)



As (37) shows, I suggest that the *le* immediately following ‘send’ is situated in the adjunct XP. This XP could be a TP or something big enough to contain another PerfP. That is to say, the structure under XP could be something like this:

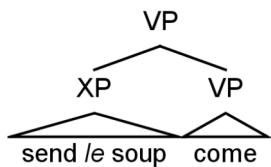
(38)



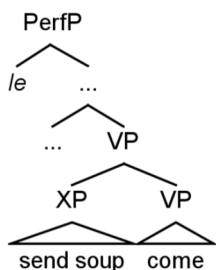
The *le* coming after ‘come’ is from the Perf head in the matrix sentence. Eventually, the whole VP will raise and derive the word order ‘send *le* soup come *le*’.

When only one *le* appears, it is because one of the PerfP is not projected, as is shown in (39) and (40).

(39)



(40)

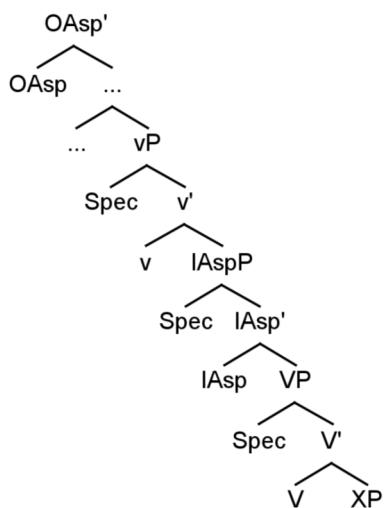


3.2.3 Telicity

How, then, to account for the fact that ‘send come soup’ is telic while ‘send soup come’ is atelic?

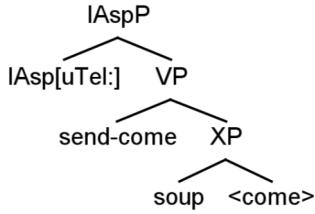
It is well known that the aspect system of human languages consists of two types: viewpoint aspect (grammatical aspect) and situation aspect (*aktionsarten*) (Smith 1997). Travis (2010) proposes that both viewpoint aspect and situation aspect are realized as a head. The situation aspect is realized as a head IAsP (Inner aspect) while the viewpoint aspect is realized as a head OAsp (Outer aspect). Furthermore, she suggests IAsp head carries a feature [+/- telic] whose value is computed by the elements within its domain. This value would decide the telicity reading. Below is a reduced version of Travis' structure for the two layer aspect:

(41)

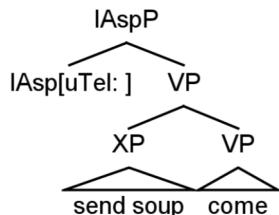


I have adapted Travis' idea of IAspP and assumed that the IAsp head bears an uninterpretable feature [uTel:] which needs to be valued by something below. The value it receives will eventually decide the telicity reading. Combined with my proposal, we enrich our original merge structures (30) and (31) into (42) and (43):

(42)



(43)



Now it is possible to explain the variation in telicity reading. I suggest that there are different ‘come’s in modern Chinese, which are different lexical items with the same pronunciation. The example shown in (42) actually means something like ‘come and arrive’. In other words, this ‘come’ entails a result. So we can assume that this ‘come’ bears a feature [res] which will value [uTel:] as [uTel: res], thus giving a telic reading. On the other hand, the ‘come’ in (43) means something like ‘move towards the speaker’, so it does not entail a result. Therefore, it bears a different feature, say [unbounded], which values [uTel:] as [uTel: unbounded], giving an atelic reading eventually.

4. Conclusion

To conclude, this paper focuses on two possible word orders of directional RVCs with *lai* ‘come’ / *qu* ‘go’ as the directional verb, which are the $V_m V_d O$ order and the $V_m O V_d$ order. I argue against the descriptive works that these two orders are not equivalent by illustrating three different syntactic and semantic characteristics, which are compatibility with a location object, position of *le* and telicity reading. To account for the syn-sem variation, I propose that the two word orders reflect two merge structures. In the $V_m V_d O$ order, V_d merges with O first and then raises to V_m . While in the $V_m O V_d$ order,

V_d order, V_d is the main predicate with $[V_m O]$ as an adjunct modifying it. With this proposal we neatly explain the differences between the two word orders.

References

- Dowty, D. R. (1979). *Word meaning and Montague grammar: The semantics of verbs and times in generative semantics and in Montague's PTQ*, Vol. 7, Springer.
- Li, C. & Thompson, S. (1981), *Mandarin Chinese: A functional reference to Chinese grammar*, Berkeley: University of California Press.
- Smith, C. S. (1997), *The parameter of aspect*, Vol. 43, Springer.
- Sybesma, R. P. (1999), *The mandarin VP*, Kluwer Academic Publishers.
- Travis, L. (2010), *Inner aspect: The articulation of VP*, Vol. 80, Springer.
- Vendler, Z. (1967), *Linguistics in philosophy*, Cornell University Press Ithaca.

Multilingual Sheffield?

A Micro-Study of the Linguistic Landscape of London Road, Sheffield

Kirsten Ehinger

University of Sheffield

k_ehinger@yahoo.com

Context

The study of linguistic landscapes is a relatively new method of assessing the status and function of language use in the public space. It is based on the results of a landmark study conducted by Landry and Bourhis in Canada and published in 1997, which found that the language that surrounded native French-speaking high school students in their daily lives had a profound impact on the way they used and how they perceived the value of their mother tongue. The paper defined the linguistic landscape as ‘the language of public road signs, advertising billboards, street names, place names, commercial shop signs, and public signs on government buildings [...] of a given territory, region or urban agglomeration’ (Landry, p.25), and after finding a link between the surrounding linguistic landscape and individual speakers’ language use it concluded that ‘ethnolinguistic groups have a vested interest in promoting the use of their own-group language within the linguistic landscape’ (Landry, p. 45). Since then many studies have been conducted into many different aspects of linguistic landscapes. The term was re-defined by Ben-Rafael et al. as ‘linguistic objects that mark the public space’ (p.7), which is a slightly broader definition than the one followed in this study.

Linguistic landscape studies generally fall into one of two categories. They either study the influence of English on the linguistic landscape of a non-English speaking country or the presence of minority languages representing real language communities in the public space. The second category is further subdivided by Gorter et al. into the study of ‘autochthonous (or “traditional”) and migrant (“new”) minority languages’ (p. 5-6). The former subcategory seems to provide the bulk of the research to date. However, relatively little work has been done on the presence of migrant languages in traditionally monolingual contexts and to the best of my knowledge no studies have yet been published on the linguistic landscape of England.

So, the question at the heart of this paper is whether the study of the linguistic landscape within a region traditionally considered monolingual English-speaking, such as England, can still yield meaningful results concerning the true make-up of the language community of a given area and the social standing of certain micro-communities within a larger context.

Methodology

In order to decide whether the study of linguistic landscapes in England (or indeed similarly monolingual English settings) would yield meaningful results, it seemed necessary to conduct one and see what it would yield. Following on from Landry and Bourhis’ definition, the first decision any researcher makes has to be the territory to be covered (Backhaus 2007, p. 61). This is absolutely vital, as the detailed results of any study will only ever be valid for that given territory.

I chose London Road, Sheffield, as the area I surveyed for a variety of reasons. Sheffield is a city that prides itself on being multilingual. Languages Sheffield, ‘a voluntary organisation supporting the learning and maintenance of all languages’ (Languages Sheffield, HoLA), claims that ‘currently more than 120 languages are spoken in Sheffield’ (Languages Sheffield, Sheffield’s Languages). London Road is a busy commercial road just outside the area mainly considered the city centre (Sheffield Tourist Information Centre) and included on the city council’s ‘Map of

Thriving Districts and Local Centres' (Sheffield City Council), so it is reasonable to expect that there would be a substantial number of signs. Backhaus suggests the 'use of roads or railway lines as orientation markers' for the area (Backhaus, p. 61), so I chose a stretch of London Road between Clarke Square/Alderson Road at the northern end and Sharrow Lane/Woodhead Road at the southern end. This choice also touches on the question of 'how representative of a city as a whole the survey area should be' (Backhaus 2007, p. 61). London Road certainly is not representative of the whole of Sheffield, as it has a reputation for being multicultural, but I felt that for a first survey it would be sensible to choose an area which would in all likelihood yield diverse linguistic objects in many languages. If these could be analysed in a meaningful way, further studies might be conducted into other areas of Sheffield to put the first study into perspective.

The next decision to be made is what exactly to include as a 'unit of analysis' (Backhaus 2007, p. 61 and Gorter, p. 3). While the definition of Landry and Bourhis seems to be fairly clear and comprehensive, it does leave room for interpretation. Gorter et al. provide a very useful discussion of the different possibilities (p. 4). I decided to follow Backhaus' definition of 'any piece of text within a spatially definable frame' as quoted in Gorter (p. 3). Some researchers like Backhaus and Ben-Rafael et al. pre-selected the signs they would sample according to specific criteria (Gorter, p. 3), but Sheffield's diverse linguistic reputation forbade any such limitations. This study records all the written discourse in the form of stationary signs found on the stretch of London Road I selected. It presents a snapshot view of the linguistic landscape of London Road on 9 December 2013, around lunchtime.

I counted everything as language that could be identified by me as writing in any script, but ignored personal names. I categorised languages even if comprising only a single word, because the use of a foreign place name could fulfil the informational function of the linguistic landscape as identified by Landry and Bourhis (p. 25f) just as well as longer texts would.

To put the data yielded by the linguistic landscape into perspective against the linguistic make-up of the whole of Sheffield, I used data from the question concerning the main language spoken by the respondent from the 2011 Census (Census 2011).

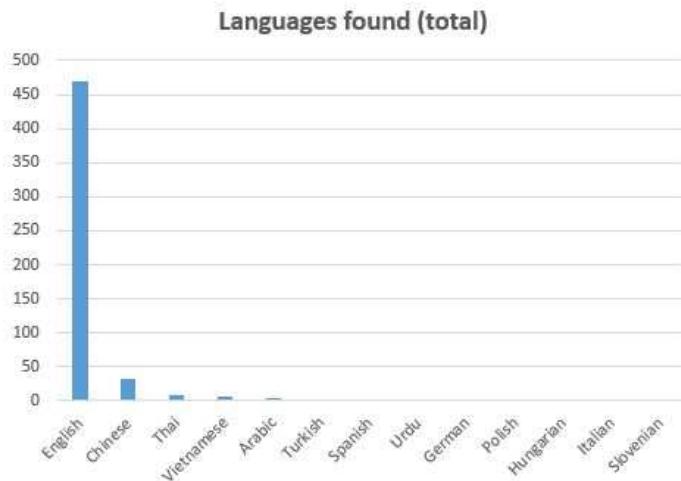
In the interpretation of the data found in the course of my micro-study I decided in my analysis to follow in the footsteps of Ben-Rafael et al. and examine my data against three basic attitudes toward social reality and social action. As presented by Ben-Rafael et al. (pp. 9f), Bourdieu sees social reality as power relations between the actors in a given setting. Thus the linguistic landscape should tell us something about the power and status of different language groups.

Goffman on the other hand sees the presentation of self as the decisive force behind social action. This means that the linguistic landscape should tell us which language communities wish, and are confident enough, to present themselves in the public space.

Finally, Boudon sees the actors as making rational choices about language use according to ‘good reasons’ that are determined by their interests and goals, such as making money.

Data analysis

A simple count of the 481 signs I collected revealed the presence of 13 different languages (Figure 1). By far the most common language was English.

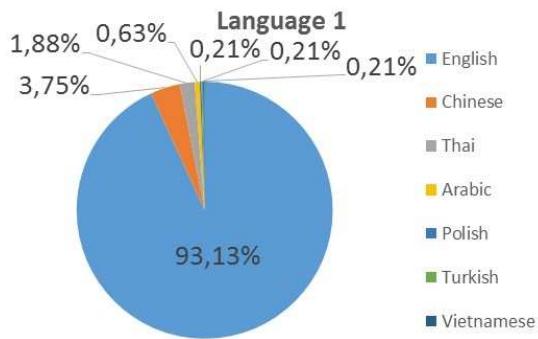
FIGURE 1

The predominance of English (89%) corresponds very closely to that of the census data (92%), whereas Chinese seems over-represented in the survey with 6% compared to 1% in the census. The same is true for Thai (1.7% in the survey compared to 0% in the census). The presence of Arabic again corresponds in both data sets (1% in the survey and 0.9% in the census). Urdu on the other hand is underrepresented (0.2% in the survey compared to 0.8% in the census).

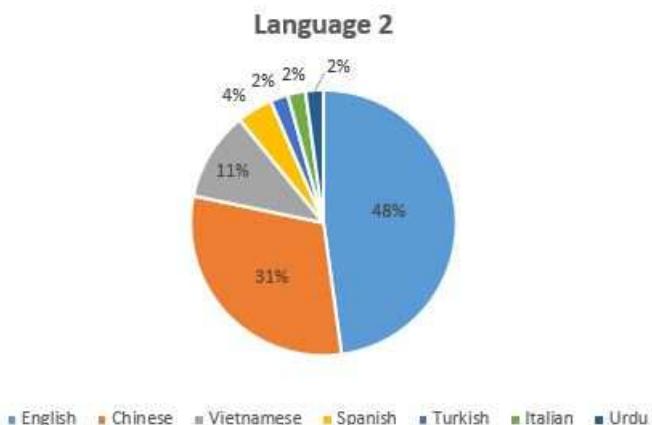
I classed the languages according to their status on each sign (for an example see Figure 2). The status was determined by position relative to reading direction and prominence (size of writing).

FIGURE 2

As Figure 3 illustrates, English constituted language 1 on 93% of signs with again Chinese coming second with 3.75%.

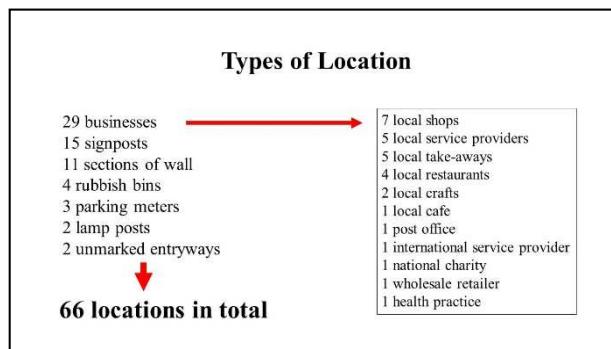
FIGURE 3

The picture is much more diverse for the language 2 (Figure 4), here English is still the most frequently used language, but the difference is not so marked.

FIGURE 4

Of the signs found, 90% were monolingual, 9% bilingual and just under half a per cent multilingual.

Figure 5 shows the types of locations found with a further breakdown of the businesses.

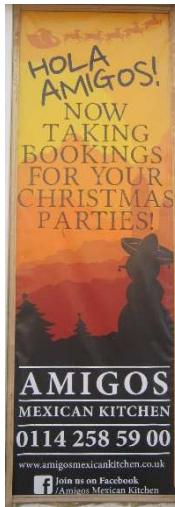
FIGURE 5

The vast majority of the signs were put up at the initiative of private persons and corporations. Most of the signs were printed professionally in the form of shop signs and posters, but self-printed computer print-outs were also found in significant numbers. Though all sizes were represented, the majority of posters were in the popular A4 to A2 format. That the government signs were all in monolingual English was not surprising as there is no policy of bilingualism or multilingualism in England. However, the large number of monolingual signs in English on the part of private actors is significant when considering the overall reputation of the area. In such a multicultural area more monolingual signs in languages other than English might have been expected.

Concentrating on the monolingual foreign, bilingual and multilingual signs as signs that are somehow connected to non-English speaking communities, interesting correlations appear. For instance, the share of PC print-outs increases significantly, while the percentage of hand written signs stays the same. This might suggest that while resorting to home printing because of the added control over scripts and language and due to the unavailability of 'official' posters in these, still great care is taken to convey a sense of professionalism that might seem lacking from hand written signs. Similarly, the percentage of medium-sized signs also increases, while the share of large signs stays the same. This indicates that the speakers of foreign languages do not need to hide evidence of their language, but use large signs in much the same proportion as the overall community.

The survey also yielded some individual examples that in my opinion form special cases defying quantitative analysis.

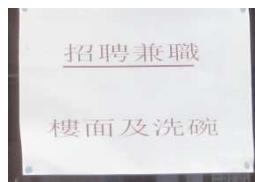
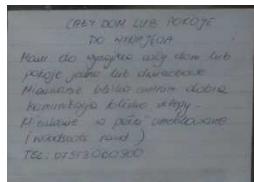
In this first one (Figure 6), the Spanish used is chosen so that most English speakers will understand and even if not, there is no information lost. Indeed, the fact that the vast majority of the signs containing languages other than English on London Road occurs in connection with restaurants, cafés and takeaways is too pronounced to be a coincidence. Clearly, these signs are not primarily aimed at the in-group, but rather at the out-group, suggesting authenticity of cuisine and promising exotic food, making them akin to the Irish language use that Moriarty quotes as having been found by Kallen in Irish tourist hot-spots (Moriarty, p.78), the purpose of which is to provide quaint local colour and a backdrop to your meal out or holiday trip.

FIGURE 6**FIGURE 7**

The second sign is that of an Anglo-Irish chain of betting shops (Figure 7). I think it is significant that an international chain not based in China chooses to use a script entirely unrelated to their business' origin to advertise. The characters are a phonetic representation of the chain's name in Chinese. As international chains usually operate under intense regulations concerning signage and corporate identity, the fact that such a chain decided to break their usual look to accommodate a minority language script strongly suggests that they expect the presence of potential readers of this script.

The two signs mentioned above draw attention to the expected readership of the signs found on London Road. One of the reasons for the dominance of English (cf. Figure 1) even in shops and restaurants with minority language names may be the fact that no language community around London Road is so predominant that its financial power would suffice to wholly support a business and that the use of English not only facilitates the custom of native speakers of English, but also of members of other minority language communities with English effectively functioning as a *lingua franca*.

This hypothesis is supported by the presence of a few specialised non-English monolingual signs, consisting mainly of job adverts for jobs in food-related businesses (e.g. Figure 8) and handwritten notices in the window of the Post Office (e.g. Figure 9).

FIGURE 8**FIGURE 9**

Their presence in the linguistic landscape of London Road only makes sense in the context of a possible community of readers, while at the same time consciously limiting the replies to these adverts to members of that community.

Discussion

When applying the three fundamentally different perspectives of Bourdieu, Goffman and Boudon as outlined in the methodology to this study, different aspects of the data are highlighted. Seen from Bourdieu's perspective, the power and highest status clearly reside with the English language. However, I think that this perspective may not present the whole picture. While being a useful angle in bilingual communities with conflicting native language communities, I feel it is too broad a brush for the

multilingual community of London Road, especially when also considering English as a possible *lingua franca* between different local language communities.

From Goffman's perspective, there were no clear delineations of language use according to side of the street or end of the stretch I surveyed. This might change if the area was extended as I saw 3 Polish shops in the distance but there was none on the stretch I studied. However, it may also be the case that no such clear areas exist in Sheffield.

The perspective of Boudon seems to me the most promising in this case. Clearly, a foreign language audience is expected in this area that would make it worth the trouble of putting up bilingual and monolingual non-English sings. The information goes beyond mere translation. Along with the example of Paddy Power and the Chinese job advert this is clear indication of actors in the linguistic landscape making conscious choices concerning the use of language in the public space according to who they wish to address and attract as customers or employees. This may also mean a decision to put up a sign in English if no one specific language community is being targeted.

Generally, the fact that Chinese was the only language other than English to be found on all three types of signs, professional as well as PC print-outs and hand-written personal messages intended for an audience that can actually read them, gives an indication that here is a language community behind the visible language. And the fact that it was also the only language found on all sizes of signs shows that this community is not afraid to visibly display its discourse, and has been established for long enough for permanent signage to have emerged.

Conclusions

The study of the linguistic landscape of London Road yielded meaningful results. It showed that linguistic objects are being used by the community occupying London Road and the adjacent residential areas for two main purposes. One purpose is that of

in-group information. Here signs are used in languages other than English to target specific audiences and convey certain information exclusively to the in-group. The other one is the conveyance of a sense of exoticism and authenticity of a product or service. In this case the language use is directed at the out-group. A third possible purpose, this time concerning the use of English, is that of a *lingua franca* to provide communication directed at the community as a whole, therefore strongly suggesting the diverse ethnic make-up of the area. The fact that no obliteration or defacing of any linguistic objects were found can also be seen as significant, as it suggests a basically peaceful co-existence of minority language groups in and around London Road.

The results of this micro study of an English linguistic landscape show that the method has its value in a majority English-speaking context as well as in those previously studied by others. However, to fully understand what the linguistic landscape can tell us about their respective communities, more research should be undertaken. Other researchers have expanded their understanding of the linguistic landscape to include buildings as well as more dynamic signage such as digital screens, mobile signs, the writing on clothing of and items carried by passers-by, and even spoken discourse (Gorter et al., pp. 3-4), which might provide angles for further research. Sayer expanded his concept of the linguistic landscape into the fourth dimension by adding to his list of signs over several months (Sayer, p. 145), another approach that may be of interest in the future. It might also prove fruitful to expand the research area to include all of London Road or to conduct similar studies into other areas of Sheffield. It would for instance be interesting to see if sign sizes or types vary according to different areas. More research into the communities itself using questionnaires to assess factors like second language use and recent migration patterns might also provide more useful data for comparison than the rather broad brush of the main language census question and explain for instance, why languages were found in the linguistic landscape that do not form part of the linguistic make-up of the city as recorded by the census.

Bibliography

- 2011 Census. QS204EW Main language (detailed), local authorities in England and Wales. In *Office for National Statistics*, <<http://www.ons.gov.uk/ons/index.html>> [released 30 January 2013].
- Backhaus, Peter (2006). Multilingualism in Tokyo: A look into the linguistic landscape. In Durk Gorter (ed.), *Linguistic landscape: A new approach to multilingualism*, (pp. 52-64). Clevedon: Multilingual Matters.
- Backhaus, Peter (2007). *Linguistic landscapes: A comparative study of urban multilingualism in Tokyo*. Clevedon: Multilingual Matters.
- Bourhis, Richard Y. and Rodrigue Landry (1997). Linguistic landscape and ethnolinguistic variety: An empirical study. *Journal of Language and Social Psychology*, 16, 23-49.
- Gorter, Durk, Marten, Heiko H. & Mensel, Luk Van. (2006). Introduction: The study of the linguistic landscape as a new approach to multilingualism. In Durk Gorter, Heiko H. Marten & Luk Van Mensel (ed.), *Linguistic landscape: A new approach to multilingualism* (pp. 1-6). Clevedon: Multilingual Matters.
- Gorter, Durk et al. (2012). Studying minority languages in the linguistic landscape. In Durk Gorter et al (eds.), *Minority languages in the linguistic landscape*, (pp. 1-15). Basingstoke: Palgrave Macmillan.
- Languages Sheffield. About us. In *Languages Sheffield*, <<http://www.languages-sheffield.org.uk/pages/about-us-0>> [accessed 22 January 2014].
- Languages Sheffield. HoLA Project. In *Languages Sheffield* <<http://www.languages-sheffield.org.uk/pages/hola-project-0>> [accessed 22 January 2014].
- Languages Sheffield. Sheffield's languages. In *Languages Sheffield* <<http://www.languages-sheffield.org.uk/pages/sheffields-languages-0>> [accessed 22 January 2014].
- Moriarty, Máiréad (2012). Language ideological debates in the linguistic landscape of an Irish tourist town. In Durk Gorter et al. (eds.), *Minority languages in the linguistic landscape* (pp. 74-88). Basingstoke: Palgrave Macmillan.
- Marten, Heiko F. (2012). “Latgalian is not a language”: Linguistic landscapes in eastern Latvia and how they reflect centralist Attitudes. In Durk Gorter et al.

- (eds.) *Minority languages in the linguistic landscape* (pp. 19-35). Basingstoke: Palgrave Macmillan.
- Sayer, Peter (2010). Using the linguistic landscape as a pedagogical resource. *ELT Journal*, 64, 143-154.
- Sheffield City Council. *Map of thriving districts and local centres*. [n.p., n.pub.]
<<https://www.sheffield.gov.uk/dms/scc/management/corporate-communications/documents/community-life/community-development/Map-of-Thriving-Districts-and-Local-Centres--PDF-3-07MB-.pdf>> [accessed 22 January 2014].
- Sheffield Tourist Information Centre, *Welcome to Sheffield: Map and mini guide*. [n.p., n.pub.]
<<http://www.welcomesheffield.co.uk/content/files/437>Welcome-to-Sheffield-map-and-miniguide.pdf>> [accessed 22 January 2014].

All pictures were taken by myself. I would like to thank Hailey Lam and Ting Feng for their help with Chinese signs.

Tracing the Emergence of Sociolinguistic Variation in L2 Speech: Polish Migrants in Edinburgh

Daniel Lawrence
University of Edinburgh
dlawrenc@staffmail.ed.ac.uk

1. Introduction

This paper explores the process by which migrant L2 speakers ‘pick up a local accent’ – how they adopt the patterns of sociolinguistic variation present in their host speech community. Previous work has shown that adult migrants may adopt patterns of sociolinguistic variation present in their new environment; however, these patterns are rarely fully acquired. In some cases, the social and linguistic constraints are replicated; in others, they are rejected; alternatively, they may be re-allocated, with the ordering of factors re-organised.

Acquiring sociolinguistic variation is a challenging task – it is not simply a case of choosing a socially-meaningful variant and adopting it categorically. Learners need to notice and replicate relative frequencies of variants, which may be sensitive to constraints at all levels of the grammar as well as social factors. Meyerhoff & Schleef (2012; 2013) highlight the *complexity* of acquiring variation by considering the steps involved. In order to fully acquire sociolinguistic variation, speakers need to learn the following (Meyerhoff & Schleef, 2012, p.409):

- a) The relative frequencies of variants.
- b) The linguistic and social constraints.
- c) Constraint orderings within factor groups.
- d) Stances, activities and styles indexing social categories such as gender.

Meyerhoff & Schleef (2012) suggest that the variable outcomes of acquisition may be explained by the cognitive demand of co-ordinating these steps. Building on this idea, it might be possible to consider variable constraints in terms of the complexity of the generalisation they demand. Phonological constraints demand only one level of generalisation – learners must notice which segments a particular variant is likely to co-occur with and replicate the pattern. However, grammatical constraints require learners to generalise over sets of lexical items based on morphosyntactic criteria. These constraints might be considered more complex, and thus more likely to undergo ‘transformation under transfer’ (Meyerhoff, 2009, p.313) or be rejected completely. Meyerhoff & Schleef (2012) suggest that social constraints may be particularly difficult to acquire, hypothesising that social meanings indexing group membership or social categories such as gender may be more accessible to learners than those indexing individual styles, stances or activities.

This paper attempts to explore the above proposals by comparing groups of learners at different stages of acquisition. The approach follows that of Meyerhoff & Schleef (2013) and Schleef (Forthcoming). Data are presented from a corpus of sociolinguistic interviews carried out with Polish migrants living in Edinburgh. Multivariate analyses are presented from a lower Length of Residence (LoR) group (14-60 months) and a higher LoR group (65-110 months), focusing on the learners’ production of word-final (t). The variable (t) is defined as any instance of the segment /t/ directly following a sonorant in coda or non-foot-initial onset position. The present work follows Schleef (2013) in focussing on glottal replacement of word-final (t), which is defined as word-final realisations of (t) which have a ‘robust glottal quality’, including both glottal stops and periods of creaky voice (p.201). Glottal replacement of (t) is a well-researched feature of British varieties of English (see e.g. Straw & Patrick (2006); Schleef (2013)), making it an excellent candidate for exploring the acquisition of sociolinguistic variation.

2. Data Collection

The sample for the present study consists of speakers who migrated after Poland’s 2004 accession to the EU. Their previous exposure to English is relatively

homogeneous – they all received instruction in Standard British English as part of their compulsory education, and would likely have had considerable communicative ability prior to their arrival in the UK. Crucially, it is unlikely that these speakers had regular pre-arrival exposure to native-like English in a range of contexts, which can be assumed to be critical for the acquisition of native-like sociolinguistic competence (Drummond, 2010).

The recruitment criteria were as follows:

- Speakers were aged between 18 and 40.
- Speakers were born and raised in Poland.
- Speakers had a minimum LoR of 1 year.
- Speakers had spent no significant time in other English-speaking countries, or other parts of the UK.

Interviews were participant-led, although a range of questions based on Hoffman & Walker's (2010) Ethnic Orientation Questionnaire were included, covering a range of factors related to speakers' language use, degree of involvement in local society and interest in local history and culture. Following Drummond (2010) and Hoffman & Walker (2010), it was thought that these factors might influence the extent to which migrants adopt glottal replacement. The interviews ended with a short reading task.

3. Analysis

Five word-final (t) variants were coded initially: released [t], glottalised [?], flap/tap [ɾ], voiced released [d] and elided [Ø]. Tokens which were ambiguous or inaudible due to speaker overlap were excluded from the analysis. This resulted in a total of 1256 tokens. The independent variables coded for (t) are summarised below, based on Schleef (2013):

Preceding context: vowel (*pot*), liquid (*built*), nasal (*pint*)

Following context: stop (*the night before*), fricative/affricate (*start the car*, nasal (*meet my parents*), liquid (*start running*), glide (*quit working*), vowel (a lot of), pause.

Grammatical category: proper noun (*Scott*), pronoun (*that*), noun (night), adjective (*cute*), adverb (*quite*), verb (*start*), conjunction (*but*), preposition (*at*).

Number of syllables in word: 1,2,3,4+

Word frequency: LOG^{10} transformation of BNC word count +1

Realisation of previous (t): [ʔ], [t], [ɾ],[d], [Ø], none

Drawing on previous work on (ing) and (t), the following social constraints were coded for both variables:

Style: Reading, Casual

Gender: Male, Female

Speaker age: Continuous measure in whole years

Length of Residence: Continuous measure in whole years

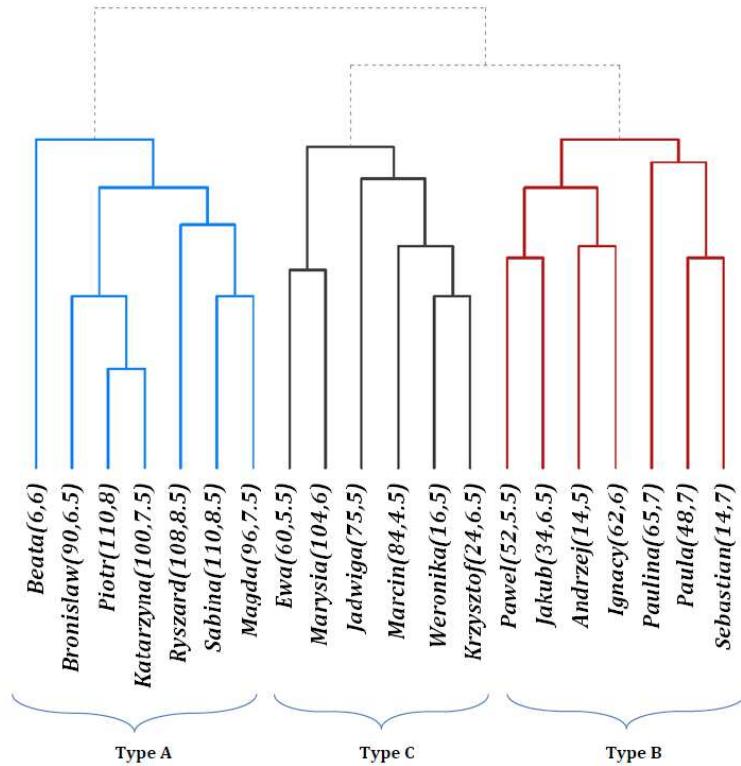
Proficiency: Ordinal scale of 1-9, judged by two trained language assessors

Integration Category: A, B, C

The *Integration Category* was derived from a statistical analysis of speakers' responses to the Ethnic Orientation questions. This questionnaire yielded a set of responses to 30 questions from each speaker. Following Hoffman & Walker (2010), each response was assigned a value between 1 and 3, where '3' represents the response that would reflect the most engagement with Scottish society, and '1' the least engagement. Indeterminate or neutral answers were given the value '2'.

These data were analysed using *complete-linkage hierarchical cluster analysis*. This algorithm operates iteratively, identifying the two speakers or groups of speakers which have the most similar responses at every stage, until all entries have been categorised. The resulting *dendrogram* expresses clusters of speakers who gave similar questionnaire responses, as shown in Figure 1:

FIGURE 1: COMPLETE-LINKAGE HIERARCHICAL CLUSTER ANALYSIS OF QUESTIONNAIRE RESPONSES. LoR (MONTHS) AND PROFICIENCY (1-9) PROVIDED IN BRACKETS.



Broadly, type ‘A’ migrants show the most integration with the local community, type ‘C’ migrants the least, with the ‘B’ group representing an intermediate category. Qualitative analysis of the interviews supports the assertion that these categories are socially relevant to the migrants; see Lawrence (2013) for a full description. The Integration Category variable thus represents a combined measure of social network characteristics and informants’ integration into the host community.

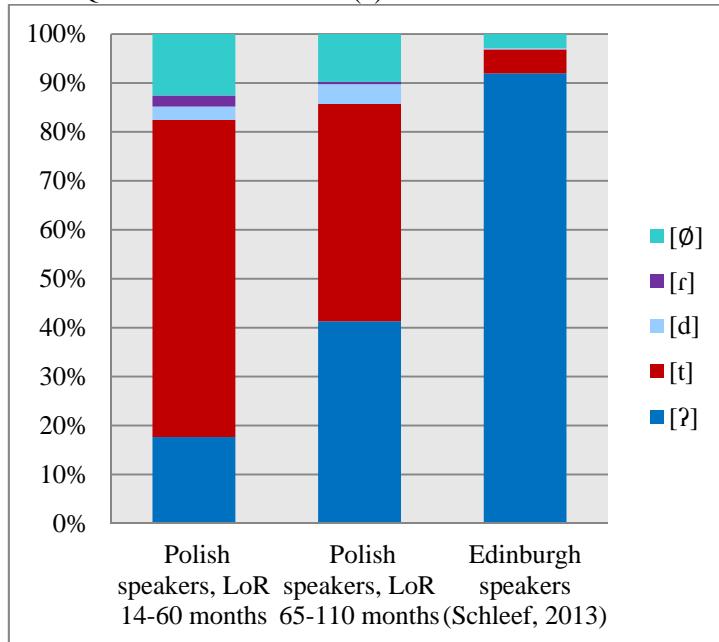
Following Schleef & Meyerhoff (2012) and Schleef (2013), the corpus was split into two groups: speakers resident for between 14 and 60 months and speakers resident for between 65 and 110 months. Splitting the data in this way was an attempt to represent speakers at different stages of the process of acquiring sociolinguistic variation. See Lawrence (2013) for a detailed explanation of the choice of LoR as a proxy for stage of acquisition.

4. Results

a) Frequency of variants

Figure 2 shows the proportion of instances of (t) realised as each variant for both groups of Polish speakers, alongside rates reported for Edinburgh natives.

FIGURE 2: OVERALL FREQUENCY OF WORD-FINAL (T) VARIANTS.



These data provide evidence that Polish migrants in Edinburgh are adopting non-standard features of the local variety, albeit to a limited extent. Despite whatever developmental limitations they may have (c.f. work on the *Critical Period Hypothesis*), these speakers seem to have identified variable features of the ambient language and adopted them in their own speech.

b) Regression analyses

Although the frequency data demonstrate the adoption of non-standard variants, they do not provide evidence of acquisition *per se*. The claim that learners have acquired a sociolinguistic variable requires evidence of native-like constraints in their production. To this end, mixed-effects multiple regression models were fitted in Rbrul (Johnson, 2009). The models were constructed exactly as in Meyerhoff & Schleef (2013). Glottal replacement [?] was contrasted with all other variants. Where independent

variables showed categorical application or non-application, they were excluded. Any exclusions or token collapses were held constant across the two groups, to allow for accurate comparison.

Based on the reasoning outlined in the introduction, the following predictions were made regarding differences in constraint adoption between the two LoR groups:

1. Phonological constraints, such as preceding and following environment, will appear early in the acquisition process, and are more likely to be adopted accurately.
2. Constraints which involve other levels of the grammar, such as grammatical category effects, will be acquired later, and more likely to undergo re-allocation or be rejected.
3. Social constraints will appear only late in the acquisition process, with macro-level constraints such as age or gender appearing before micro-level constraints such as style.

Table 1 summarises the constraints selected as significant across the LoR group.

Polish speakers, LoR 14-60 months					Polish speakers, LoR 65-110 months									
	Log odds	N	% [?]	Factor weight		Log odds	N	% [?]	Factor weight					
Following segment (p=1.05e-07)														
plosive	1.078	124	33.9	0.746	Preceding segment (p=3.39e-08)									
nasal	0.589	45	33.3	0.643	vowel	0.583	875	52.1	0.642					
fricative/affricate	0.012	168	19.0	0.503	liquid/nasal	-0.583	275	24.0	0.358					
glide	-0.002	99	20.2	0.499	Grammatical category (p=7.16e-08)									
liquid	-0.334	35	14.3	0.417	conjunction	1.095	159	67.9	0.749					
pause	-0.424	214	16.8	0.396	pronoun	0.626	155	64.5	0.652					
vowel	-0.919	201	13.4	0.285	adverb	-0.126	203	52.0	0.468					
Preceding segment (p=0.00129)														
vowel	0.428	668	22.6	0.605	preposition	-0.132	104	40.4	0.467					
liquid/ nasal	-0.428	218	11.9	0.395	verb	-0.214	314	33.3	0.447					
Style (p=0.0157)														
casual	0.274	512	22.2	0.568	adjective	-0.611	100	32.0	0.352					
reading	-0.251	374	16.2	0.432	noun	-0.638	115	26.1	0.364					
<i>Not significant: age, gender, integration category, proficiency, no. syllables in word, lexical frequency, previous realisation of (t)</i>														
Deviance: 683.796					Style (p=0.00196)									
Df: 10					casual	0.485	753	53.5	61.9					
Speaker ID random SD: 1.253					reading	-0.485	397	30.0	0.381					
N= 886					Integration category (p=0.00325)									
<i>A</i>														
1.281														
<i>B</i>														
0.394														
<i>C</i>														
-1.675														
Previous realisation of (t) (p=0.0174)														
none	0.403	21	52.4	0.6	Following segment (p=0.0242)									
glottal replacement	0.024	421	67.2	0.506	nasal	0.424	53	52.8	0.604					
other stop	0.428	708	32.2	0.395	liquid	0.385	43	55.8	0.595					
Following segment (p=0.0242)														
nasal	0.277	143	43.4	0.569	plosive	0.000	118	47.5	0.5					
liquid	-0.189	217	45.6	0.453	glide	-0.339	293	47.4	0.416					
plosive	-0.339	293	47.4	0.416	fricative/affricate	-0.558	283	40.3	0.364					
glide	-0.558	283	40.3	0.364	pause									
fricative/affricate					vowel									
pause					<i>Not significant: age, gender, proficiency, no. syllables in word, lexical frequency</i>									
vowel					<hr/>									
Deviance: 1055.716										<hr/>				
DF: 20										<hr/>				
Speaker ID random SD: 0.917										<hr/>				
N= 1,150										<hr/>				

TABLE 1: SIGNIFICANT CONSTRAINTS ON WORD-FINAL (T) FOR THE TWO LOR GROUPS. APPLICATION VALUE = [?]

Table 1 indicates that both groups have adopted native-like phonological constraints on their production of word-final (t). The general pattern here appears to be PreC>PreP>PreV, which is widely reported in accounts of native speech (Schleef, 2013; Straw & Patrick, 2006). A grammatical category constraint has been adopted by the longer-LoR group, whereby conjunctions and pronouns strongly favour glottal replacement. This constraint is not significant for the shorter-LoR group. A similar pattern is reported for Edinburgh speakers in Meyerhoff & Schleef (2013). However, the Edinburgh pattern broadly follows a nominal-verbal continuum (with verbs favouring [?]), while the effect found in the present study is carried by conjunctions and pronouns.

Turning to the social constraints, native-like stylistic constraints can be noted; speakers seem to prefer glottal replacement in casual conversation and avoid it in monitored speech. This effect is present in both groups. Integration Category emerges as a significant constraint for the longer-LoR speakers: the ‘Type A’ speakers, or most integrated, exhibit the highest rates of glottal replacement, followed by ‘Type B’ speakers; ‘Type C’ speakers seem to avoid this variant. This suggests that migrants are adopting glottal replacement as an index of their involvement in the local community. Crucially, this constraint is absent for the shorter-LoR group, suggesting that the migrants’ integration into the host community is reflected in their use of local non-standard variants only at later stages of acquisition.

Table 2 summarises the differing outcomes of acquisition across the two LoR groups.

TABLE 2: SIGNIFICANT CONSTRAINTS FOR WORD-FINAL (T) ACROSS BOTH GROUPS
*(in order of effect size; + attested in British English; * factors re-ordered; arrows show shared constraints; bold text shows constraints added in the Longer-LoR model)*

Shorter-LoR Migrants	Longer-LoR Migrants
Following segment +	Preceding segment +
Preceding segment +	Grammatical category +*
Style +	Style +
	Integration category
	Previous realisation of
	variable
	Following segment +

```

graph LR
    A[Following segment +] --> C[Grammatical category +*]
    B[Preceding segment +] --> C
    D[Style +] --> E[Integration category]
    D --> F[Previous realisation of variable]
    F --> G[Following segment +]
    
```

Overall, learners at a later stage of acquisition seem to have adopted a more complex set of constraints. Not only do they employ the non-standard variant more frequently as they live longer in the host community, they also become more native-like in terms of the linguistic and social contexts in which the different variants are used. This acquisition of variable patterns appears to happen implicitly – while learners might notice different features and try to imitate them, it seems unlikely that they are explicitly aware, for example, that a word-final glottal stop is more likely before a consonant than a vowel. Nevertheless, these patterns emerge in learners’ production, broadly in order of relative complexity. Phonological constraints are adopted earlier and more accurately than grammatical category or word-length effects. In terms of linguistic constraints, the data appear to support the predictions outlined in the introduction; it seems that phonological constraints are relatively straightforward for learners to acquire, while grammatical category and syllable effects are more challenging.

Predictions (1) and (2) (p.48), which concern the acquisition of linguistic constraints, seem to be borne out by these data. Prediction (3) suggested that stylistic constraints would emerge a) only late in the acquisition process, and b) after more ‘concrete’ social indices had been acquired. However, the data do not support this prediction. Speakers exhibited stylistic variation in (t) even at early stages of acquisition, and the group who had been resident for a longer period of time appeared to produce different rates of glottal replacement depending on their integration into the host community. The fact that style emerges before any social category indices

can only suggest that the way this variable is being acquired is quite different from what Schleef & Meyerhoff (2012) suggest. Rather than associating variants first with transparent group identities and then abstracting to individual styles, stances or activities, it seems that the acquisition of variation in (t) is proceeding in the opposite direction; speakers seem to first associate the variants with different styles, stances and activities, then later connect them with their knowledge of broader social categories in the speech community.

5. Conclusion

Where most previous work has focused on the *outcome* of the transfer of variation under contact, the account of constraint complexity described in the present work allowed predictions to be made about the *process* of acquiring variation – it was hypothesised that phonological constraints might be acquired earlier and more accurately than grammatical category and word length effects. Predictions about the acquisition of socio-indexical meaning were drawn from Schleef & Meyerhoff (2012) – it was suggested that social category indices are more readily acquired than those related to stylistic variation. Although the predictions regarding linguistic constraints were borne out by the data, the pattern observed for social constraints was the opposite of what was expected. This suggests that in some cases, learners may exhibit stylistic variation in a given variable at an early stage, only later learning to co-ordinate it as an index of group identity.

This project adds to a growing body of work which shows that migrants' acquisition of sociolinguistic variation is a complex and interesting phenomenon. Adult L2 speakers are sensitive to probabilistic patterns of linguistic variation in the speech they encounter, and seem to be able to acquire these patterns implicitly. As well as learning the linguistic constraints on variation, speakers also acquire knowledge of how variation is distributed socially, and with it, knowledge of the social structure of the speech community.

References

- Drummond, R. J. (2010). *Sociolinguistic variation in a second language: The influence of local accent on the pronunciation of non-native English speakers living in Manchester* (Unpublished PhD Dissertation). University of Manchester.
- Hoffman, M. F., & Walker, J. A. (2010). Ethnolects and the city: Ethnic orientation and linguistic variation in Toronto English. *Language Variation and Change*, 22(01), 37-67.
- Johnson, D. E. (2009). Getting off the GoldVarb standard: Introducing Rbrul for mixed-effects variable rule analysis. *Language and Linguistics Compass*, 3(1), 359-383.
- Lawrence, D. (2013). *Fitting in: adult migrants' acquisition of sociolinguistic variation in Edinburgh English* (Unpublished MSc Thesis). University of Edinburgh.
- Meyerhoff, M. (2009). Replication, transfer and calquing: Using variation as a tool in the study of language contact. *Language Variation and Change*. 21(3), 1-21.
- Meyerhoff, M. & Schleef, E. (2012). Variation, contact and social indexicality in the acquisition of (ing) by teenage migrants. *Journal of Sociolinguistics*, 16(3), 398-416.
- Meyerhoff, M. & Schleef, E. (2013). Hitting an Edinburgh target: Immigrant adolescents' acquisition of variation in Edinburgh English. In R. Lawson (ed.), *Sociolinguistic perspectives on Scotland* (pp. 103-108). Basingstoke: Palgrave Macmillan.
- Schleef, E., Meyerhoff, M., & Clark, L. (2011). Teenagers' acquisition of variation: A comparison of locally-born and migrant teens' realisation of English (ing) in Edinburgh and London. *English World-Wide*, 32(2), 206-236.
- Schleef, E. (2013). Glottal replacement of /t/ in two British capitals: effects of word frequency and morphological compositionality. *Language Variation and Change*, 25(2), 201-223.
- Schleef, E. (Forthcoming). Developmental sociolinguistics and the acquisition of T-glottalling by immigrant teenagers in London. in G. De Vogelaer & M. Katerbow (eds.), *Variation in language acquisition*. Amsterdam: John Benjamins.

Straw, M & Patrick, P.L. (2007). Dialect acquisition of glottal variation in /t/:
Barbadians in Ipswich. *Language Sciences*, 29(2), 385-407.

Bringing theoretical L2 acquisition research findings to the language classroom: A materials development challenge

Elaine Lopez
University of York
esl505@york.ac.uk

Since the majority of second language (L2) acquirers receive instruction at some point, it may seem obvious that second language acquisition (SLA) research would be relevant to language teachers and learners. However, theoretical SLA research conducted within the framework of generative linguistics (henceforth GenSLA) does not commonly influence classroom practice. Increasingly, there has been a disconnect between language teaching and acquisition research, as demonstrated by this statement from Lightbown (2000 p.437): “SLA researchers whose work is focused on solving theoretical puzzles [are] increasingly separating their research activities from those of researchers whose questions [are] more pedagogical in nature.” Possible reasons for this perceived separation could be an overuse of subject specific terminology in research publications, or disagreement about the theoretical underpinnings of SLA. GenSLA research may be perceived by teachers as inaccessible and irrelevant. However, as Bruhn de Garavito (2013, p.32) states, “Research carried out within the generative tradition has something to say regarding language teaching. It is a valuable resource that should not be squandered.” This paper will discuss applying the results of GenSLA research to language teaching using the example of specificity in the English article system.

There are four parts to the paper. Section 1 gives an overview of GenSLA research into the acquisition of English articles which has highlighted the importance of specificity for speakers whose first language (L1) does not have articles. Section 2 shows how articles are currently taught, providing evidence of standard teaching materials focusing only on definiteness and excluding specificity; and also of inaccurate use of the term 'specific' in some textbooks. Section 3 discusses the process of developing the results of GenSLA research into a teachable concept, and finally, an example of linguistically-informed teaching materials will be given.

1. Specificity in article acquisition research

Article errors take two forms, misuse or omission, and are widely documented in L2 English. Example (1) shows an article misuse error, taken from the written work of an L1 Chinese student at Sheffield Hallam University. Here, the definite article (underlined) has been used in a context that is obligatorily indefinite.

- 1) Tomorrow I going shopping because I need to buy the new suit.

In last decade, a growing body of GenSLA work has investigated how articles are acquired by L2 learners of English. The possibility that two-article languages, such as English, select articles on the basis of either definiteness or specificity was suggested by Ionin (2003), who proposed a parametric variation between languages which set articles on the basis of definiteness (like English) or specificity. The definition of definiteness and specificity given by Ionin, Ko and Wexler (2004 p.5; henceforth IKW) is shown in (2).

- 2) Definiteness and Specificity: Informal definitions

If a Determiner Phrase (DP) of the form [D NP] is...

1. [+definite], then the speaker and hearer presuppose the existence of a unique individual in the set denoted by the NP.
2. [+specific], then the speaker intends to refer to a unique individual in the set denoted by the NP and considers this individual to possess some noteworthy property.

Ionin hypothesised that more L2 errors would be made in contexts which were either definite and non-specific or indefinite and specific. Examples (3) and (4) show the potentially difficult contexts.

- 3) definite/non-specific

I read a very good book recently. I don't remember the name of the author.

- 4) indefinite/specific

Two ladies are sitting in a restaurant. They are waiting for a friend but she is late.

Results of a forced-choice elicitation task conducted on groups of L1-Russian and L1-Korean learners of English by IKW (2004) supported the proposal that learners struggle to produce articles correctly in these two contexts. The error-rates for definite non-specific dialogues were 33% for the L1-Russian learners of English, and 14% for the L1-Korean learners of English. The indefinite specific dialogues had slightly higher error rates of 36% and 22% respectively (IKW, 2004). In contrast, the error rates for definite specific and indefinite non-specific contexts were between 4 and 8% for each group of learners. IKW (2004) concluded that both specificity and definiteness have a significant effect on article choice for both groups of learners.

Other research, whilst not always agreeing with Ionin's explanation of a parametric variation, nonetheless supports her finding that learners' with an article-less L1 will fluctuate between correct and incorrect use of English articles on the basis of specificity. Ionin, Zubizarreta and Bautista Maldonado (2008) found an effect of specificity on Russian learners, but not Spanish learners. The presence of article errors in the L1-Spanish group appeared to be caused by L1 transfer since Spanish, like English, selects articles on the basis of definiteness. Likewise, Hawkins et al. (2006) found an effect of specificity on Japanese learners, but not Greek learners. Finally, Tryzna (2009) tested L1-Polish and L1-Mandarin Chinese adult learners of English and found that indefinite specific contexts had a higher misuse of the definite article amongst learners from both language groups. This context also showed the highest error rates in the IKW (2004) study.

Whilst generative researchers focus on explanations for these errors, for example parameter resetting or a feature re-assembly account, the important point for teaching is knowing that these problem areas exist. If learners can be made aware of these problematic contexts in the L2 classroom, it may improve article accuracy. Therefore, the next section will consider how articles are currently taught to L2 learners of English.

2. Standard teaching materials

A review of four series of general English coursebooks has uncovered some trends in the way that articles are currently taught to L2 learners of English. These titles were

chosen because they are some of the most widely-used general English coursebooks in the UK. The books are New English File (Oxenden, Latham-Koenig and Seligson, 2004; 2005; 2008; 2009), Cutting Edge (Cunningham and Moor, 2005; 2007), Language Leader (Cotton, Falvey and Kent, 2008; Lebeau and Rees, 2008), and New Headway (Soars and Soars, 2003; 2005; 2006). Across these courses, articles were generally introduced at an elementary level, although with a simplification of the rules. Most then continue to teach articles at every level of proficiency. In theory at least, this approach should allow for the gradual acquisition of articles, with learners being presented with progressively more complex rules. The widely documented problems with article use amongst L2 learners, however, seem to suggest otherwise. All of the materials in the four series of books focused on definiteness; no published teaching materials could be found which provide instruction on the specific/non-specific contrast. This suggests a disconnect between what research has demonstrated about the acquisition of the English article system, and how it is taught to L2 learners of English.

In addition, the term ‘specific’ is used within several published textbooks as a synonym of ‘definite’. Not only is this linguistically inaccurate, but could create problems if it leads L2 learners to believe that all specific contexts are definite and all non-specific contexts are indefinite. Such a misconception could lead to the exact type of errors uncovered by the GenSLA research discussed in Section 1.

One such example of linguistically inaccurate use of the term ‘specific’ comes from a textbook for students of academic English (Bailey, 2006). Examples (5) and (6) both contain problematic uses of the term ‘specific’, which have been underlined. First is an explanation of the rules for article use, where ‘specific’ is used as a synonym of ‘definite’ when describing the examples.

5) Unless they are uncountable, all nouns need an article when used in the singular.

The article can be either *a/an* or *the*. Compare:

- a) Research is *an* important activity in universities.
- b) *The* research begun by Dr Mathews was continued by Professor Brankovic.

c) A survey was conducted among 200 patients in the clinic.

In (a) research, which is usually uncountable, is being used in a general sense.

In (b) a specific piece of research is identified.

In (c) the survey is not specified and is being mentioned for the first time.

Bailey (2006, p.130)

Secondly, in this exercise ‘specific’ is used as a synonym of ‘definite’ when use of the definite article is contrasted with generic uses of the zero article.

6) ***In the following sentences, decide if the words in italic are specific or not.***

Insert the if specific.

Example: inflation was the greatest problem

for Brazilian government.

Inflation was the greatest problem for *the* Brazilian government.

a) *engineering* was the main industry in the region.

b) *global warming* is partly caused by *fossil fuels*.

c) *Russian revolution* was partly a result of *First World War*.

[d) ... n)]

Bailey (2006, p.131)

Examining these teaching materials in the context of what GenSLA research has shown about the importance of specificity demonstrates one area where such research could be applied to teaching. Research findings strongly suggest specificity is important and yet learners are only taught about definiteness. The next section of this paper will go on to explain the present study’s application of IKW’s (2004) findings to the language classroom through the development of new, linguistically-informed teaching materials.

3. Process of materials development

In order to develop linguistically-informed grammar instruction materials based on the results of IKW’s study, a consultation with practising English teachers was carried

out. The teachers had no background in generative linguistics, and the consultation continued at all stages of materials development. Feedback was given by the teachers on the positive and negative points of the materials, as well as whether they understood the concepts being presented. There was an initial objection to instruction on specificity, as the teachers did not feel confident teaching a property of which they had no detailed knowledge, and which they considered too abstract to be teachable. Therefore, although the materials needed to be linguistically accurate, there was also a requirement to make them simple enough so that teachers could use them without a prior knowledge of specificity.

Once an initial version of the materials had been produced, extensive piloting was conducted with learners of different proficiency levels, all of whom were university students at Sheffield Hallam University. The use of a traditional presentation/practice lesson format was decided on so that input could be controlled when the materials were used by different teachers. Additionally, the presentation materials were put onto PowerPoint, again due to the need to control input across different groups. A decision was made to adapt current article instruction materials for the practice materials.

One challenge when developing the teaching materials was changing the linguistic definition of definiteness and specificity given in Section 1 into something that was comprehensible and teachable for non-linguists. After consultation with the teachers and piloting, the definitions were simplified for teaching use, as shown in (7).

7) Definiteness and Specificity: pedagogical definitions

If a noun phrase is...

1. [+definite], then *both* the speaker *and* the listener can identify the noun, and answer the question ‘Which one?’
2. [+specific] then the speaker is referring to one particular individual.⁹

⁹ This definition refers to the singular form of the noun since it is an adaptation of IKW’s (2004) informal definition which refers to a ‘unique individual’. Plural noun phrases, like singular nouns phrases, can be either specific or non-specific, and examples of plural nouns requiring the definite article were included in the teaching materials.

The simplification was particularly important when describing the concept of specificity. Therefore, ‘speaker intent to refer’ to an individual was maintained from the original definition, but the concept of that individual possessing some ‘noteworthy property’ was removed.

At the end of the consultation period and after piloting, the final version of these materials was used as part of a larger research project looking at the role of instruction on the L2 acquisition of the English article system. A pre-test and post-test was administered to the learners, and results will be contrasted with two other groups of learners, one of whom was taught using standard materials, and one group who received no instruction on articles. All of the learners who participated in this project are L1 speakers of Mandarin Chinese, a language that does not have articles.¹⁰

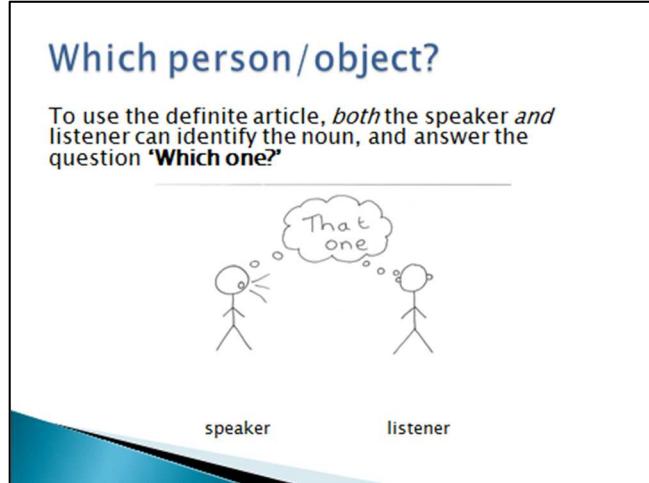
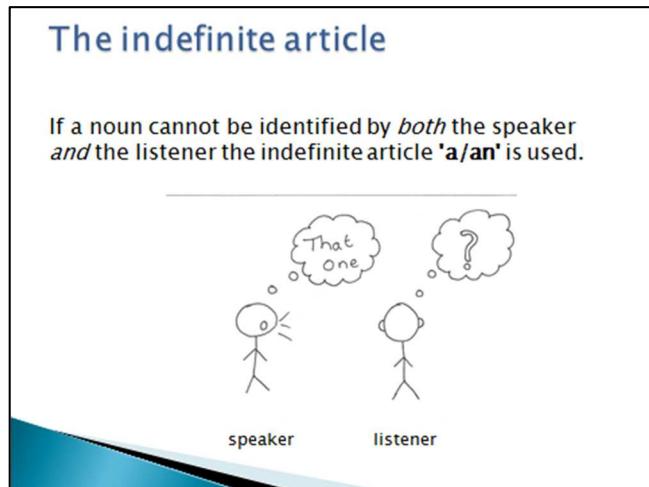
A lesson plan and materials were provided for three 90-minute grammar lessons. Lesson 1 focused on definiteness only, Lesson 2 contrasted definiteness and specificity, and Lesson 3 was an error correction lesson including examples of article errors.

4. Presentation of teaching materials

This section provides an overview of the materials used to present definiteness and specificity to the learners. In addition, a number of practice exercises were adapted from currently published teaching materials in order to make reference to specificity. As demonstrated above, definiteness was presented as shared knowledge between a speaker and listener and the definite/indefinite contrast formed the basis of the first 90-minute grammar lesson. Pictures were provided for both definite and indefinite uses of the article, as shown in Figures 1 and 2, respectively.

The use of thought bubbles in Figures 1 and 2 allows the learners and teachers to visualise the concept of ‘shared knowledge’.

¹⁰ Further details of the research project, as well as cross-linguistic facts about the role of specificity in Chinese, are not included here due to the length limitations of the paper.

FIGURE 1: SLIDE USED TO ILLUSTRATE USE OF THE DEFINITE ARTICLE**FIGURE 2:** SLIDE USED TO ILLUSTRATE USE OF THE INDEFINITE ARTICLE

Following these slides, the PowerPoint presentation goes on to provide example sentences, and was designed in an interactive manner so that teachers could engage with their students and maintain a communicative teaching style.

Specificity was taught during Lesson 2, and the materials were also designed to last approximately 90 minutes, including time for discussion. This concept was presented as 'speaker intent to refer' and, as specificity is not normally taught to L2 learners of English, it was necessary for teachers to ensure that their learners' understood the concept. Pictures and an example sentence for specific and non-specific contexts can be seen in Figures 3 and 4, respectively.

The difference between specific and non-specific reference was again visualised with the use of thought bubbles. In Figure 3, the specific sentence shows the speaker visualising a particular person, whereas in Figure 4, which is non-specific, the speaker has no such person in mind. As with definiteness, more example sentences were provided before students worked in groups to complete the practice exercises. Typical exercises included labelling sentences as either specific or non-specific, and recognising the difference between similar sentences on the basis of specificity.

FIGURE 3: SLIDE USED TO EXPLAIN SPECIFIC AND NON-SPECIFIC REFERENCE, WITH AN EXAMPLE OF A SPECIFIC SENTENCE

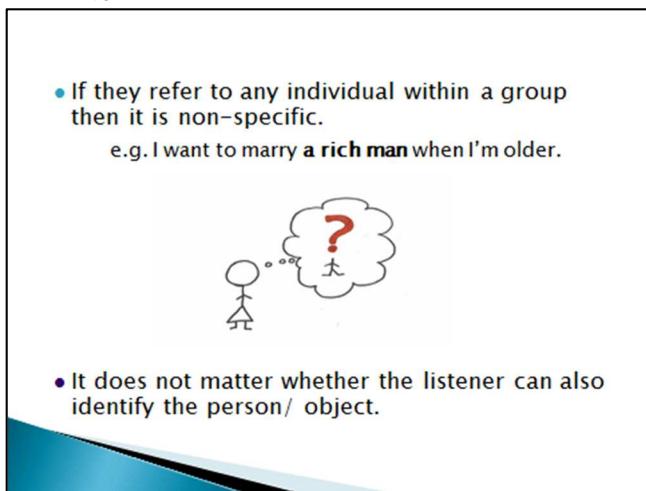
Specific and non-specific

- Specific and non-specific noun phrases are different depending on who/ what the speaker intends to refer to.
- If the speaker refers to one particular individual then it is specific.
e.g. I am meeting a **rich man** after work today.



At the end of the teaching period, none of the students or teachers reported any problems with these materials and, whilst the specificity lesson appeared to provoke more discussion amongst students, they were generally able to arrive at the correct answer and explain the reasons for their choice. A presentation of the results of the pre-test and post-test goes beyond the scope of this paper, but it is clear from the response of the teachers and learners that it is possible to teach specificity to L2 learners of English.

FIGURE 4: SLIDE USED TO ILLUSTRATE, AND PROVIDE AN EXAMPLE SENTENCE FOR, NON-SPECIFIC REFERENCE



6. Conclusion

As both an English language teacher and a generative linguist, it is my belief that GenSLA research *is* relevant to language teaching. This paper has demonstrated this using the example of the English article system. Furthermore, it has shown that it is possible to develop linguistically-informed teaching materials, and that the seemingly abstract property of specificity can be taught. There is no reason why there cannot be more collaboration between GenSLA researchers and teaching professionals in the future; however, for such collaboration to be successful I would recommend both groups avoiding the use of specialist terminology. Furthermore, to make GenSLA research relevant to teaching professionals, I believe that the focus should be on what research results say about potential sources of learner errors, which could then be targeted in the classroom with the aim of increasing learner accuracy.

References

- Bailey, S. (2006). *Academic writing: A handbook for international students*. (2nd ed.). London: Routledge.
- Cotton, D., Falvey, D. and Kent, D. (2008). *Language leader. Intermediate and upper-intermediate*. Harlow, Essex: Pearson Longman.
- Cunningham, S. and Moor, P. (2005; 2007). *Cutting edge*. Harlow, Essex: Pearson Longman.

- Hawkins, R., Al-Eid, S., Almahboob, I., Athanasopoulos, P., Chaengchenkit, R., Hu, J., Rezai, M., Jaensch, C., Jeon, Y., Jiang, A., Leung, Y.I., Matsunaga, K., Ortega, M., Sarko, G., Snape, N., and Velasco-Zarate, K. (2006). Accounting for English article interpretation by L2 speakers. *EUROSLA Yearbook*, 6, 7-25.
- Ionin, T. (2003). Article semantics in second language acquisition. (PhD dissertation). MIT.
- Ionin, T., Ko, H. and Wexler, K. (2004). Article semantics in L2 acquisition: The role of specificity. *Language Acquisition*, 12(1), 3-69.
- Ionin, T., Zubizarreta, M.L. and Bautista Maldonado, S. (2008). Sources of linguistic knowledge in the second language acquisition of English articles. *Lingua*, 118, 554-576.
- Lebeu, I. and Rees, G. (2008). *Language leader. Elementary and pre-intermediate*. Harlow, Essex: Pearson Longman.
- Lightbown, P. (2000). Anniversary article. Classroom SLA research and second language teaching. *Applied Linguistics*, 21(4), 431-462.
- Oxenden, C., Latham-Koenig, C. and Seligson, P. (2004; 2005; 2008; 2009). *New English file*. Oxford: Oxford University Press.
- Soars, L. and Soars, J. (2003; 2005; 2006). *New Headway English course*. (3rd ed.). Oxford: Oxford University Press.
- Tryzna, M. (2009). Questioning the validity of the Article Choice Parameter and Fluctuation Hypothesis: Evidence from L2 English article use by L1 Polish and L1 Mandarin Chinese Speakers. In M. Garcia Mayo & R. Hawkins (eds.), *Second language acquisition of articles* (pp. 67-86). Amsterdam: John Benjamins.

A Contrastive Rhetorical Study of Critical Essays in English by Tunisian EFL Students and American Students

Enneifer Salwa

L'Institut Supérieur des Langues de Tunis « ISLT ». Tunisia.

salwa2083@hotmail.fr

1. Topic area

The present study intends to examine the rhetorical structure of argumentative essays written by some American and Tunisian students in English. It seeks to investigate whether Tunisians, as native Arabic speakers, write different essays from those written by American native speakers of English in terms of rhetoric. The study also aims at identifying whether these differences, if they exist, stem from Tunisian students' first language. The researcher assumes that the transference of Arabic rhetorical structures, if it exists, detracts from the quality of the argument in English.

It is motivated by the prevalence attributed to contrastive rhetoric (CR) as an 'international' project aiming at building "an international databank available for other international researchers" (Connor, 1996: 173). It endeavors to test the validity of the contrastive rhetoric hypothesis; at least in the Tunisian context.

CR was first introduced by Robert Kaplan (1966). CR hypothesis claimed that non-native speakers of English write different texts than those produced by their native English speaking counterparts. The mastery of syntactic structures, lexicon or spelling cannot by themselves make an EFL student use English as a native English speaker does. The difference lies in the fact that non-native English speaking learners have internalised the rhetorical structures of their mother tongues. Therefore CR, as a theory, aims to study and identify these differences.

Some previous studies have dealt with the difficulties experienced by Arab EFL students when writing essays in English, such as Ouaouicha (1986) or Ismail (2010).

They showed the differences and/or similarities between Arab students' essays written in English and English native speakers' essays. For instance, Aouicha identified differences only in terms of audience awareness and emotional appeals for persuasion. Americans used both subjects aspects significantly more often than did Moroccans. Ismail (2010) also suggests that no significant differences exist between US and Arab advanced writers. He assumes that within-group variation is more important than between-group variation in the rhetorical performance of his participants.

Nevertheless, there is still a gap in the studies involving Tunisians as a particular population. No previous studies have studied Tunisians. Moreover, previous studies results cannot be generalised to the whole Arab population. The current research endeavors to fill this gap.

This article, then, is an investigation of the structure of critical essays written by some American and Tunisian students in English. It studies the rhetoric in essays by postgraduate Tunisian students of English written in Arabic and in English, in order to find out whether there is interference between the rhetorical structures of Arabic to English. Then, it compares the rhetoric of English critical essays written by postgraduate Tunisian students of English with American native speakers' English critical essays to identify similarities and differences. The current article starts by presenting the hypotheses tested, the methodology and approach used and it ends with the results and implications of the study.

2. Hypotheses

The hypotheses to be tested are:

- 1- Analytic variables selected to study critical essays in this study, namely:
argument superstructure, reasoning, persuasive appeals and persuasive adaptiveness are valid, as they can predict the holistic score.

The current study also hypothesized that, if the contrastive rhetoric claims presented first by Kaplan (1966) are correct, the analysis of essays in the current study will confirm that:

- 2- Significant differences exist between critical essays written by postgraduate Tunisian students of English and American native speaking students' essays written in English, in terms of selected analytic variables.
- 3- Postgraduate Tunisian students of English produce similar argumentative essays in English and Arabic, in terms of selected analytic variables.
- 4- Differences between critical essays written by postgraduate Tunisian students of English written in English and those written by American native speaking students in English are due to Arabic language rhetoric transfer.

3. Methodology

3.1. Corpus

The corpus for this study includes three sets of critical essays, thus: essays written in English by American native English speakers (AEE); essays written in English by postgraduate Tunisian students of English (TEE); and essays written in Arabic by postgraduate Tunisian students of English (TEA).

The researcher used the LOCNESS corpus for the American students' essays. LOCNESS is one of the available corpora of the 'Centre for English Corpus Linguistics' in the '*Université Catholique de Louvain*' by Sylvianne Granger. For this study, the researcher has selected only twenty seven essays, all written at Indiana University in Indianapolis. The latter are timed critical essays written in March 1995 by twenty seven English native speakers. The age of the writers ranged from 22 to 48. The topics of the essays were:

- 1- Money is the root of all evil
- 2- Crime does not pay
- 3- A man/woman's financial reward should be commensurate with their contribution to the society in which they live
- 4- Feminists have done more harm to the cause of women than good

The Tunisian essays are written by Tunisian students who were studying English at first year Masters' degree level at the 'Faculty of Letters and Humanities of Kairouan'

in Tunisia during the academic year 2011-12. Students at this particular level of education were chosen because it was assumed that they have less linguistic difficulties; such as grammatical errors, than undergraduate students. The essays were timed essays and the students answered one of the questions from the LOCNESS' Indiana University topic list already presented. Each Tunisian student was required to answer one of the four questions.

The choice of Indiana University essays among all the essays in LOCNESS is explained by the fact that these essays were timed essays. Also, the age of the students was the closest to the age of first year Masters' degree students. The American and Tunisian essays needed to share a maximum number of similarities in their characteristics in order to be comparable. The same assignments already presented were translated into Arabic and administered to the same postgraduate Tunisian students of English. Each was required to answer one of the four questions, in Arabic. In order to obtain an equally distributed corpus, containing equal numbers of both Tunisian and American students' essays, the Tunisian students sample also includes twenty seven essays written in Arabic and twenty seven essays written in English written by the same Tunisian students.

3.2. Approach

Ulla Connor (1990) presents a comprehensive model to assess persuasive writing. The model is multi-dimensional because it considers both linguistic and discourse-level features of essays and assesses a specific kind of discourse by developing analytical scales. In Connor's article (1990: 70), six variables are used to analyse persuasive student writing. Two variables are concerned with the study and analysis of syntax and coherence in students' critical essays. Both variables are not included in the following study, as it is basically and only concerned with the rhetorical differences between Tunisian and American students' critical essays.

Four variables of argumentative/persuasive texts are used to evaluate persuasive features in the sample essays, namely: the superstructure of the argument, the strength of logical reasoning involved; the use of persuasive appeals; and the adaptation of the

text to one's audience. The following table presents the independent variables and the scales used to describe them.

TABLE 1: INDEPENDENT VARIABLES AND SCALES FOR ANALYSING RHETORIC IN STUDENTS' CRITICAL ESSAYS

Independent variables	Scales
Argument Superstructure	Situation, problem, solution, evaluation
Reasoning	Claim, data, and warrant scale (Connor & Lauer, 1988)
Persuasive appeals	Rational, affective, and credibility appeals scale
Persuasive adaptiveness	Audience awareness and persuasive adaptiveness scale

For more details concerning the independent variables used in this study and presented in the table above, the reader can consult Connor (1990).

In addition to the independent variables presented in the above table, the corpus' essays are also holistically rated. The overall writing quality of each essay in this study is rated using a 0-5 point holistic scale.

3.3. Techniques used in corpus analysis

3.3.1. Rating essays

Both the holistic and analytic aspects of the corpus' essays were scored in this research by two independent raters. Once all the corpus' essays were rated, the researcher performed statistical tests in order to obtain quantitative results.

3.3.2. Quantitative analysis

All statistical analyses in the current study were performed via SPSS 19. In order to study the raters' agreement, an intra-class correlation coefficient was calculated to measure if there was an agreement between the scores each rater assigned to each essay, using each variable. Furthermore, a multinomial logistical regression was performed in order to study the relationship between the holistic score and analytic measures. The purpose of multinomial logistic regression analysis is to check the validity of the selected variables used to study critical essays. If the selected variables are valid, they can predict the overall writing quality and hence the holistic score. If the analytic variables are valid, the focus will be on testing the contrastive rhetoric hypothesis; or they can be used to identify whether significant rhetorical differences exist between postgraduate Tunisian students of English critical essays and American

native speaking students' essays written in English, using MANOVA as a statistical tool. If a difference was found, the researcher investigated whether it stemmed from the Tunisian students native language; that is Arabic.

4. Results

4.1. Results of intra-class correlation

The results of intra-class correlations revealed that there was a high level of agreement between both raters. The results of intra class correlation are summarised in table 2.

TABLE 2: INTERRATER AGREEMENT FOR ALL VARIABLES

Variable	Intra class correlation coefficient
Situation	1.00
Problem	1.00
Solution	.96
Evaluation	.81
Claim	.70
Data	.83
Warrant	.83
Rational appeals	.79
Credibility appeals	.83
Affective appeals	.83
Persuasive adaptiveness	.86
Overall mark	.84

As shown in the table the agreement rate between raters ranges from 70% to 100 % agreement. Hence, the researcher was confident that the scores she would use for later statistical analyses would reliably measure the rhetorical performance of the participants.

4.2. Results of multinomial logistic regression

In order to identify predictors of overall score, a multinomial logistic regression was made. It included four covariates or variables, namely: argument superstructure; added Toulmin; persuasive appeals; and persuasive adaptiveness. It should be mentioned here that the researcher calculated the mean of individual claim, data, and

warrant scores to obtain only one score or; an added Toulmin's informal reasoning score. The purpose is to facilitate later statistical analysis. The resulting added Toulmin score for each participant was a potential value of one to three.

The researcher also included grouping factors in the multinomial logistic regression model. The results of the regression analysis are summarized in table 3.

TABLE 3: MULTINOMIAL LOGISTIC REGRESSION

Model	Effect(s)	Model Fitting Criteria		Effect Selection Tests			Pseudo R-Square Nagelkerke
		AIC	-2 Log likelihood	Chi Square	df	Sig	
Step 0	Intercept	155.38	137.38	.			
Step 1	Group Added Toulmin	108.78	84.78	52.60	3	.000	
Step 2	Persuasive appeals	95.09	65.09	19.69	3	.000	
Step 3	Argument superstructure	93.04	57.04	8.06	3	.045	.82

As shown in Table 3, multinomial logistic regression indicates that the predictor variables of the overall writing scores were added Toulmin, persuasive appeals and argument superstructure. The pseudo R-Square equals 82%, which means that the four previously identified variables together can predict 82% of the variance in the holistic scores. The regression excluded persuasive adaptiveness as a predictor of overall scores.

Consequently, the added Toulmin score is the best predictor of the overall score. The second best predictor of overall mark is the persuasive appeals variable, which includes credibility appeals and affective appeals. The third predictor of the overall score is argument superstructure. It should be noted here that persuasive adaptiveness was dismissed as a predictor variable; as the regression excluded persuasive adaptiveness as a predictor of overall scores. Participants demonstrated a low performance on the persuasive adaptiveness scale, therefore cross-cultural variations of the use of persuasive adaptiveness in the corpus essays was insignificant for the current study.

When adding grouping factors into the model, the AIC improved. It reached 155.38 (Table 3) against 167.095 (Table 4). Therefore, grouping had an effect on the improvement of predictability of the score range. As shown in Table 5, the final model explains score range variations with a significance equaling $p < 0.05$.

TABLE 4: BEST PREDICTOR OF OVERALL MARK

Model	Likelihood Ratio Tests				
	AIC	-2 Log Likelihood	Chi-Square	df	Sig.
Intercept Only	167,095	161,095			
Final	93,038	57,038	104,057	15	,000

The variables added Toulmin, argument superstructure and persuasive appeals are significant at the 0 and 05 level, as shown in table 5. It implies that the latter variables may explain overall mark variations. Moreover, grouping factors have an effect on the overall score (with a P-value= 0.00), as shown in Table 5.

TABLE 5: LIKELIHOOD RATIO TESTS

Effect(s)	Likelihood Ratio Tests		
	Chi Square	Df	P-value
Intercept	.000	0	.
Group	26.002	6	.000
Added	54.860	3	.000
Toulmin			
Argument	8.056	3	.045
superstructure			
Persuasive	21.003	3	.000
appeals			

Table 6 presents the predicted overall scores according to their distribution. For instance, five essays were given a holistic score ranging between 1 and 2. Three out of the five essay scores were predicted via the model. According to multinomial logistic regression, two essays should be given a score ranging from 2 to 3. It also suggests that the scores of 60% of essays which obtained a holistic score, ranging from 1 to 2, are predicted by the four predictor variables, namely grouping, added Toulmin, persuasive appeals and argument superstructure. The predictability of essays which obtained a holistic score ranging from 2 to 3 ($n = 39$), equaled 84.6%. Essays which obtained a score ranging from 3 to 4 can be predicted at 71%. Finally, the essays

which scored between 4 and 5 are predictable at 66.7%. The four predictor variables achieved 76.5% correct overall percentage.

TABLE 6: CLASSIFICATION OF PREDICTED OVERALL SCORE RANGES

Observed	Predicted				Percent Correct
	[1,2[[2,3[[3,4[[4,5[
[1,2[3	2	0	0	60,0%
[2,3[2	33	4	0	84,6%
[3,4[0	8	22	1	71,0%
[4,5[0	0	2	4	66,7%
Overall Percentage					76,5%

4.3. Testing the contrastive rhetoric hypothesis

The second part of the statistical analysis aimed to identify whether significant rhetorical differences existed between critical essays written by postgraduate Tunisian students of English and essays written in English by American native speaking students. Therefore, it aimed to test the contrastive rhetoric hypothesis, and, in so doing, to answer research questions 2, 3 and 4, as previously presented.

In order to answer the following questions, MANOVA is used. MANOVA was conducted with holistic and analytic measures as the dependent variables, and group membership as the independent variable. The statistical comparative analysis began by checking whether either grouping essays, on one hand, or the categorization of essays in terms of TEA, TEE and AEE, on the other, had an impact on the analytic variables, namely: argument superstructure, added Toulmin, persuasive appeals and persuasive adaptiveness, as well as overall score.

TABLE 7: PAIRWISE COMPARISONS

Dependent Variable	(I) Groups	(J) Groups	Mean Difference (I-J)	Std. Error	Sig. ^a
Added Toulmin	TEA	TEE	-,086	,116	,459
		AEE	-,469*	,116	,000
	TEE	TEA	,086	,116	,459
		AEE	-,383*	,116	,001
	AEE	TEA	,469*	,116	,000
		TEE	,383*	,116	,001
Persuasive Appeals	TEA	TEE	-,130	,133	,333
		AEE	-,296*	,133	,029
	TEE	TEA	,130	,133	,333
		AEE	-,167	,133	,214
	AEE	TEA	,296*	,133	,029
		TEE	,167	,133	,214
Overall mark	TEA	TEE	,093	,155	,551
		AEE	-,648*	,155	,000
	TEE	TEA	-,093	,155	,551
		AEE	-,741*	,155	,000
	AEE	TEA	,648*	,155	,000
		TEE	,741*	,155	,000

Based on estimated marginal means

a. Adjustment for multiple comparisons: Least Significant Difference (equivalent to no adjustments).

**. The mean difference is significant at the, .05 level.*

A significant difference is signaled between TEE and TEA, on one hand, and AEE, on the other, in terms of added Toulmin. No significant difference exists between TEA and TEE in terms of added Toulmin ($p > 0.05$).

AEE and TEA have a significant difference in terms of persuasive appeals. Still, no significant differences exist between TEA and TEE ($p > 0.05$) or AEE and TEE ($p > 0.05$) in terms of persuasive appeals.

A significant difference exists between TEE and AEE in terms of overall score; ($p < 0.05$). Also, as shown in table 7, a significant difference exists between AEE and TEA ($p < 0.001$). However, no significant difference exists between TEA and TEE in terms of overall mark: ($p > 0.05$).

Through the previously stated findings, hypotheses two, three and four for the current study are validated. As a result, the current study assumes that:

- (1) Significant differences exist between the critical essays of postgraduate Tunisian students of English and American native speaking students' essays written in English, in terms of added Toulmin and overall score.
- (2) Postgraduate Tunisian students of English produced similar argumentative essays in English and Arabic in terms of added Toulmin, persuasive appeals and overall score.
- (3) A significant difference existed between TEA and AEE essays in terms of added Toulmin, persuasive appeals and overall score.
- (4) Tunisian students of English produced similar argumentative essays in English and Arabic in terms of added Toulmin and persuasive appeals, furthermore differences between critical essays written by postgraduate Tunisian students of English and those written in English by American native speaking students in English are also signalled in the use of added Toulmin and persuasive appeals, the researcher claims that the latter differences stem from Arabic language rhetoric transfer.

5. Implications

The implications for the field of contrastive rhetoric include the validity of the analytic model (Connor, 1990) selected to study rhetorical performance in critical essays. It is a valid tool of analysis for the investigation of various rhetorical aspects of Tunisian and American rhetorical structures and identifies their similarities and differences.

Furthermore, the current research found that both Tunisian and American groups displayed a relatively low performance on persuasive adaptiveness measures. One implication of this study is to enhance training on the effective use of persuasive adaptiveness.

The study revealed that no significant difference existed between TEE and AEE essays of the same Tunisian students. The same rhetorical techniques were found in both groups. Hence, the writing problems experienced by EFL Tunisian students are the byproduct of negative Arabic language transfer. Therefore, the researcher claims that contrastive rhetoric is a valuable resource. It may explain one of the reasons why ESL/EFL students produce different English texts in comparison to native English speakers. It may also locate the differences. Hence, course designers and teachers may use contrastive rhetoric results to improve their teaching methods and focus more on the difficulties experienced by their students. That way, teaching will satisfy the special needs of each population or discourse community.

References

- Connor, U. (1990). Linguistic/Rhetorical measures for international persuasive student writing. *Research in the Teaching of English*, 24(1), 67-87.
- Connor, U. (1996). *Contrastive rhetoric: Cross-cultural aspects of second language writing*. Cambridge: Cambridge University Press.
- Ismail, S. (2010). *Arabic and English persuasive writing of Arabs from a contrastive rhetoric perspective*. (Doctoral dissertation) Indiana University of Pennsylvania. The School of Graduate Studies and Research. Department of English.
- Kaplan, R. B. (1966). Cultural thought patterns in inter-cultural education. *Language Learning*, 16, 1- 20.
- Ouaouicha, D. (1986). *Contrastive rhetoric and the structure of learner-produced argumentative texts in Arabic and English*. (Doctoral dissertation). The University of Texas at Austin.

Dubbing can be an agent in pragmatic language change: Evidence from the use of *esatto* in Italian as an agreement marker

Lorella Viola
University of East Anglia
L.Viola@uea.ac.uk

1. Introduction

There exists a number of studies (i.e., Pavesi 2005; Alfieri *et al* 2003; Alfieri & Bonomi 2008) concerning Italian dubbing which claim interference phenomena derived from the audio-visual translation (AVT) process have affected the end users' language. However, no empirical evidence of such an influence has been provided so far and the alleged impact of dubbed Italian on spoken Italian relies on personal, anecdotal and/or circumstantial evidence.

In this paper, the results of the use of *esatto* (exactly) – considered as an AVT interference phenomenon when used as an agreement marker – are given. The results provide, for the first time, empirical evidence that Italian dubbing can be considered as an agent in language change at the pragmatic level.

The fact that interference is “an intrinsic factor in any translation” (Newmark, 1991: 78) is well established; in fact, studies in AVT (among the others Zaro 2001; Gottlieb 2001) have shown this form of translation to be even more vulnerable to interferences from the source text than other forms. Notoriously, linguistic interference can also take place at different levels (semantic, phonetic, morphological, syntactic). In AVT, the factor which paves the way specifically for pragmatic interference is the nature itself of film dialogues which try to imitate spontaneous spoken language (Chaume 2001). Pragmatic interference occurs in fact in day-to-day contact situations between different cultures and languages, whereby linguistic formulas that appear

similar in their meaning or structure are transferred from one language into the other (Thomas 1983). These transferred formulas are pragmatic interference phenomena.

The linguistic expressions claimed in previous studies to be interference phenomena from dubbing have been gathered and selected by applying the definition of pragmatic interference adopted throughout the study (see § 2). At the end of this first phase, 72 pragmatic interference phenomena were obtained and categorized in pragmatic markers, formulaic language, and fixed expressions. However, after running diachronic, lexicographic and quantitative analyses, only 54 items most likely to have been affected by pragmatic interference from dubbing have been identified. Finally, their presence and frequency have been searched within spoken Italian corpora (LIP, LABLITA, CLIPS) to ultimately assess whether spontaneous Italian has been affected by dubbing. In this paper, in particular, the results of the investigation of one expression (*esatto* - exactly) are given.

After giving a brief overview of the concept of linguistic interference within key fields, the definition of pragmatic interference adopted throughout this study is presented. Then, the stages of the study are described followed by an outline of the corpora used for the analysis. Finally, the results for *esatto* are presented and the conclusions discussed.

2. Pragmatic interference

The concept of pragmatic interference is derived from the classic definition of linguistic interference given by Weinreich (1953: 1) who described it as “Those instances of deviation from the norm of either language which occur in the speech of bilinguals as a result of their familiarity with more than one language, i.e. as a result of language contact”. In translation studies, interference is said to take place when “apparently inappropriately, any feature of the source or a third language [...] is carried over or literally translated as the case may be into the target language text.” (Newmark, 1991: 78). In translation, as well as in other forms of language contact situations, interference is assumed to take place in both directions (from L1 to L2 and *vice versa*).

Notoriously, linguistic interference can also take place at different levels (semantic, phonetic, morphological, syntactic). In this study, I consider linguistic interference from L2 to L1 from a pragmatic perspective as taking place during the AVT process and which from now onwards it will be referred to as *pragmatic interference*. Pragmatic interference is defined here in the sense adopted by Thomas (1983: 101):

The inappropriate transfer of speech act strategies from one language to another, or the transferring from the mother tongue to the target language of utterances which are semantically/syntactically equivalent, but which, because of different 'interpretive bias', tend to convey a different pragmatic force in the target language.

It has already been said that AVT, among all the forms of translation, has been proven to be particularly vulnerable to interferences from the source text. The factor which paves the way specifically for pragmatic interference is the nature itself of film dialogues which try to imitate spontaneous spoken language (Chaume 2001). Thus, in their attempt of recreating, indeed imitating spontaneous conversation in the dubbed script, dubbing translators resort to the linguistic tools typical of spoken language to confer vitality and authenticity to film dialogues. Being features of spoken language, these linguistic tools make film dialogues look unplanned and realistically spontaneous. At the same time, however, they make pragmatic interference highly likely to occur, for it takes place in dialogue interaction.

Translation in itself is well-known for being one of the most productive processes through which new features are brought into a language. What makes dubbing more powerful than other forms of translation is the channel through which it operates. While before the language contact situations between different cultures traditionally took place through books (accessible only to a minority of highly educated people), starting from the last century TV and cinema have become more and more easily accessible to everyone, thus eliminating any distinction between social classes. Moreover, both TV and cinema have become predominant in people's everyday life who have been exposed to them repetitively for about eight decades now and, in the case of TV, also for many hours a day. Considering that dubbing was first employed in Italy in 1932 and that today about 92% of AV products are dubbed (Paolinelli & Di

Fortunato, 2005) of which 80% are imported from English-speaking countries (Antonini, 2008: 135), Italian dubbing has become the most powerful and influential process through which interference phenomena between the Italian and the English culture have been brought into the Italian language.

3. Stages of the study

The analysis starts from the gathering of the expressions claimed in previous studies to be interference phenomena from dubbing (i.e., Alfieri *et al* 2003; Di Fortunato & Paolinelli 1996; Pavesi 2005). However, deriving from so many different studies, the expressions performed very different functions and have been defined in multiple ways. For instance, depending on the study, the phenomena are referred to as “stock translations” (Maraschio 1982), “translational routines” (Pavesi 1994, 2005), “phraseological calques” (Alfieri *et al* 2003; Alfieri & Bonomi 2008), or simply “dubbese” (Di Fortunato & Paolinelli 1996). For this reason, the expressions have been selected in the light of the definition of pragmatic interference adopted throughout (§ 2), thus providing a homogenous and consistent set of data (72 items) to work on.

It was similarly necessary to qualitatively analyse and categorise the relevant interference phenomena to identify more rigorously their range and type. The qualitative analysis helped disambiguate those expressions which performed different functions in different contexts and it made possible to differentiate between overlapping meanings.

Because film dialogues imitate spontaneous talk (Chaume 2001), the categories of pragmatic interference have been determined from the analysis of the features shared by dubbing and spoken language. These are: pragmatic markers (sub-categorised in metadiscourse markers, attention getters, modality markers), formulaic language (including forms of address, rituals, greetings and farewells), and fixed expressions (simple formulas, sayings, similes).

It has already been said how in previous studies the supposition that dubbing has an influence on spontaneous Italian remained limited to the personal perception of the

scholar; this also applies to the methodology used to identify the interference phenomena themselves. In other words, no historic linguistic analysis had been run to assess whether the alleged interference phenomena could be considered as expressions originated and/or used specifically in dubbing or, in any case, as the result of the interference from AVT.

Etymological, lexicographic and quantitative analyses of the pragmatic interference phenomena were carried out within grammars, dictionaries, and corpora to analyse the origin and the actual use of these expressions in Italian over the years. Corpora of written Italian (DiaCORIS and CORIS), in particular, have been used to perform a diachronic and quantitative analysis of the selected phenomena. From the initial list of 72 expressions, only 54 phenomena most likely to have been affected by pragmatic interference from dubbing were identified.

The last phase of the study is focussed on the analysis of these 54 phenomena throughout corpora of spoken Italian (LIP, LABLITA, CLIPS) to count their presence and frequency of occurrence so as to establish whether the initial hypothesis of the influence of dubbing on spoken Italian is correct.

4. The corpora

The CORIS Corpus (Corpus dell' Italiano Scritto – Corpus of written Italian) is made up of 130 million words and it collects authentic written texts from 1980 to 2011 chosen by virtue of their representativeness of modern Italian. The DiaCORIS (Diachronic CORIS) collects texts from 1861 to 2001 to monitor potential changes in the use of the phenomena analysed.

The LIP Corpus (De Mauro *et al* 1993) - Lessico di frequenza dell' italiano parlato (LIP Corpus – Frequency lexicon of Italian) was realised in 1990-1992 and contains a total of approximately 490,000 words.

The LABLITA Corpus as we refer to in this study is in fact the combination of the Stammerjohann Corpus (Stammerjohann 1970) and the C-ORAL-ROM (Cresti & Moneglia 2005) which together gather about 410,000 words.

The CLIPS (Corpora e Lessici dell'Italiano Parlato e Scritto – Corpora and Lexicons of written and spoken Italian) used in this study consists of about 550,000 words of oral speech collected between 1994 and 2004.

5. Results – *Esatto!* (exactly!)

We present here the results obtained from the analysis of *esatto* (exactly) as an agreement marker. After presenting the etymological and lexicographic searches, the results of the corpora queries are displayed in graphs in parts per million (p.p.m.). The calculation of the proportion is essential to understand the actual frequency of each item since the corpora used are different in size and proportional statistics are a better approach to present frequencies (McEnery & Wilson, 1996/2001: 82-83).

Esatto has been defined by Pavesi (2005: 51) as a translational routine of Italian dubbing. The scholar argues that more natural Italian expressions in similar contexts would be *sì*, *hai ragione* (yes, you're right), *sono d'accordo* (I agree) or similar.

Eco (1992: 169), on the contrary, affirmed that the spreading of the use of *esatto* in Italian has to be ascribed to TV and in particular to quiz shows of American origin. What is striking about Eco's claim is the fact that at the time of his writing (1992) the use of *esatto* as an agreement marker was perceived as "non Italian". This supports the hypothesis that such a use may be a relatively recent interference phenomenon.

Our etymological (DELI) and lexicographic search (TB, Crusca 1863-1923, GDLI, Garzanti 1965, Panzini 1905, Migliorini 1950) confirmed such a hypothesis since no results were found for *esatto* employed as an agreement marker while contemporary dictionaries (Devoto-Oli 2014, Treccani, GDI, Hoepli, DISC, Zing. 2008, DM) all report this specific use. In particular, the DM marks the use of *esatto* in answers as being colloquial which is particularly relevant to us. Since our study focusses on spoken Italian this finding already is, in itself, a confirmation of the use of *esatto* in spoken Italian.

To investigate the actual frequency of *esatto* as an agreement marker throughout time, we queried the DiaCORIS and the CORIS Corpus; the graphs below show the results.

FIGURE 1: ESATTO - DIACORIS RESULTS

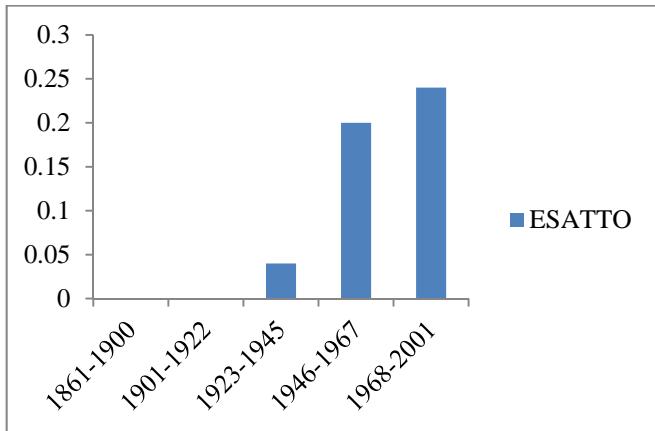
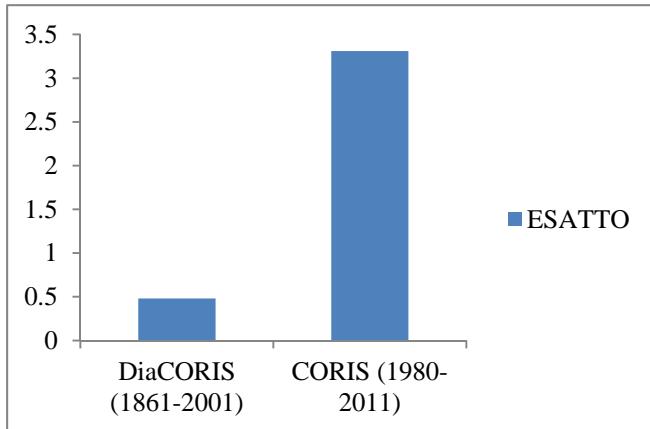
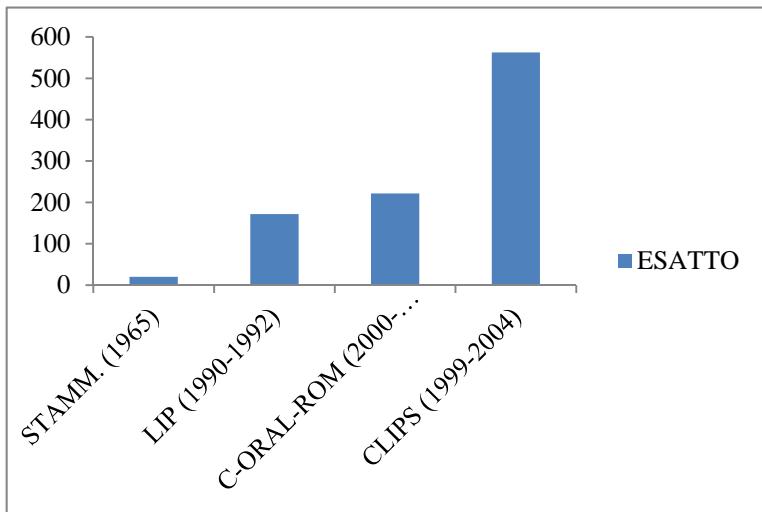


FIGURE 2: DIACORIS vs CORIS RESULTS



The graphs show how the use of *esatto* as an agreement marker has increasingly spread throughout time. To verify whether this trend is mirrored in the spoken level as well, its frequency has been analysed within spoken Italian corpora. The results are shown here below.

FIGURE 3: ESATTO - SPOKEN CORPORA RESULTS



The results amply confirm that *esatto* as an agreement marker has started to be used in modern times and then become more and more spread both at the written and the spoken level.

6. Conclusions

There exists a number of sociolinguistic studies (i.e., Muhr 2003; Carvalho 2004) which claim that exposure to media language affects language change. As far as the Italian language is concerned, the hypothesis that broadcasting mass media, and dubbing in particular, have an impact on spoken language has never been validated by any systematic evidence. In fact, previous claims, allegations and conclusions have been based on personal opinions.

This study introduces for the first time the use of empirical data to assess whether dubbing can be considered as an agent in language change. In this paper, in particular, *esatto* has been considered as a pragmatic interference phenomenon from dubbing when used as an agreement marker. The etymological, lexicographic and quantitative investigations have supported this hypothesis, for this device was not found in earlier stages of Italian while contemporary dictionaries list it.

Moreover, in the CORIS *esatto* as an agreement marker is far more frequent than in the DiaCORIS. These results mirror the searches within contemporary spoken Italian

corpora. Considering these findings, it is reasonable to infer that the repetition over time in dubbed films and programs of *esatto* in answers may have induced a change in spoken Italian.

It should be said, however, that this study does not aim to determine that dubbing is *the only* mechanism of language change, nor the exclusive origin of the entering of new expressions in the language (unlike previous studies). Rather, the aim was to assess that dubbing plays a role in these processes, also in terms of simply intensifying the use of such expressions, and to develop an empirical methodology to effectively contribute to the study of media-induced language change.

From our analysis, we conclude that dubbed Italian can be considered as having affected spontaneous Italian, that is as being an agent in pragmatic language change, as far as the use of *esatto* is concerned.

References

- Alfieri G., Contarino, S. and Motta, D. (2003). Interferenze fraseologiche nel doppiaggio televisivo: l’italiano di *E.R.* e *Beautiful*. In A. V., Sullam Calimani, (ed.), *Italiano e inglese a confronto: Atti del convegno Italiano e inglese a confronto: problemi di interferenza linguistica* (Venezia, 12-13 aprile 2002), *Quaderni della Rassegna* 33 (pp. 127-149). Firenze: Franco Cesati Editore.
- Alfieri, G. & Bonomi, I. (eds.). (2008). Gli italiani del piccolo schermo. Lingua e stili comunicativi nei generi televisivi. Firenze: Franco Cesati Editore.
- Antonini, R. (2008). The perception of dubbese. An Italian study. In D. Chiaro Nocella, C. Heiss, and C. Bucaria (eds.), *Between text and image: Updating research in screen translation* (pp. 135-148). Amsterdam/Philadelphia: Benjamins.
- Carvalho, A. M. (2004). I speak like the guys on TV: Palatalization and the urbanization of Uruguayan Portuguese. *Language Variation and Change* 16: 127–51.
- Di Fortunato, E. & Paolinelli, M. (eds.). (1996). Barriere linguistiche e circolazione delle opere audiovisive: la questione doppiaggio. Roma: AIDAC.
- Eco, U. (1992). *Il secondo diario minimo*. Milano: Bompiani.

- Gottlieb, H. (2001). Anglicisms and TV subtitles in an Anglified world. In Y. Gambier and H. Gottlieb (eds.), *(Multi)Media translation: Concepts, practices, and research* (pp. 249-258). Amsterdam: John Benjamins.
- McEnery, T. and Wilson, A. (1996/2001). *Corpus linguistics. An introduction*. Edinburgh: Edinburgh University Press.
- Muhr, R. (2003). Language change via satellite: The influence of German television broadcasting on Austrian German. *Journal of Historical Pragmatics*, 4, 103-27.
- Newmark, P. (1991). *About translation*. Clevedon, UK: Multilingual matters.
- Paolinelli, M. & Di Fortunato, E. (2005). *Tradurre per il doppiaggio. La trasposizione linguistica dell'audiovisivo: teoria e pratica di un'arte imperfetta*. Milano: Hoepli.
- Pavesi, M. (1994). Aspetti (socio)linguistici del doppiaggio. In R. Baccolini, R. M. Bollettieri Bosinelli, L. Gavioli (eds.), *Il doppiaggio: trasposizioni linguistiche e culturali* (pp. 129-142). Bologna: CLUEB.
- Pavesi, M. (2005). *La traduzione filmica. Aspetti del parlato doppiato dall'inglese all'italiano*. Roma: Carocci.
- Stuart-Smith, J. (2006). The influence of the media. In C. Llamas, L. Mullany, and P. Stockwell (eds.), *The Routledge companion to sociolinguistics* (pp. 140-148). London: Routledge.
- Thomas, J. (1983). Cross-cultural pragmatic failure. *Applied Linguistics*, 4(2), 91-112.
- Weinreich, U. (1953). *Languages in contact, findings and problems*. Linguistic Circle of New York: New York.
- Zaro, J. J. (2001). Conceptos traductológicos para el análisis del doblaje y la subtitulación. In M. Duro (ed.), *La traducción para el doblaje y la subtitulación* (pp. 47-63). Madrid: Cátedra.

Corpora

CLIPS (Corpora e Lessici di Italiano Parlato e Scritto). (n. d.). Retrieved from
<http://www.clips.unina.it/it/>

CORIS (Corpus di italiano scritto). (n. d.). Retrieved from
<http://corpora.dslo.unibo.it/TCORIS/>

- Cresti, E. & Moneglia, M. (eds.). (2005). *C-ORAL-ROM (Integrated Reference Corpora for Spoken Romance Languages)*. Amsterdam/Philadelphia: John Benjamins Publishing Company.
- De Mauro, T., Mancini, F., Vedovelli, M. and Voghera, M. (1993). LIP: Lessico di frequenza dell’italiano parlato, Milano: Etaslibri. Retrieved from <http://www.parlaritaliano.it/index.php/it/visualizza-corpus>
- DiaCORIS (Corpus diacronico di italiano scritto) (n.d.). Retrieved from <http://corpora.dslo.unibo.it/DiaCORIS/>
- Stammerjohann, H. (1970). Strukturen der Rede. Beobachtungen an der Umgangssprache von Florenz. *Studi di Filologia Italiana*, XXVIII, 295-397.

Dictionaries

- Battaglia, S. (1961-2002). *Grande dizionario della lingua italiana*. Torino: UTET = GDLI.
- Cannella, M. (ed.). (2007). *Lo Zingarelli 2008. Vocabolario della lingua italiana di Nicola Zingarelli con CD-ROM*. Bologna: Zanichelli = Zing. 2008.
- Cortelazzo, M. & Zolli, P. (eds.). (1999/2008). *Il nuovo etimologico. Dizionario etimologico della lingua italiana. Con CD-ROM*. Bologna: Zanichelli = DELI.
- Cusatelli, G. (ed.). (1965). *Dizionario Garzanti della lingua italiana*. Milano: Aldo Garzanti Editore = Garzanti 1965.
- De Mauro, T. (2000). *Dizionario della Lingua Italiana De Mauro*. Versione elettronica. Torino: Paravia = DM.
- Devoto, G. & Oli, G. C. (2013). *Il Devoto-Oli. Vocabolario della lingua italiana 2014. Versione elettronica*. Serianni, Luca and Trifone, Maurizio (eds.). Firenze: Le Monnier = Devoto-Oli 2014.
- Gabrielli, A. (2011). *Grande Dizionario Italiano - Speciale 150 anni*. Milano: Hoepli = Hoepli. Retrieved from www.hoepli.it.
- Grande dizionario di Italiano. (2013). Milano: Garzanti Linguistica = GDI. Retrieved from www.garzantilinguistica.it.
- Migliorini, B. (1950). Appendice al Dizionario moderno. In A. Panzini, *Dizionario moderno (nona edizione)* (pp. 765-997). Milano: Hoepli.
- Panzini, A. (1905). *Dizionario Moderno delle parole che non si trovano nei dizionari comuni*. Milano: Hoepli.

- Sabatini, F. & Coletti, V. (2008). *Il Sabatini Coletti. Dizionario della Lingua Italiana.* 2007. Milano: Rizzoli Larousse = DISC. Retrieved from www.dizionari.corriere.it.
- Tommaseo, N. & Bellini, B. (1861-1879). Dizionario della lingua italiana con oltre centomila giunte a precedenti dizionari. In M. Cannella (ed.), *Lo Zingarelli 2008. Vocabolario della lingua italiana di Nicola Zingarelli con CD-ROM.* 2007. Bologna: Zanichelli = TB.
- Vocabolario Degli Accademici Della Crusca. Quinta edizione 1863-1923. Firenze: Tipografia Galileiana = Crusca 1863-1923.
- Vocabolario Treccani. (n.d.). Roma: Istituto della Enciclopedia Italiana = Treccani. Retrieved from www.treccani.it/vocabolario/.

@Sum1, #Everyone: A Corpus-Based Analysis of Language Performances and Attitudes on Twitter

Matthew Voice

University of Sheffield, School of English

mavoice1@sheffield.ac.uk

1. Introduction

As a website, Twitter exemplifies the complex discursive relationships intersecting professionalism, commercialism, and personal social interaction online today. While computer-mediated communication (CMC) has been studied considerably, its ephemeral nature requires continual reanalysis of the features in regular use, as well as their contexts of production. Through explicit comparison to earlier studies of internet language (e.g. Crystal, 2001), this paper provides a contemporary account of the presence and nature of CMC-specific features, before working toward an understanding of the modern contexts in which such language appears, updating and preserving findings in a rapidly shifting linguistic environment.

To more accurately describe CMC-specific discourse, the data for this study has been taken from the Illocution Inc. 2012 Twitter corpus, which is composed of one million tweets from across 2012, thus allowing for an estimation of linguistic patterns across Twitter's English-speaking user-base through the range and quantity of its contents. While corpus linguistics as a research method allows efficient processing of large bodies of data, it has been observed that 'many corpus linguists are still very much concerned with issues such as representativeness, structure, balance, documentation, and replicability' (Hundt et al., 2007: 4). Though substantial, Illocution Inc. contains 0.002% of tweets sent each day in 2012 (Twitter, 2012). That said, it remains as reasonable a representation of the site as can presently be processed.

2. Defining “Netspeak”

Logically, CMC-specific features must be defined before the corpus’s data can be compared to non-internet language. However, the components of CMC are ‘not a choice among a repertoire of fixed choices made available through a specific language code, but rather an unfixable constantly mutating act of partial intersubjectivity’ (Gillen and Merchant, 2013: 57), resisting a definitive list of properties, as “netspeak” and its constituent structures change dramatically over relatively short spaces of time. This is best demonstrated by considering Crystal’s 2001 collection of some of the ‘commonest’ features of CMC-language (2001: 84-6). Though Crystal’s collection limits itself to acronyms and initialisms, searching for these terms in the 2012 Twitter corpus demonstrates a marked shift away from many of these specific phrases. Only 36 out of 109 occur within the corpus, and several of the most productive (‘ASAP’, ‘FYI’, and ‘RIP’) predate their CMC-usage, originating in 1954, 1941, and 1613 respectively (OED). While they belong to the same lexical category as many CMC-specific terms, their wider acceptance and use necessitates classification in a separate register. This leaves nine terms found more than 100 times in the million-tweet resource, emphasising the need for continual reassessment of “netspeak” as category of language.

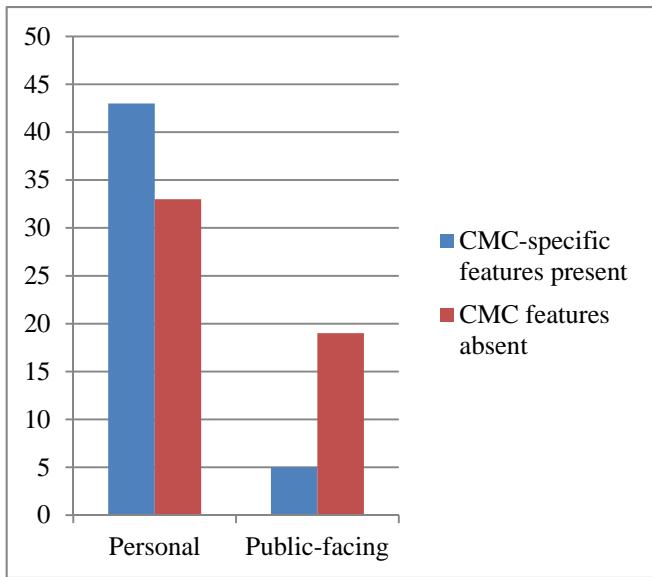
While this provides an insightful starting point, the 2012 Twitter corpus can be explored more effectively through manual comparison. A wordlist of the most frequent non-standard features in the corpus not present in Crystal’s list (Appendix), illuminates CMC’s ‘constantly mutating’ nature. Though indicating productivity beyond acronymy, the number of frequent terms is still notably lower than those catalogued by Crystal. While Twitter’s character-limit requires an ‘intuitive response to a technological problem’ comparable to Crystal’s (2008: 69) observation of non-standard English in texting, the decrease in non-standard features rebuffs ‘human ludic temperament’ (71) as a sufficient explanation of linguistic creativity. Indeed, although Crystal argues that ‘internet users are continually searching for vocabulary to describe their experiences [...] and to overcome the communicative limitation of its technology’ (2001: 67), the physical limitations alone cannot account for productivity, or lack thereof, in online language variation. Social attitudes toward language change

are also a variable in CMC analysis: as Squires puts it, ‘categorisation [of CMC-specific language] is based largely on what features are enregistered as belonging to... Standard English, and *which social settings it is perceived as appropriate to*’ (2010: 483, my emphasis). My definition of CMC-specific features encompasses non-standard features of English employed for spatial or emotive efficiency, such as abbreviations and pictographs. However, their limited presence, compared to similarly restricted media such as texting, necessitates an exploration into additional variables influencing language choices on Twitter.

3. Language Production and Twitter-specific Contexts

3.1. Public/Personal

Considering the nature and purpose of communication on Twitter, Marwick and boyd conclude that all users ‘must formulate tweets... based on an imagined audience judgement’ (2011: 124), thereby implying a conscious identity performance between likeability and authenticity. In other words, consideration of the audience(s) is central to communication on Twitter, and thus an analysis of its construction. While all tweets reach an audience of the account’s followers, some messages are targeted at specific user, by prefixing ‘@’ to their username. If used in reply to another message, tweets marked in this way will be automatically structured into a linear, dialogic presentational format. Thus, by searching the corpus for tweets containing ‘@’, it was possible to develop broad thematic categories of discourse, indicated through a common, searchable linguistic feature (fig. 1).

FIGURE 1: CMC-SPECIFIC FEATURES IN ‘@’ TWEETS

From a sample of 100 tweets containing ‘@’ within the 2012 corpus, the graph shows CMC-specific features in 60% of one-on-one discursive tweets. Additionally, the ‘public-facing’ category accounts for the fact that ‘@’ can also address a broad audience. For example:

I really enjoyed spending the day with my gorgeous @username1

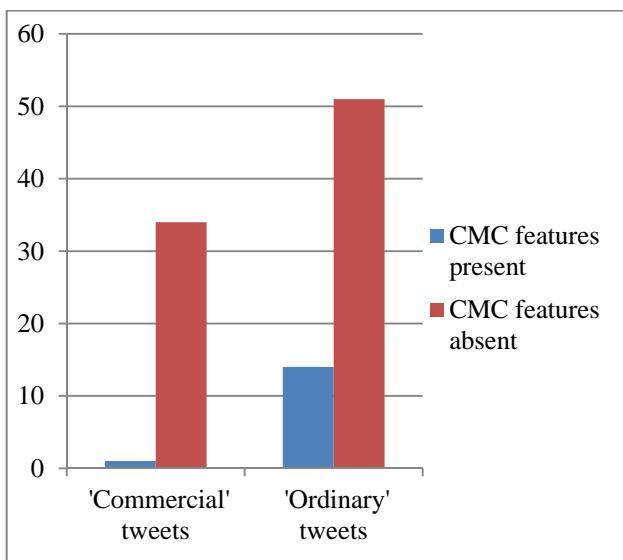
While ‘@’ is present, its object-positioning within the tweet’s grammatical structure does not make it the primary addressee. In such cases, identification of the tweet’s discursive function relies on interpretation at the analysis stage, as opposed to the inclusion of a specific feature. Such a distinction is valuable, however, as fig. 1 suggests non-standard features appear in only 20% of public-facing ‘@’ tweets; a significant contrast warranting further investigation into the distinction between public and personally oriented tweets. Thus a study was developed using the hashtag (#) as a second Twitter-specific feature functionally indicative of a tweet’s audience. Unlike, ‘@’, the hashtag marker is designed to ‘increase the ‘loudness’ of [a user’s] discourse by increasing the likelihood that their words will be found’ (Zappavigna, 2011: 800), and recording of ‘#’ alongside standard and/or CMC-specific English features is intended to expand on findings in public-facing ‘@’ tweets, as the inclusion of hashtags makes even ostensibly one-on-one communication more searchable within the wider Twitter user-base. Essentially, it aims to capture

language attitudes considering the ‘presuppose[d] virtual community of interested listeners’ (803) suggested by non-personal ‘@’ findings.

3.2. Commercial/‘Ordinary’

Before this investigation’s results can be considered, an important distinction within its findings must be explained. Page (2012) has highlighted hashtags in corporate/celebrity Twitter use as commercially driven, and separating these from ‘ordinary’ tweets through manual processing will allow for the influence and extent of corporate hashtag use at the analysis stage. Though the anonymity of the corpus makes identification of specifically corporate authors challenging, classification can be made at the level of content, described by Page as ‘declarative forms or imperatives that seek in turn to persuade the addressed audience to engage with the promoted product’ (198). As fig. 2 shows, highlighting this form within hashtags tweets provides an enlightening comparison:

FIGURE 2: FREQUENCIES OF CMC-SPECIFIC FEATURES IN ‘#’ TWEETS



Using the same definition of CMC-specific features as in the ‘@’ tweet collection, this graph reveals reduced non-standard language in explicitly public and searchable discourse. Moreover, the ‘commercial’/‘ordinary’ distinction demonstrates this change is not an anomaly caused by corporate presence, but a distinct change in style made by regular site users.

3.3. Discussion

What these studies appear to show is a tendency toward Standard English (SE) by the majority of Twitter users when their discourse is emphatically public-facing. With commercial tweets, these findings are unsurprising: hashtagging behaviour ‘is rooted in, and reinforces, offline asymmetries of economic power and status’ (Page: 193), and SE is equally enregistered as indicative of professionalism online and off. That said, fig. 2 does show one instance of CMC-specific language in a commercial tweet:

RT @JustinBieber: Holidays are about #GIVINGBACK. @NookBN is
doubling your #Schools4All donations & u can meet me at your school!

Here, the commercial content comes via a celebrity with a young demographic, associated with flexible language likely to incorporate CMC-specific features (cf. Bucholz, 2000). Hence, use of ‘u’ signifies an adaptation to the linguistic standards of the tweeter’s (and encoded advertisement’s) target audience, demonstrating the Community Norm Adherence principle: ‘the extent to which an individual adheres to the norms that have evolved in the use of the medium’ (Fischer and Reuber, 2011: 2). A marketing strategy, CNA embraces language models that authors associate with their intended audience, and has been shown to positively affect discourse around individuals and companies who employ it (Wigley and Lewis, 2012). However, the infrequency of CMC-features in commercial discourse samples indicates a more common form of CNA: conformity to forms enregistered as Standard English.

Pressure to follow existing linguistic standards influences not just commercial language, but also the ways everyday users of Twitter present themselves to their audiences. According to Page, ‘one-to-many broadcasts... are the primary site of self-branding via hashtags’ (187) for all users; a contrast in frequency of non-standard language features between personal and expressly public communication may be indicative of a variety of social and performative factors. For instance:

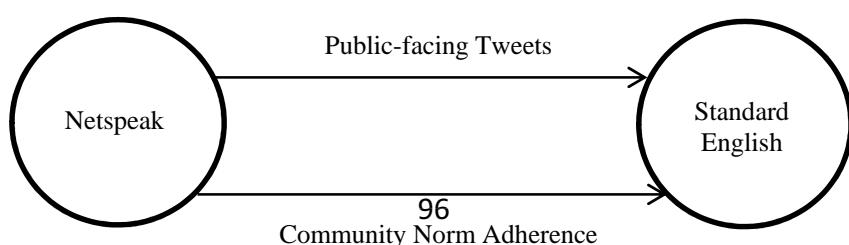
Chelsea have to be commended for their spirit of togetherness. #CFC

From an ‘ordinary’ user provides an evaluative commentary within an established sports-related debate (around Chelsea football club), making use of a standardised

hashtag ('#CFC') employed by professional and amateur commentators alike. Ostensibly equal, as a result of the perceived 'asymmetries of economic power', users may feel their writing must conform and perform in the same way as language developed from offline journalism and commentary. Hence, there exists a variation of Community Norm Adherence, in which 'ordinary' users aspire to commercial and professional Twitter styles to contribute to widely-viewed discourses.

Just as users may be inclined to adopt a particular register based on the established modalities of prominent on- and offline writing with significant cultural/economic capital, they may be influenced by popular attitudes of other 'ordinary' users. Returning to the dialogic '@' tweets, established linguistic features 'can determine others' responsive positions under the complex conditions of speech communication in a particular cultural sphere' (Bakhtin, in Gillen and Merchant: 57). Essentially, a participant within a cultural sphere already inclined toward non-standard English features will be more willing to employ these when others within the sphere are perceived to be the tweet's audience (i.e. some users of informal, personal '@' tweets). To borrow Goffman's metaphor of an actor adopting a role, when a user engages with a social group, 'usually he [sic] finds that a particular front [a set of linguistic and discursive norms] has already been established' (1969: 37). In circles already inclined toward it, CMC-specific language remains a permissible front-feature reinforced by its recurrence in particular personal discourse. However, informal deviation from SE is still exceptional, as the processes permitting such language use only occur in some '@'-personal tweets. According to Marwick and boyd, the unknowable audience 'is often imagined as its most sensitive members: parents, partners, and bosses', meaning users, 'may frame Twitter as a place where the strictest standards apply' (125-6). Alongside commercial influence, offline perceptions of language use take precedent in CMC, at least on Twitter.

FIGURE 3: VISUALISING THE PROCESSES OF PERFORMING/TRANSFORMING LINGUISTIC REGISTERS.



Active avoidance of non-standard, CMC-specific linguistic features is hardly a recent phenomenon. Regarding early internet writing, Elmer-Dewitt (in Cummings, 1995: 7) suggests that it ‘can be very bad indeed: sloppy, meandering, puerile, ungrammatical, poorly spelled... a Darwinian survival principle has started to prevail... good writing on the Net tends to be clear, vigorous, witty, and above all brief’. This attitude discourages deviation from existing, offline notions of Standard English, conflating elliptical respellings with failure to communicate. While this might be dismissed as early resistance to the emergence of a new form, similar views persist in contemporary analysis of CMC language. For instance, van der Laan argues, ‘when we speak and write in accord with technological principles and values, we conform to technology and, along with our language, transform into mechanisms’ (2012: 251), explicitly enregistering CMC-driven innovations as being negative to, hence belonging outside, Standard English. Typical acronym-like features aside, other innovations documented within the corpus, such as ‘tho’, ‘bout’, ‘Ima’ and ‘tryna’ (Appendix), associate CMC linguistic productivity with an ‘informalization of discourse’ (Montero-Fleta et al., 2009: 771) through orality, counter to the aims of commercial users. The consistency of standardised usage in (most) consciously public tweets suggests that English registers are distinguished ‘by social function, rather than modality’ (Squires: 477), and the continued expression of offline-rooted, professional language-values impact significantly on the practices of the majority of publicly-orientated users.

3.4. Further Features

Finally, the 2012 Twitter corpus was parsed through WordSmith’s ‘Keywords’, against a representative million-word sample of the British National Corpus (BNC), in order to calculate the ‘keyness’ of words used on Twitter, where a greater value indicates a higher usage frequency in comparison to the second corpus. As CMC-specific words will, by definition, occur with greater frequency in the Twitter corpus, the table below represents the Standard English words with greatest keyness against the BNC sample.

FIGURE 4: TOP 5 STANDARD ENGLISH ‘KEY’ WORDS IN 2012 TWITTER CORPUS

Rank	Word	Keyness
1	My	2,449.73
2	I	2,045.94
3	Me	1,774.04
4	Love	1,198.50
5	I'm	1,063.01

After this, Standard English words with significant keyness tend to be expletives not usually socially permissible in many of the registers recorded in the BNC sample. Such a finding indicates comparative informality within some uses of Twitter, though less pronounced than the keyness of all first person singular pronouns (genitive ‘I’ve’ is 7th). The dramatic keyness of all its forms highlights author-as-subject as a sustained and distinctive stylistic feature of tweeting. This style, identified as indicative of narcissism by Twitter users themselves (Qiu et al. 2012; Panek et al. 2013) can be found throughout personal and publicly-orientated tweets, reinforcing the value of linguistic regulation within users’ performance of self to a variety of audiences.

4. Conclusions

While its forms have changed drastically since Crystal (2001), perceptions of CMC-features as outside SE persist; its usage demonstrably limited to particular imagined audiences. By identifying broad categories of personal/public and commercial/everyday performances on Twitter, the influence of offline ideals of language are made clearer. This demonstrates the challenges in conceptualising a single “language of Twitter”, let alone CMC-specific language more generally.

Twitter represents only a fragment of the internet, and the frequency of CMC-specific features likely varies dramatically across contexts. Just as significant changes in internet language have been recorded since Crystal (2001), so they are likely to continue changing, along with personal and global attitudes toward Standard English and its classification. This paper has aimed to synchronically document this relationship in a particular online setting, in a style accessible for future research. Comparison of these results to other on- or offline corpora may elucidate the

relationship between public/private and commercial/ordinary discursive practices in other contexts, while a review of Twitter at a later date may incorporate these findings into a diachronic commentary on shifting digital language performances.

Appendix

Non-standard language features from 2012 Twitter corpus through WordSmith's wordlist:

Phrase	Occurrences
LMAO	7,437
OMG	6,893
Ur	4,970
SMH	4,789
Tho	4,478
Bout	4,277
WTF	2,929
Ima	2,707
LMFAO	2,524
PPL	2,459
Tryna	1,636

Bibliography

- 2012, *Twitter SRS, English, 1,000,000 Tweets*. Illocutioninc.com, 2013.
[<http://www.illocutioninc.com/Corpora/>](http://www.illocutioninc.com/Corpora/) Accessed 14th November 2013.
- ‘asap, adv. (and adj.)’ (2013). *OED Online*.
 [<http://www.oed.com.eresources.shef.ac.uk/view/Entry/306110?redirectedFrom=asap>](http://www.oed.com.eresources.shef.ac.uk/view/Entry/306110?redirectedFrom=asap). Oxford: Oxford University Press. Accessed 24th November 2013.
- The BNC Sampler*, XML version. (2005). Distributed by Oxford University Computing Services on behalf of the BNC Consortium.
 [<http://www.natcorp.ox.ac.uk/>](http://www.natcorp.ox.ac.uk/) Accessed 17th November 2013.
- Bucholz, M. (2000). Language and youth culture, *American Speech* 75(3): 280-283.
- Crystal, D. (2001) *Language and the Internet*. Cambridge: Cambridge University Press.
- Crystal, D. (2008). *Txtng: The Gr8 Db8*. Oxford: Oxford University Press.
- Cumming, J.D. (1995). The Internet and the English language, *English Today*, 11(1): 3-8.

- Fischer, E., and Reuber, A.R. (2011) Social interaction via new social media: (How) Can interactions on Twitter affect effectual thinking and behaviour?, *Journal of Business Venturing* 26(1): 1-18.
- ‘FYI, phr., adj., and n.’ (2013). *OED Online*.
<http://www.oed.com.eresources.shef.ac.uk/view/Entry/248803?redirectedFrom=yi>. Oxford: Oxford University Press. Accessed 24th November 2013.
- Gillen, J., and Merchant, G. (2013). Contact calls: Twitter as dialogic and social linguistic practice, *Language Sciences* 35, 47-58.
- Goffman, E. (1969). *The presentation of self in everyday life*. London: Allen Lane.
- Hundt M., N. Nesselhauf, and C. Biewer. (2007). Corpus linguistics and the Web. In M. Hundt, N.Nesselhauf, and C. Biewer (eds) *Corpus linguistics and the Web*, pp. 1-6. Amsterdam: Rodopi.
- Marwick, A.E., and boyd, d. (2011). ‘I tweet honestly, I tweet passionately: Twitter users, context collapse, and the imagined audience, *New Media & Society* 1(1): 114-133.
- Montero-Fleta, B., A. Montesinos-Lopez, C. Perez-Sabater, and E. Turney. (2009). ‘Computer mediated communication and the informalization of discourse: The influence of culture and subject matter, *Journal of Pragmatics* 41(4), 770-779.
- Page, R. (2012). ‘The linguistics of self-branding and micro-celebrity in Twitter: The role of hashtags’, *Discourse and Communication* 6(2), 181-201.
- Panek, E.T., Y. Nardis, and S. Konrath. (2013). ‘Mirror or megaphone?: How relationships between narcissism and social networking use differ between Facebook and Twitter, *Computers in Human Behaviour* 29(5), 2004-2012.
- Qiu, L., H. Lin, J. Ramsay, and F. Yang. (2012). You are what you tweet: Personality expression and perception on Twitter, *Journal of Research in Personality* 46(6), 710-718.
- ‘RIP, int. and v.3’. (2013). *OED Online*.
<http://www.oed.com.eresources.shef.ac.uk/view/Entry/270781?rskey=qaMIXy&rезультат=10&isAdvanced=false>. Oxford: Oxford University Press. Accessed 24th November 2013.
- Scott, M. (2004). *WordSmith Tools version 4*, Oxford: Oxford University Press.
- Squires, L. (2010). Enregistering Internet language, *Language in Society* 39(4): 457-492.

Twitter Turns Six. (2012). <<https://blog.twitter.com/2012/twitter-turns-six>> Accessed 20th November 2013.

Van der Laan, J.M. (2012). Language and being human in technology, *Bulletin of Science Technology* 32(3), 241-252.

Wigley, S., and Lewis, B. (2012). Rules of engagement: Practice what you tweet, *Public Relations Review* 38(1): 165-167.

Zappavigna, M. (2011). Ambient affiliation: A linguistic perspective on Twitter, *New Media & Society* 13(5): 788-806.